WHO Technical Report Series

1031

The selection and use of essential in vitro diagnostics

Report of the third meeting of the WHO Strategic Advisory Group of Experts on In Vitro Diagnostics, 2020
(including the third WHO model list of essential in vitro diagnostics)

World Health Organization

The World Health Organization was established in 1948 as a specialized agency of the United Nations serving as the directing and coordinating authority for international health matters and public health. One of WHO's constitutional functions is to provide objective and reliable information and advice in the field of human health, a responsibility that it fulfils in part through its extensive programme of publications

The Organization seeks through its publications to support national health strategies and address the most pressing public health concerns of populations around the world. To respond to the needs of Member States at all levels of development, WHO publishes practical manuals, handbooks and training material for specific categories of health workers; internationally applicable guidelines and standards; reviews and analyses of health policies, programmes and research; and state-of-the-art consensus reports that offer technical advice and recommendations for decision-makers. These books are closely tied to the Organization's priority activities, encompassing disease prevention and control, the development of equitable health systems based on primary health care, and health promotion for individuals and communities. Progress towards better health for all also demands the global dissemination and exchange of information that draws on the knowledge and experience of all WHO's Member countries and the collaboration of world leaders in public health and the biomedical sciences.

To ensure the widest possible availability of authoritative information and guidance on health matters, WHO secures the broad international distribution of its publications and encourages their translation and adaptation. By helping to promote and protect health and prevent and control disease throughout the world, WHO's books contribute to achieving the Organization's principal objective – the attainment by all people of the highest possible level of health.

The WHO *Technical Report Series* makes available the findings of various international groups of experts that provide WHO with the latest scientific and technical advice on a broad range of medical and public health subjects. Members of such expert groups serve without remuneration in their personal capacities rather than as representatives of governments or other bodies; their views do not necessarily reflect the decisions or the stated policy of WHO.

For further information, please contact: WHO Press, World Health Organization, 20 avenue Appia, 1211 Geneva 27, Switzerland (tel. +41 22 791 3264; fax: +41 22 791 4857; email: bookorders@who.int; order on line: www.who.int/bookorders).

WHO Technical Report Series
1031

The selection and use of essential in vitro diagnostics

Report of the third meeting of the WHO Strategic Advisory Group of Experts on In Vitro Diagnostics, 2020
(including the third WHO model list of essential in vitro diagnostics)

World Health Organization

The selection and use of essential in vitro diagnostics: report of the third meeting of the WHO Strategic Advisory Group of Experts on In Vitro Diagnostics, 2020 (including the third WHO model list of essential in vitro diagnostics).
(WHO Technical Report Series, no. 1031)
ISBN 978-92-4-001910-2 (electronic version)
ISSN 978-92-4-001911-9 (print version)
ISSN 0512-3054

Suggested citation. The selection and use of essential in vitro diagnostics: report of the third meeting of the WHO Strategic Advisory Group of Experts on In Vitro Diagnostics, 2020 (including the third WHO model list of essential in vitro diagnostics). Geneva: World Health Organization; 2021 (WHO Technical Report Series, No. 1031). Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at http://apps.who.int/iris.

Sales, rights and licensing. To purchase WHO publications, see http://apps.who.int/bookorders. To submit requests for commercial use and queries on rights and licensing, see http://www.who.int/about/licensing.

Printed in Switzerland

Contents

Abbreviations vii

List of participants xiii

Declarations of interests xvi

Acknowledgements xvii

Introduction 1

1. EDL and the review process 2

1.1. The EDL so far 2
1.2. Methods used to develop EDL 3
1.2.1. Changes to submission forms 4
1.2.2. 2020 remote process for deliberations 4

2. EDL-related products and processes 7

2.1. Progress updates 7
2.1.1. eEDL 7
2.1.2. Country guidance 8
2.1.3. EDL technical specifications 8
2.1.4. Harmonization activities 9
2.2. Prioritization of the EDL 10

3. Applications for additions to the EDL 13

EDL Section I.a: General IVDs for community settings and health facilities without laboratories 14
3.1. Erythrocyte sedimentation rate 14
EDL Section I.b: Disease-specific IVDs for community settings and health facilities without laboratories 21
CORONAVIRUS DISEASE 21
3.2. SARS-CoV-2 antigen 21
SICKLING DISORDERS 28
3.3. Sickle cell testing 28
STREPTOCOCCAL PHARYNGITIS 36
3.4. Group A *Streptococcus* antigen 36
EDL Section II.a: general IVDs for use in clinical laboratories 42
HAEMATOLOGY 42
3.5. Cerebrospinal fluid profile: leukocyte count, differential leukocyte count, protein and glucose 42
EDL Section II.b: Disease-specific IVDs for clinical laboratories 55
ASPERGILLOSIS 55
3.6. *Aspergillus* antigen 55
3.7. *Aspergillus* immunoglobulin G antibody 62
CANCER 67
3.8. Serum and urine protein electrophoresis for multiple myeloma 67
3.9. Epidermal growth factor receptor gene mutation 72

CORONAVIRUS DISEASE 81
3.10. SARS-CoV-2 nucleic acid test 81
ENDOCRINE DISORDERS 89
3.11. Cortisol (total) 89
3.12. Estradiol 106
3.13. Follicle-stimulating hormone 116
3.14. Luteinizing hormone 128
3.15. Progesterone 139
3.16. Prolactin 148
NEGLECTED TROPICAL DISEASES 160
3.17. *Trypanosoma* immunoglobulin G antibody 160
3.18. Visceral leishmaniasis direct agglutination test 166
PNEUMOCYSTIS PNEUMONIA 174
3.19. *Pneumocystis jirovecii* nucleic acid test 174
PRIMARY IMMUNODEFICIENCIES 182
3.20. Serum and urine protein electrophoresis for primary immunodeficiency 182
STREPTOCOCCAL PHARYNGITIS 190
3.21. Group A *Streptococcus* nucleic acid test 190
VACCINE-PREVENTABLE DISEASES 195
3.22. Measles immunoglobulin G antibody 195
3.23. Measles immunoglobulin M antibody 201
3.24. Measles nucleic acid test 207
3.25. Rubella immunoglobulin G antibody 213
3.26. Rubella immunoglobulin M antibody 219

4. Applications for revisions to the EDL 226

CLINICAL CHEMISTRY 226
4.1. Changes requested by SAGE IVD 226
HEPATITIS B 226
4.2. Hepatitis B e antigen 226
HIV INFECTION 227
4.3. Qualitative HIV virological nucleic acid test 227
TUBERCULOSIS 228
4.4. Lipoarabinomannan antigen 228
BACTERIOLOGY, MYCOLOGY AND PARASITOLOGY 229
4.5. Changes requested by SAGE IVD 229
CLINICAL CHEMISTRY 229
4.6. Changes requested by SAGE IVD 229
4.7. C-reactive protein 230
4.8. Creatinine 230
4.9. Glucose-6-phosphate dehydrogenase activity 231
HAEMATOLOGY 231
4.10. Changes requested by SAGE IVD 231
4.11. Complete blood count automated 232
CLINICAL PATHOLOGY (NEW SECTION) 232
4.12. Changes requested by SAGE IVD 232
SEROLOGY 232
CANCER 233
4.13. Alpha-fetoprotein 233

4.14. Basic panel for immunohistochemistry testing for diagnosis of lymphoma 234
4.15. Basic panel of immunohistochemical markers for diagnosis of solid tumours 235
4.16. Essential flow cytometry panel of antibodies for leukaemia 238
4.17. Basic panel of immunohistochemical markers for diagnosis of solid tumours 239
4.18. Tyrosine-protein kinase receptor or human epidermal growth factor receptor 2 overexpression 240
4.19. Lactate dehydrogenase activity 240
HEPATITIS B 243
4.20. Quantitative HBV virological nucleic acid test 243
HEPATITIS C 243
4.21. Qualitative or quantitative HCV virological nucleic acid test 243
HIV INFECTION 244
4.22. Quantitative HIV virological nucleic acid test 244
HUMAN PAPILLOMAVIRUS INFECTION 245
4.23. Human papillomavirus nucleic acid test 245
PRIMARY IMMUNODEFICIENCIES 245
4.24. Lymphocyte subtype enumeration 245
SEXUALLY TRANSMITTED INFECTIONS 246
4.25. Qualitative test for *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infections 246
TUBERCULOSIS 247
4.26. *Mycobacterium tuberculosis* DNA nucleic acid test 247
4.27. *Mycobacterium tuberculosis* DNA loop-mediated isothermal amplification 248
4.28. *Mycobacterium tuberculosis* DNA mutations associated with resistance 249
OTHER TRANSFUSION-TRANSMITTED ORGANISMS 249
4.29. IVDs to screen for pathogens in blood donations 249

5. Applications for Do Not Do recommendations in the EDL 250

5.1. Commercial serodiagnostic tests to diagnose tuberculosis 250
5.2. Western blot and line immunoassays to diagnose HIV infection 251

6. Applications for reversing conditional listings in the EDL 254

HIV INFECTION 254
6.1. *Histoplasma* antigen test 254
NEGLECTED TROPICAL DISEASES 256
6.2. Kato-Katz test for soil-transmitted helminths and intestinal schistosomes 256
PRIMARY IMMUNODEFICIENCIES 258
6.3. Enumeration of lymphocyte subtypes: CD4, CD8, CD20 and CD16/56 cells, B cells and NK cells 258
6.4. Immunoglobulin plasma levels 260
6.5. Zika virus immunoglobulin M 262

7. Summary and recommendations 265

7.1. IVDs adopted for EDL 3 265
7.1.1. Additions and revisions 266
7.1.2. Do Not Do recommendations 268
7.1.3. Rejections 268
7.2. IVDs recommended for EDL 4 269
7.2.1. New submissions 269
7.2.2. Additional evidence 270

7.3. General recommendations 271
7.3.1. EDL format 271
7.3.2. EDL processes 271
7.3.3. EDL-related products 271
7.3.4. EDL strategy 272

Annex 1 273

The third WHO model list of essential in vitro diagnostics (EDL 3) 273

Abbreviations

ACR	albumin : creatinine ratio
ACTH	adrenocorticotropic hormone
AFP	alpha-fetoprotein
AMH	anti-Müllerian hormone
APL	acute promyelocytic leukaemia
ARDS	acute respiratory distress syndrome
ART	assisted reproductive treatment
BAL	bronchoalveolar lavage
BDG	(1-3)-β-D-glucan
CADTH	Canadian Agency for Drugs and Technologies in Health
CBC	complete blood count
CC	clomiphene citrate
CCA	circulatory cathodic antigen
CDGP	constitutional delay of growth and puberty
CL	cutaneous leishmaniasis
CLL	chronic lymphocytic leukaemia
CMIA	chemiluminescent microparticle immunoassay
CND	Classificazione Nationale dei Dispositivi medici
CNS	central nervous system
COPD	chronic obstructive pulmonary disease
COVID-19	coronavirus disease
CPA	chronic pulmonary aspergillosis
CRP	C-reactive protein
CRS	congenital rubella syndrome
CSF	cerebrospinal fluid
CT	computed tomography
CVID	common variable immune deficiency

DALY	disability-adjusted life year
DAT	direct agglutination test
DBS	dried blood spot
DIA	digital immunoassay
DSD	disorder of sexual development
DOR	diagnostic odds ratio
DSD	disorder of sexual development
DUB	dysfunctional uterine bleeding
ECBS	Expert Committee on Biological Standardization
ECIL-5	Fifth European Conference on Infections in Leukaemia
EDL	WHO Model List of Essential In Vitro Diagnostics
EDTA	ethylenediaminetetraacetic acid
eEDL	electronic EDL
EGFR	epidermal growth factor receptor
EIA	enzyme immunoassay
EID	early infant diagnosis
ELISA	enzyme-linked immunosorbent assay
eEML	electronic EML
EML	WHO Essential Medicines List
EP	ectopic pregnancy
ESR	erythrocyte sedimentation rate
E2	estradiol
FISH	fluorescence in situ hybridization
FSH	follicle-stimulating hormone
GAS	group A streptococcus
GBD	global burden of disease
GCA	giant cell arteritis
GM	galactomannan
GMDN	Global Medical Devices Nomenclature

GMS	Grocott's methenamine silver
GnRH	gonadotropin-releasing hormone
G6PD	glucose-6-phosphate dehydrogenase
HBV	hepatitis B virus
hCG	human chorionic gonadotrophin
HCV	hepatitis C virus
HH	hypogonadotropic hypogonadism
HIA	haemagglutination inhibition assay
HIC	high-income country
HLA	human leukocyte antigen
HPG	hypothalamic-pituitary-gonadal
HPLC	high-performance liquid chromatography
HPO	hypothalamic-pituitary-ovarian
HPV	human papillomavirus
HSV	herpes simplex virus
HTA	health technology assessment
IA	invasive aspergillosis
ICD	International Classification of Diseases
ICT	immunochromatographic test
ICU	intensive care unit
IEF	isoelectric focusing
IFA	immunofluorescence assay
IFAT	immunofluorescence assay test
IgG	immunoglobulin G
IgM	immunoglobulin M
IHA	indirect haemagglutination
IHC	immunohistochemistry
IIF	indirect immunofluorescence
IUP	intrauterine pregnancy

IVD	in vitro diagnostic
IVF	in vitro fertilization
LAMP	loop-mediated isothermal amplification
LDH	lactate dehydrogenase
LFA	lateral flow assay
LFIA	lateral flow immunochromatographic assay
LH	luteinizing hormone
LMIC	low- and middle-income country
LOINC	Logical Observation Identifiers Names and Codes
LPA	line probe assay
LPD	luteal phase deficiency
MCBS	Magnitude of Clinical Benefit Scale
MEDEVIS	WHO Priority Medical Devices information system
MGUS	monoclonal gammopathy of undetermined significance
MoH	ministry of health
MRI	magnetic resonance imaging
NAT	nucleic acid test
NAAT	nucleic acid amplification test
NEDL	national EDLs
NGS	next-generation sequencing
NPV	negative predictive value
NSCLC	non-small cell lung carcinoma
NTD	neglected tropical disease
ODI	Oswestry Disability Index
OIA	optical immunoassay
ORR	objective response rate
PAX	paired box
PCOS	polycystic ovary syndrome
PCP	*Pneumocystis* pneumonia

PCR	polymerase chain reaction
PEP	protein electrophoresis
PFS	progression-free survival
PID	primary immunodeficiency
PLAP	placental alkaline phosphatase
POC	point of care
POI	primary ovarian insufficiency
PP	precocious puberty
PPE	personal protective equipment
PPV	positive predictive value
PRISM	Primary Care Streptococcal Management
PRL	prolactin
QALY	quality-adjusted life-year
qPCR	quantitative PCR
RADT	rapid antigen detection test
RBC	red blood cells
RCT	randomized clinical trial
RDT	rapid diagnostic test
ROC	receiver operating characteristic
RT-PCR	reverse transcription-PCR
SAB	spontaneous abortion
SAGE IVD	Strategic Advisory Group of Experts on In Vitro Diagnostics
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SCD	sickle cell disease
SCLC	small cell lung carcinoma
SEM	standard error of the mean
SFLC	serum-free light-chain
SLE	systemic lupus erythematosus
SLL	small lymphocytic lymphoma

SSPE	subacute sclerosing panencephalitis
STI	sexually transmitted infection
TB	tuberculosis
THS	thyroid stimulating hormone
TKI	tyrosine kinase inhibitor
TNF	tumour necrosis factor
TTP	thrombotic thrombocytopenic purpura
UHC	universal health coverage
UMDNS	Universal Medical Device Nomenclature System
VL	visceral leishmaniasis
WB	western blot
WBC	white blood cells
WHO	World Health Organization
WT1	Wilms' tumour antigen
ZIKV	Zika virus

List of participants

SAGE IVD members

Rashad Abdul-Ghani, Assistant Professor, Sana'a University, Sana'a, Yemen

Ghaia Alwahdani, Head, Quality Control Department, Ministry of Health, Jordan

Anurag Bhargava, Professor, Department of Medicine, Yenepoya University, India

Catharina Boehme, CEO, Foundation for Innovative New Diagnostics, Switzerland

Jean-Pierre Chanoine, Head, Endocrinology and Diabetes Unit, University of British Columbia, Canada

Gvantsa Chanturia, Head, Department for Virology, Molecular Biology and Genome Research, National Center for Disease Control and Public Health, Georgia

Wenxiang Chen, Director, National Center for Clinical Laboratories and Research, Ministry of Health, China

Ronnie Gavilán, Research Biologist, National Institute of Health, Peru

Ravnit Grewal, Principal Investigator, South African National Bioinformatics Institute, South Africa (until November 2020); Principal Pathologist/Senior Lecturer, NHLS Universitas and University of the Free State, South Africa

Isao Hamaguchi, Director, Department of Safety Research on Blood and Biological Products, National Institute of Infectious Diseases, Japan

Michelle McNiven, Director, IVD Medical Device Reforms Section, Medical Devices Branch, Therapeutic Goods Administration, Australia

Shahin Sayed, Associate Professor, Department of Pathology, Aga Khan University Hospital, Kenya

Lee Schroeder, Associate Professor, Department of Pathology, University of Michigan Medical School, USA

Valerie Wilson, Director, Caribbean Med Labs Foundation, Port-Of-Spain, Trinidad and Tobago

Sangay Zangmo, Senior Laboratory Officer, Royal Centre for Disease Control, Ministry of Health, Bhutan

Nwe Oo, Deputy Director of Medical Devices, Department of Food and Drug Administration, Ministry of Health and Sports, Myanmar

Alice Emeka, Deputy Director, Medical Laboratory Services, Federal Ministry of Health, Nigeria

SAGE IVD members with no more than limited participation in deliberations:

Chansy Dalavong, Physician and Trainer, Obstetrics and Gynecology, Sethathieat Hospital, Lao People's Democratic Republic (until September 2020)

Ciprian Tomuleasa, General Secretary of Haematology Committee, Iuliu Hatieganu University of Medicine and Pharmacy, Romania (until June 2020)

Alpha Keita, Head, Molecular Biology Unit, Institut National de Santé Publique de Guinée, Guinea (until September 2020)

Observers

The following experts conducted the evidence review for all submissions received to support the addition of a test category to the list.

Jonathan Deeks, Deputy Director of the Institute of Applied Health Research, University of Birmingham, United Kingdom (UK)

Lotty Hooft, Associate Professor, Julius Center Methodology Research Program, Utrecht University, Netherlands

EDL Secretariat

Adriana Velázquez Berumen, Team Lead Medical Devices and IVDs (MDD)

Dorcas Appiah Agbogla, Administrative Assistant (MDD)

Francis Gabriel Moussy, Scientist (MDD)

Lucy Hattingh, Technical Adviser (Consultant) (MDD)

Mercedes Pérez González, Technical Officer (MDD)

Noelle Mendy, Administrative Assistant (MDD)

Magdalena Rabini, Programme Manager (MDD) (until March 2020)

Reza Sasanto, Programme Manager (MDD) (March to July 2020)

Sian Lewis, Technical Editor (Consultant)

Ana Elisa Aceves, Technical Officer (MDD)

Nikhil Moorchung, Technical Officer (until March 2020)

Giselle Weiss, Technical Editor (Consultant)

WHO invited experts

Albert Figueras (Consultant), Essential Medicines (EML)

Anita Sands, Technical Officer, Incidents, Substandard and Falsified Products (ISF)

Bernadette Cappello, Technical Officer, Essential Medicines (EML)

Felipe Roitberg (Consultant), Management of Noncommunicable Diseases, (MND)

James Kiarie, Unit Head, Contraception and Fertility Care (CFC)

Lorenzo Moja, Scientist, Essential Medicines (EML)

Mick Mulders, Scientist, Immunization Analysis and Insights (IAI)

Mohamed Ismail, Technical Officer, Essential Drugs and Medicines, WHO Bangladesh Country Office

Muhammad Jamil, Technical Officer, Testing, Prevention and Populations (TPP)

Rita Kabra, Technical Officer, Contraception and Fertility Care (CFC)

Saurabh Jain, Scientist, Prevention, Treatment and Care (PTC)

Declarations of interests

Management of conflicts of interest is a priority in preparing the WHO Model Essential In Vitro Diagnostics List (EDL) and in establishing the Strategic Advisory Group of Experts on In Vitro Diagnostics (SAGE IVD).

Before participating in SAGE IVD work, all members, advisers, consultants, reviewers and observers submitted written disclosures of relevant competing interests for consideration. Possible conflicting interests include employment by a commercial entity, consultancy or board, advisory board membership, lecture fees, expert witness income, industry-sponsored grants including for contracted research, patents received or pending, royalties, stock ownership or options, other personal financial interests or any financial relation between the institution or employer and a commercial entity with an interest in the field of the IVDs evaluated by SAGE IVD.

SAGE IVD members were also asked to disclose academic or scientific activities (including leading research or grant applications) in primary clinical studies or reviews with direct bearing on a decision about IVDs. In addition, all members were asked at the start of the meeting to update their declarations if any new conflicts had arisen in the meantime.

After analysing each declaration, the EDL Secretariat, assisted by the Office of Compliance, Risk Management and Ethics, concluded that all conflicts of interests declared were minor and there were none that would exclude any member from participating fully in SAGE IVD.

Acknowledgements

WHO acknowledges the technical support of all WHO staff who collaborated in the submission and review of tests for the EDL and participated in regular meetings, including Ana Aceves Capri, Dorcas Appiah Agbogla, Jorge Alvarez, Lucy Hattingh, Andre Ilbawi, Saurabh Jain, Cheryl Johnson, Rita Kabra, Alexei Korobitsyn, Nikhil Moorchung, Mick Mulders, Mercedes Perez Gonzalez, Magdalena Rabini, Reza Sasanto and Adriana Velazquez Berumen.

WHO also thanks the following experts for reviewing applications and providing feedback to SAGE IVD:

- **Susan Best**, Former director, National Serology Reference Laboratory, Fitzroy, Australia
- **Diana Hardie**, NHLS Pathology Laboratory, Groote Schuur Hospital, Cape Town, South Africa
- **Sally A. Hojvat**, Consultant, Rockville, USA
- **Joseph P. Icenogle**, US Centers for Disease Control and Prevention, Atlanta, USA
- **Katharina Kranzer**, London School of Hygiene and Tropical Medicine, London, UK
- **Adrian Newland**, The Royal London Hospital, Barts Health NHS Trust, London, UK
- **Olga Perovic**, University of Witwatersrand, Johannesburg, South Africa
- **Craig Platt**, Boston Children's Hospital, Boston, USA
- **Fernando Bellissimo Rodrigues**, University of São Paulo, São Paulo, Brazil
- **Sadia Shakoor**, Aga Khan University, Karachi, Pakistan
- **Lynette M. Sholl**, Harvard Medical School, Boston, USA
- **Dario Trapani**, European Institute of Oncology (IEO), IRCCS, Milan, Italy

WHO thanks the UK Department for International Development (DfID) for its financial support to WHO which allowed the establishment of SAGE IVD and the development of the EDL. Financial support from the Bill & Melinda Gates Foundation is also acknowledged.

Introduction

SAGE IVD advises the Director-General of WHO in the area of in vitro diagnostics (IVDs). With members appointed from a roster of experts, the group provides WHO with technical advice on global policies and strategies related to priority, essential and neglected IVDs.

The group also oversees maintenance of the WHO Model List of Essential In Vitro Diagnostics (EDL), which serves as an evidence-based reference point for countries to develop their own national lists to guide how they choose and use IVDs. The EDL recognizes that IVDs are essential for advancing universal health coverage (UHC), addressing health emergencies and promoting healthier populations, which are the three strategic priorities of WHO's thirteenth general programme of work covering 2019–2023 (GPW13).

The EDL is updated every year through a broad consensus-building process, including expert review and public consultation. As part of that process, SAGE IVD members usually meet face to face to jointly assess individual applications to update the EDL. This annual meeting is also used as a platform for exchanging experience and expertise: key stakeholders attend open sessions to share their views and concerns about the EDL; WHO technical staff update SAGE IVD on related areas of ongoing work; and SAGE IVD members debate key issues in IVD policy and EDL strategy.

In March 2020, the group's annual face-to-face meeting was cancelled due to the global pandemic of coronavirus disease (COVID-19). Instead, SAGE IVD established a remote process to update the EDL and provide WHO with IVD-related advice. Through successive rounds of remote voting and online deliberations from April to July 2020, and two extraordinary sessions in October and November, SAGE IVD managed to consider and make recommendations on every application made to update the EDL, including late-breaker submissions of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) tests, as well as to share progress updates and discuss strategic topics.

This report provides a summary of the group's deliberations, decisions and recommendations achieved through this process.

1. EDL and the review process

1.1. The EDL so far

The EDL lists IVDs that are recommended by WHO for use in countries to improve access to IVD testing. The list is not intended to be prescriptive with respect to the specific tests or levels at which IVDs should be used. Rather, it aims to serve as a reference for programme and laboratory managers, procurement officers and reimbursement officers who are developing or updating their own national lists of essential diagnostics within the framework of UHC. In all cases, countries are expected to decide for themselves which IVDs to select and where to use them, depending on their epidemiology, funding, human resources and infrastructure.

The EDL is also used by UN agencies and nongovernmental organizations to support selection, procurement, supply, donation or provision of IVDs, and can be used by the private medical technology sector to gain insight into the diagnostic priorities and related IVDs needed to address global health issues.

The first EDL (EDL 1), which was published in 2018, comprised 62 test categories and 107 test formats that were considered essential to every health care system in the world. It included general diagnostic tests and disease-specific tests as well as tests for screening blood donations for transfusion purposes. EDL 1 defined several features of each test, including relevant health care setting, test purpose, assay format and specimen type, alongside a list of any related WHO guidance and recommendations.

In 2019, the list was updated to cover 122 test categories, based on proposed additions and changes submitted by stakeholders and reviewed by SAGE IVD. This second EDL (EDL 2) included a selection of IVDs that were listed conditionally, pending further evidence of test performance, clinical utility or acceptability.

The third EDL (EDL 3, see Annex I) similarly reflects additions and edits submitted by stakeholders and reviewed by SAGE IVD. Importantly, and following a SAGE IVD recommendation in 2019, EDL 3 introduces a list of "Do Not Do" recommendations to identify tests that are not useful in informing clinical management, performing surveillance or informing critical aspects of population health status (see Section 5). SAGE IVD recommends against the procurement or use of any IVD category listed in the Do Not Do recommendations.

SAGE IVD members conducted a thorough review of the list and further harmonized terms used to describe test purposes. In addition, given the ongoing COVID-19 pandemic, SAGE IVD requested the EDL Secretariat to consider the addition of SARS-CoV-2 tests to the EDL 3 through the normal submission process. Two extraordinary sessions were held to address these tests, causing a delayed release of EDL 3.

Find out more about the EDL at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics

1.2. Methods used to develop EDL 3

EDL 3 builds on EDL 2, taking into consideration proposals for additions and edits that emerged from public consultation. EDL 3 was updated using a stepwise approach:

i. **Open pre-submissions.** Online pre-submissions for EDL 3 were accepted from August 2019. They were open to all stakeholders and required for all four types of EDL updates: the addition of a new test category, an edit to an existing entry, the removal of an existing test category and the addition of a test category to the Do Not Do recommendations. Each pre-submission was screened by the EDL Secretariat and WHO technical experts for their completeness of information and alignment with WHO policies and guidelines. A form was available to submit additional evidence in support of the lifting of conditions for all conditional listings in EDL 2.

ii. **Invited full submissions.** Successful pre-submissions were invited to submit a full submission in October 2019. Full submissions must include all the evidence needed to support the test's inclusion in the EDL, including systematic reviews or primary studies of the test's clinical accuracy, utility and impact on patients as well as guidance and recommendations on the test's use from expert bodies.

iii. **Expert review.** All full submissions were reviewed by SAGE IVD members (or external experts, where necessary), and applicants were invited to respond to reviewers' comments.

iv. **Public consultation.** Each full submission, alongside SAGE IVD reviews and applicant responses, was published online on 15 February 2020 for 8 weeks of public review and comment.

v. **SAGE IVD deliberations.** Following the declaration of COVID-19 as a pandemic, the third SAGE IVD meeting planned for end of March 2020 was cancelled, and individual submissions were instead presented, discussed and decided upon using a remote process (see Section 1.2.2 below) from April to July 2020. This process did not alter the criteria for selecting tests for inclusion in the EDL, which include public health and clinical need, availability of validated commercial tests, clinical utility, diagnostic accuracy, cost–effectiveness, infrastructure requirements and appropriateness to specified health care setting. Extraordinary sessions to review SARS-CoV-2 tests took place in October and November 2020.

vi. **WHO Director-General approval.** All SAGE IVD recommendations were presented to the WHO Director-General for his consideration and final approval before the meeting report and EDL are published.

Find out more about EDL processes at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics

1.2.1. Changes to submission forms

In 2019, the SAGE IVD recommended reviewing the submission process to make it easier to understand and to improve the quality of submissions. A working group of 2019 SAGE IVD members updated the submission forms and drafted guidelines for applicants.

Notable differences from the old process at the pre-submission stage include:

- greater focus on the availability of evidence about the proposed IVD's performance and utility (rather than on the summary and description of IVD characteristics);
- new questions about the proposed IVD's links to essential medicines; and
- new focus on how the proposed IVD will be applied (triage, replacement or add-on, in testing algorithms, etc.).
- Notable differences at the full submission stage include greater clarity on expected content for sections on:
- systematic reviews and primary studies of the proposed IVD's clinical accuracy;
- systematic reviews and primary studies of the clinical utility for the proposed indication;
- infrastructure requirements for the proposed IVD;
- cost–effectiveness evaluations of the proposed IVD; and
- potential ethics, equity and human rights concerns associated with using the proposed IVD.

1.2.2. 2020 remote process for deliberations

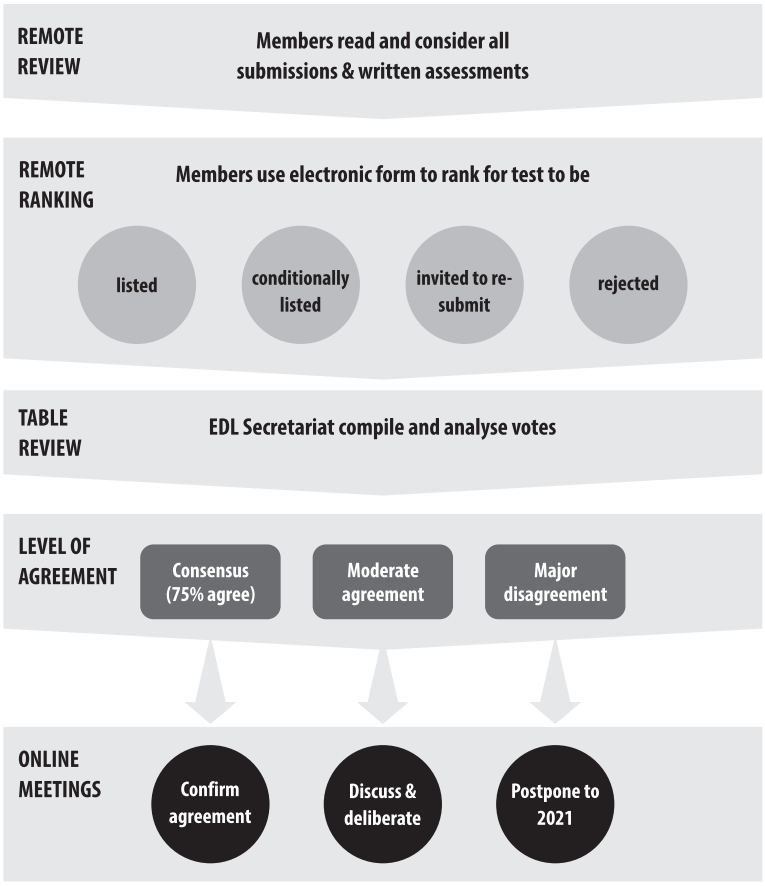
The remote process for SAGE IVD deliberations was designed to advance, as far as possible, the group's work, while acknowledging that it could not hope to replace the face-to-face meeting. In that context, the process had a defined methodology that was both practical and realistic but that did not compromise

the quality of the group's decision-making. The process covered two main types of activity: discussing key issues and reviewing full submissions to the EDL.

Members received working documents by correspondence. Online meetings were scheduled for them to receive brief progress updates and to discuss key issues around new or ongoing EDL-related products and processes (see Section 2).

The process to review full submissions used a combination of remote ranking of submissions and online meetings (see Fig. 1).

Fig. 1
Remote process for SAGE IVD decision-making



All SAGE IVD members were first asked to examine all the information that would have been presented during the face-to-face meeting and use a standard electronic form to rank the categories of tests in question to be:

1 – listed in EDL 3: because the evidence convincingly shows the test is accurate, clinically useful and accessible;

2 – conditionally listed in EDL 3: because the evidence around certain aspects or uses of the tests is lacking in the submission but available, in which case its listing is contingent on the provision of this additional evidence;

3 – invited for submission to EDL 4 (with further evidence): because while the tests are acknowledged to be useful, there is not enough evidence in the submission to recommend their inclusion; or

4 – rejected: because there is not enough evidence on the tests' performance or value.

SAGE IVD members were also given the opportunity to abstain from issuing an opinion about a category of tests outside their area of expertise and for which they felt unqualified to review the evidence presented.

In each case, SAGE IVD members were asked to provide a rationale for their ranking, including any recommended changes to the wording and any major or minor concerns. Once all the rankings were received, the EDL Secretariat analysed them to establish which categories of tests had a quorum (60% of SAGE IVD members) or consensus (75% of SAGE IVD members issuing a ranking) for decision-making.

For all categories of tests that achieved a quorum of ranking members and consensus for listing, the discussion at online meetings was limited to agreeing the specific test purpose, format, specimen type and relevant health setting. Those tests that failed to achieve this during the remote ranking, or displayed major disagreements in ranking patterns, were fully discussed by SAGE IVD members at their online sessions, with inputs from evidence reviewers and WHO experts. In each case, members evaluated the evidence available for the test's clinical utility, performance (accuracy, precision, variability between IVD products) and cost–benefit ratio. The quality of available evidence was also considered.

Applications to edit the EDL or reverse conditional listings were similarly considered through online deliberations.

2. EDL-related products and processes

2.1. Progress updates

2.1.1. eEDL

EDL 3 will be released both electronically and in print. Having an electronic EDL (eEDL) is important to increase the list's accessibility and availability across all countries and contexts. It is also important to improve the links between the EDL and other global IVD initiatives and WHO priority lists.

IVDs will be listed alphabetically in the eEDL. But the database will be much easier to navigate than the print version, because users can filter IVDs by recommended health care setting or search for individual IVDs within specific fields (specimen type, assay formats and test purposes). Information on each field is provided in full on each individual IVD webpage, alongside SAGE recommendations and evidence summaries. Individual IVD webpages are printable; and the list as a whole can be exported as a Microsoft Excel worksheet or a pdf.

As well as being easier to use, the eEDL will also enable more efficient and effective updating of the EDL. The entire submission process will be moved to the online platform; and all edits to the EDL will be automatically archived, along with any additional data used to support the change. In this way, users will be able to clearly see exactly what has changed, when and why.

Development of the eEDL was informed by an online survey of laboratory technologists, pathologists, clinicians and public health experts to find out which aspects of the EDL were most useful. The survey, which was distributed through departmental newsletters and sent directly to external partners and stakeholders, generated more than 70 responses. These indicated that the most useful information for clinicians is test names and purposes, while laboratorians find information on assay formats and specimen types of more use.

The eEDL will include tests that are conditionally listed, but these are clearly marked as such.

SAGE IVD recommendations

SAGE IVD noted the progress made and established a working group to review the current beta version of eEDL before publication in late 2020.

2.1.2. Country guidance

In 2019, during its second meeting, SAGE IVD recommended developing a new guidance document to help countries implement the EDL by developing or updating their own national EDLs (NEDLs).[1]

The EDL Secretariat has worked with colleagues and in-country experts to draft such a document, which was shared discussed with SAGE IVD for initial feedback. The new guidance details the steps countries need to take to develop an NEDL (or update an existing one).

The guidance is essentially a "how-to" for ministries of health (MoHs) to develop or update an NEDL to further improve access to IVD testing in support of UHC programmes and national emergency preparedness and response efforts. But it will also be relevant to the many diverse individuals and organizations who are stakeholders in regulating, selecting and using IVDs in-country, such as regulatory, planning, procurement and reimbursement agencies, and laboratory managers.

The guidance is essentially a "how-to" for ministries of health (MoHs) to develop or update an NEDL to further improve access to IVD testing in support of UHC programmes and national emergency preparedness and response efforts. But it will also be relevant to the many diverse individuals and organizations who are stakeholders in regulating, selecting and using IVDs in-country, such as regulatory, planning, procurement and reimbursement agencies, and laboratory managers.

SAGE IVD recommendations

SAGE IVD noted the progress made and established a working group to finalize the country guidance before publication.

2.1.3. EDL technical specifications

EDL technical specifications are designed to support the procurement of specific products in line with test categories listed in the EDL.[2] Made up of a range of pre-defined criteria, they provide information on the minimum requirements needed to ensure products selected under each category conform to minimum predefined standards. They are intended for use by countries and laboratories.

[1] The selection and use of essential in vitro diagnostics: report of the second meeting of the WHO Strategic Advisory Group of Experts on In Vitro Diagnostics, 2019 (including the second WHO model list of essential in vitro diagnostics). Geneva: World Health Organization; 2019 (WHO Technical Report Series, No. 1022).

[2] For more information on recommended procurement practices for IVDs and laboratory equipment see: WHO Manual for procurement of diagnostics and related laboratory items and equipment. Geneva: World Health Organization; 2013 (https://apps.who.int/iris/handle/10665/255568, accessed 21 April 2020).

But, like the EDL itself, they are not absolute and must be tailored to meet individual country needs.

The EDL technical specifications are being developed based on all existing WHO guidance and specifications that support the procurement, donation or lease of good-quality IVDs and medical devices.[3] In developing the technical specifications, the EDL Secretariat acknowledges a need to support more uniformity in tendering. Because while most organizations specify details of a diagnostic kit at the time of tendering, these can vary widely.

All the EDL technical specifications will be shared with subject matter experts, procurement agencies and specialists for feedback; they will also be open to public comment on the WHO website for 2 months. All feedback received will be used to inform a set of revisions, and the final list of EDL technical specifications for procurement will then be ratified by the EDL Secretariat and shared with the SAGE IVD.

SAGE IVD recommendations

SAGE IVD noted the progress made and established a working group to provide technical feedback before publication.

2.1.4. Harmonization activities

Harmonization activities over the past year can be grouped into two broad categories.

First has been the work to align IVD tests in the EDL with the International Classification of Diseases 11th Revision (ICD-11), which provides a common language for classifying diseases, injuries and causes of death, and for consistently reporting and monitoring health conditions. Every test entry in the eEDL that is related to a disease or specific health condition includes a link to the corresponding ICD-11 code.

A mapping between the different nomenclature systems has been drafted linking available information from the Global Medical Device Nomenclature (GMDN), Universal Medical Device Nomenclature System (UMDNS), Classificazione Nazionale dei Dispositivi medici (CND) and Logical Observation Identifiers Names and Codes (LOINC), noting that CND will become the European Medical Device Nomenclature (EMDN). This is an ongoing effort that will continue in 2021 and will be reflected in the eEDL once the nomenclature system is defined and complete.

[3] Guidance for procurement of in vitro diagnostics and related laboratory items and equipment. Geneva: World Health Organization; 2017 (https://apps.who.int/iris/handle/10665/255577, accessed 21 April 2020).

Second has been the EDL Secretariat's contribution to developing a new WHO global public good to support priority setting for benefits packages for UHC at country level, known as the UHC compendium. This new tool will link the diagnostic tests in the eEDL to the clinical interventions listed in the UHC compendium, which is a single interactive database. Work is in progress with many other departments in WHO, and eventually users will be able to search the online database by various criteria (e.g. disease or level of care or clinical intervention) to find information on health products required to carry out the intervention, and clinical guidelines for the use of the products and the health workforce required to conduct the intervention. It is also foreseen that the UHC compendium will interact with the eEML, the electronic version of the Essential Medicines List, creating an interaction between the electronic databases.

WHO's Medical Devices information system, called MEDEVIS, will also interact with the UHC compendium. This electronic database includes both the eEDL and the Priority Medical Devices. MEDEVIS is under development to be finalized in 2021.

SAGE IVD recommendations

SAGE IVD noted the progress made and established a working group to contribute to the different harmonization initiatives.

2.2. Prioritization of the EDL

In the context of UHC, all countries have to undertake some prioritization when defining which services they want to cover in the face of limited health financing. Such decisions are usually based on the availability of key resources, including workforce, guidelines for use, health technologies (including diagnostics), medicines, infrastructure and information.

In the realm of health technologies, this type of prioritization for UHC is generally called health technology assessment (HTA). It takes various factors into consideration, including clinical effectiveness, ethics, social issues and organizational frameworks. HTA happens as part of the natural process of ensuring access to appropriate and safe health technologies, alongside regulation and management.

The extent to which countries require HTA depends on their income level: the more limited the resources, the more HTA is needed. In high-income countries (HICs) with strong health systems and high levels of health coverage, HTA is only really used to decide on the inclusion of innovative tests. But in middle- and high-income countries with medium coverage, HTA is used to define a broader package of interventions, while in low-income countries with low levels of health coverage, HTA may even be required for the most essential health care interventions.

The EDL was conceived as a tool to help countries in their prioritization task, whatever their circumstances. By setting out a group of WHO-recommended IVD tests for different levels of a tiered testing system, the EDL aims to provide guidance to countries that are creating their own lists of national essential IVDs for defining UHC interventions and choosing the right IVDs to implement them.

The question is how best to package that guidance in a way that both responds to individual country needs and integrates and aligns with WHO's other prioritization tools, including the EML, the WHO Priority Medical Devices list and the WHO Priority Assistive Products List.

At its virtual meetings in 2020, SAGE IVD considered two key options:

1. **Embed prioritization into the EDL**, for example by tagging test categories in the list as core or complementary. This is the approach that the EML takes, in which the core list sets out the minimum medicine needs for a basic health care system, while the complementary list sets out essential medicines for priority diseases that require specialized equipment, infrastructure or care.

2. **Develop multi-decision criteria and methodology** that countries can use themselves to identify which tests to focus on during their HTA. Such criteria may include, for example, the availability of WHO guidelines, the test type (general, disease-specific, blood and organ screening) and the intervention type (high impact, early diagnosis, high prevalence).

The first option is more centralized and has the advantage of being aligned with the EML's approach. But because the EDL is so new, most tests currently listed are likely to be core. And even if the list were differentiated, a full country-specific HTA would still be needed, because health resources and contexts vary so much from country to country.

Still, there are potentially other ways that the EDL could be restructured to better support country prioritization. One way might be to review whether or not to assign tests to different tiers of the health care system (because the conditions that are addressed at each tier may vary across countries). Another approach might be to further refine the "with labs" tier (EDL, Section II) to distinguish tests that "can" be done with limited laboratory resources from those that require an established laboratory.

The second option of developing a multi-decision criteria and methodology for prioritization has the advantage of supporting a more tailored HTA that enables countries to address their own needs and priorities. This approach better reflects how the EDL was designed to be implemented (as a high-level policy tool); and it also naturally follows and fits with SAGE IVD's recommendation to develop country guidance for implementing the EDL (see Section 2.1.2).

Ultimately, the EDL is not the only tool available to support prioritization for UHC; and it should not be used in isolation. This means that whatever approach is recommended must be integrated not only with the other priority lists published by WHO but also with other WHO-led tools to support UHC (see Sections 2.1.3 and 2.1.4).

SAGE IVD recommendations

SAGE IVD agreed to establish a working group to develop a proposal for how best to enable prioritization, which will be presented to the 4th SAGE IVD for approval.

3. Applications for additions to the EDL

A total of 26 full applications for additions to the EDL were submitted by academia, industry and WHO technical departments, which were reviewed according to the methods set out in Section 1.2. In addition, SAGE-IVD also considered 26 applications for other revisions to the EDL; 5 applications to reverse the conditional listing of specific IVD categories in the EDL; and 2 applications for Do Not Do recommendations that advise countries against the procurement and use of specific IVD categories.

Online meetings convened to discuss additions to the EDL were led by a rotating chair, which included Ms Valerie Wilson, Professor Lee Schroeder, Dr Catharina Boehme and Professor Anurag Bhargava.

In considering the recommended test purpose for each accepted application, SAGE IVD used standard EDL definitions, which identify eight different test purposes:[4]

- **Screening tests.** Screening tests are used to determine the status of a disease, disorder or other physiological state in an asymptomatic individual. Depending on the nature of the condition and the targeted patient population, screening tests may be used routinely or may be restricted to "at risk" patients. These tests are designed to evaluate an individual's current state.
- **Diagnostic tests.** Diagnostic tests are used to determine, verify or confirm a patient's clinical condition as a sole determinant. This type of testing also includes sole confirmatory assays (to verify results of previous testing) and sole exclusion assays (to rule out a particular condition). These tests are designed to evaluate a patient's current state.
- **Aids to diagnosis.** Tests that are used as aids to diagnosis provide additional information to assist in the determination or verification of a patient's clinical status. The test is not the sole determinant. These tests are designed to evaluate a patient's current state.
- **Monitoring tests.** Monitoring tests are used for measuring levels of analytes for the purpose of adjusting treatments or interventions as required. Monitoring tests include:
  - Assays which are used to ensure that an analyte remains within physiological levels or within an established therapeutic drug

[4] Adapted from: Clinical evidence for IVD medical devices – clinical performance studies for in vitro diagnostic medical devices. Study group 5 final document GHTF/SG5/N8:2012. International Medical Device Regulators Forum; 2012 (http://www.imdrf.org/docs/ghtf/final/sg5/technical-docs/ghtf-sg5-n8-2012-clinical-performance-studies-ivd-medical-devices-121102.pdf, accessed 9 December 2020).

range. These types of monitoring tests are designed to evaluate an individual's current state.

  - Assays which are used for serial measurement, whereby multiple determinations are taken over time. These types of monitoring tests are typically used for the detection/assessment of disease progression/regression, disease recurrence, minimum residual disease, response/resistance to therapy and/or adverse effects due to therapy. These types of monitoring tests are designed to evaluate changes in an individual's state.

- **Prognostic tests.** These tests are used to measure factors linked to clinical outcome irrespective of treatment. Such tests may be used to estimate the natural progression of a disease (i.e. outcome in the absence of treatment), or to determine the likelihood of a clinical outcome irrespective of therapeutic intervention. These tests are designed to evaluate a patient's future state.
- **Surveillance tests.** Performed on populations of interest to track the progression of disease incidence and/or prevalence.
- **Staging tests.** Staging test: Performed on patients with a confirmed disease or condition to determine its state at the time of diagnosis and establish a baseline to make relevant treatment decisions.

EDL Section I.a: General IVDs for community settings and health facilities without laboratories

3.1. Erythrocyte sedimentation rate

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding an erythrocyte sedimentation rate (ESR) test to the EDL as an IVD to aid diagnosis and monitoring of certain infections and immune diseases; and as an alternative to a C-reactive protein (CRP) test where this is not available.

B. Applicant

University of New South Wales, Sydney, Australia

In March 2019, SAGE IVD discussed including ESR in the EDL and requested a full submission.

C. WHO technical department

None

D. Background (from application)

Disease condition and impact on patients

The ESR – with or without a concurrent CRP test – contributes to the diagnosis and management of certain infections and diseases of the immune system, including orthopaedic infections (*1*), rheumatoid arthritis (RA) (*2*) and systemic lupus erythematosus (SLE), as well as less common conditions, including giant cell arteritis (GCA) and polymyalgia rheumatica. RA and SLE typically present in young adulthood and are lifelong diseases causing significant disability and substantial socioeconomic impact.

Does the test meet a medical need?

The ESR was the mainstay of laboratory testing for inflammation during the 20th century. In inflammation, there is increased production of fibrinogen and other plasma proteins, which cause erythrocytes to adhere to each other, increasing the rate at which erythrocytes sediment when blood is placed in a vertical tube. The ESR is therefore an indirect measure of inflammation. Factors other than inflammation also have a substantial influence on the rate at which erythrocytes sediment (*3*, *4*).

How the test is used

In certain infections, particularly orthopaedic infections, an elevated ESR may be an early clue to infection, justifying microbiological investigations that, if positive, provide a specific diagnosis.

In disorders of the immune system, an elevated ESR, with or without CRP, provides objective evidence of significant inflammation. But it is too non-specific for final diagnosis, as ESR can be raised in many immune and infectious diseases. In certain immune disorders, CRP may not be as elevated as clinically expected because of a Type I interferon response (*5*). In these cases, ESR still has a role in assessment, often alongside CRP. But diagnosing and monitoring immune diseases is much more complex than infectious disorders, and so also includes clinical assessment, often with scoring systems, imaging and other blood tests. In the case of GCA, biopsy is also part of the diagnostic algorithm.

In many cases, if CRP is not available to diagnose or monitor infection, then ESR may be used. Elevated ESR is a non-specific finding, however, and diagnosis depends on infection-specific tests.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Rheumatoid arthritis is a chronic disease that typically presents in young adulthood and is prevalent in around 0.25% of the world population. In 2010, it had a burden of 4.8 million disability-adjusted life years (DALYs) (*6*).

Juvenile idiopathic arthritis affects around 10 in 100 000 people, although estimates vary widely (*7*). Some patients continue to be affected in adulthood.

SLE affects approximately 0.1% of the global population (*8*). It is a chronic disease that typically presents in teenagers or young adults. In South Korea, people with SLE have a standardized mortality ratio 2.6 times greater than the general population (*9*).

GCA and polymyalgia rheumatica affect 0.1–0.3% and 0.6% of the US population, respectively (*10*). Onset is in older adults and the conditions are chronic; they may occur in all racial groups but are best recognized in Europeans. Other non-organ-specific systemic autoimmune diseases are less common, and diagnosis is challenging. Presenting features may overlap the above conditions. There is limited evidence about ESR in these conditions.

Around 1–2% of recipients of hip and knee replacements experience prosthetic joint infections (*11*).

F. WHO or other clinical guidelines relevant to the test

The American College of Rheumatology (ACR) recommends using ESR (or CRP) as a disease activity measure in its DAS 28-ESR score, where the ESR contributes a sizeable proportion of the information in the final score (*2*).

The Royal Australian College of General Practitioners recommends the use of ESR (or CRP) in diagnosing idiopathic juvenile arthritis.

The European League against Rheumatism (EULAR) updated its guidelines for large vessel vasculitis (GCA and Takayasu arteritis) in 2018 (*12*), giving the use of ESR and CRP a 3b level of evidence and indicating that this dual measurement is useful in diagnosis.

The EULAR/ACR guidelines for polymyalgia rheumatica recommended assessing ESR and/or CRP for monitoring disease progress (*13*).

Armstrong et al. (*14*) compiled a set of "unified" guidelines for the diagnosis of prosthetic joint infections, which indicated that 92% of patients who went to surgery had a pre-operative ESR and CRP.

G. Basic test characteristics (from application)

Test formats available	Westergren, microphotometer; automated or manual
Specimen types	Whole blood: citrated or ethylenediaminetetraacetic acid (EDTA)

Table *continued*

Equipment required	Level benchtop, sedimentation rack, tubes
Regulatory status	Some CE-marked and US Food and Drug Administration (FDA)-approved products
Availability	Global
Price per test range	Manual ESR: £0.30 Automated ESR: £0.90 (4)
Instrument price range	Manual ESR: £60–300 Automated ESR: £30 000 (4)

H. Evidence for diagnostic accuracy (from application)

Lapic et al. (*15*) analysed 29 studies comparing the clinical accuracy of ESR and CRP. The studies included 16 on orthopaedic infections, 3 on rheumatic conditions (RA, GCA and Takayasu's disease) and 10 on miscellaneous conditions (mostly infective). Overall clinical accuracy was reasonable, CRP performed somewhat better than ESR, and combined use of CRP and ESR was often better than either alone.

Carli et al. (*16*) did a systematic review of articles on periprosthetic joint infection and found that serum markers including ESR and CRP performed moderately well, although in their view synovial examination was preferable.

Berbari et al. (*17*) did a meta-analysis on the performance of inflammatory markers in assessing prosthetic joint infections. CRP performed somewhat better than ESR though both were satisfactory. The combination of CRP and ESR was not assessed. Interestingly, the three studies on interleukin-6 (IL-6) showed it performed better than either CRP or ESR. IL-6 has been considered as a routine diagnostic test, but it is more expensive and technically more complex than CRP or ESR.

I. Evidence for clinical usefulness and impact (from application)

Gwinnutt et al. (*18*) reviewed studies on factors predicting outcome in inflammatory arthritis. They found that ESR performed better than CRP, but they concluded that clinical assessment was more informative than laboratory markers. Dejaco et al. (*19*) reviewed studies on prognostic factors in polymyalgia rheumatica, and found evidence from some but not all studies that a high ESR at the time of diagnosis is a poor prognostic factor.

Clinical utility in SLE has also been demonstrated. Feldman et al. (*20*) provide supportive evidence for the fact that diseases of the immune system such as SLE may be discordant, with high ESR and low CRP. Other studies provide data on usefulness in SLE for monitoring disease activity (*11*, *21*).

J. Evidence for economic impact and/or cost–effectiveness (from application)

Although the ESR is a long-established test, there is limited literature on its cost–effectiveness. A 2015 Canadian review (*22*) assessed the value of testing both ESR and CRP compared with testing either alone and found it reduced the rate of misdiagnosis in prosthetic joint infection. The cost of the combined test per misdiagnosis avoided, compared with the cost of CRP alone, was Can$ 240.62.

The benefits, opportunities, costs and risks of various diagnostic approaches to prosthetic joint infection were assessed in a US study on Medicare patients (*23*). The study found that the most effective strategy was to start with both ESR and CRP, and then follow these with arthrocentesis in patients testing positive for both.

K. Ethics, equity and human rights issues (from application)

None identified.

L. Summary of evidence evaluation

Some guidelines recommend measuring either CRP or ESR as part of the initial diagnostic workup for juvenile idiopathic arthritis, GCA, polymyalgia rheumatica and prosthetic joint infections, or for monitoring disease for GCA. But there is little evidence behind these recommendations, and they are rated as weak. The DAS28 disease activity score for RA includes either CRP or ESR, but other scores exist which seem to perform as well.

Systematic reviews exist of the accuracy of ESR for inflammation and joint infection which show some diagnostic relationship between ESR and these diseases, but one which is weak. CRP has similar, sometimes superior, performance. The sensitivity and specificity noted for different conditions are, respectively: 78% and 68% for orthopaedic infections; 77% and 59% for other inflammatory infections; 79% and 81% for prosthetic joint infection; and 84% and 30% for GCA (single study).

There were no real studies of impact, although the studies cited did give some evidence on other uses of ESR. There was evidence that discrepant high ESR and low CRP results had some diagnostic value in diseases of the immune system such as lupus, and that ESR was related to disease activity in lupus.

In summary, although ESR is recommended in guidelines to aid in the diagnosis of several diseases (including RA, juvenile idiopathic arthritis, bone infections, prosthetic joint infections, GCA, SLE and polymyalgia rheumatica), such recommendations are mainly based on expert opinion with little supporting evidence. There is no strong evidence supporting ESR as an essential test, although there are strong opinions that its use, despite its limitations, is helpful in the workup, diagnosis and monitoring of several different diseases.

Evidence of the accuracy of ESR suggests high rates of false positives and false negatives in all applications, similar to those of CRP.

Evidence showing the utility and impact of measuring ESR either for diagnosis or for monitoring is not available.

M. Summary of SAGE IVD deliberations

SAGE IVD noted that the evidence provided suggests that ESR may be inaccurate and that there is little evidence of the test's clinical utility and impact. But despite the test's limited sensitivity and specificity overall, a marked elevation of ESR > 100 mm/h is strongly associated with an underlying infection, neoplasm or rheumatic disease.

ESR is included in the guidelines of several inflammatory conditions and is a well-established test that is embedded in clinical practice around the world. It is used in most laboratories as a non-specific marker for inflammation for a broad range of infectious and autoimmune diseases. In particular, the test is regularly used to aid the diagnosis of conditions such as GCA and polymyalgia rheumatica; it is also often used to monitor rheumatic diseases, including RA and SLE; and it is sometimes used to aid the prognosis of conditions like Hodgkin's lymphoma.

ESR is a very simple test that has minimal requirements (it can even be done without a kit), which means that access to it can be ensured in resource-constrained environments. Importantly, SAGE IVD emphasized that in most cases, CRP is more sensitive than ESR, with a more robust reference range. This means that in most cases, CRP should remain the preferred choice of test, with ESR providing an alternative only in settings where CRP is not available. But SAGE IVD members highlighted two examples – SLE and low-grade bone and joint infections – where ESR may be the preferred method, because in both cases there is evidence that the condition elevates ESR without causing a rise in CRP.

N. SAGE IVD recommendations

SAGE IVD recommended including the erythrocyte sedimentation rate (ESR) test category in the third EDL:

- as a general IVD for use in community settings and health facilities without laboratories (EDL 3, Section I.a Haematology);
- using a Westergren method;
- to detect inflammation as an indicator of various conditions when CRP is not available.

O. References

1. Osmon DR, Berbari EF, Berendt AR, Lew D, Zimmerli W, et al. Diagnosis and management of prosthetic joint infection: clinical practice guidelines by the Infectious Diseases Society of America. Clin Infect Dis. 2013;56:e1–e25. doi:10.1093/cid/cis803.
2. Anderson J, Caplan L, Yazdany J, Robbins ML, Neogi T, et al. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. Arthritis Care Res. 2012;64:640–647. doi:10.1002/acr.21649.
3. Harrison M. Erythrocyte sedimentation rate and C-reactive protein. Aust Prescr. 2015;38(3):93–94. doi:10.18773/austprescr.2015.034.
4. Osei-Bimpong A, Meek JH, Lewis SM. ESR or CRP? A comparison of their clinical utility. Hematology. 2007;12(4):353–357. doi:10.1080/10245330701340734.
5. Dima A, Opris D, Jurcut C, Baicus C. Is there still a place for erythrocyte sedimentation rate and C-reactive protein in systemic lupus erythematosus? Lupus. 2016;25(11):1173–1179. doi:10.1177/0961203316651742.
6. Cross M, Smith E, Hoy D, Carmona L, Wolfe F, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis. 2014;73(7):1316–1322. doi:10.1136/annrheumdis-2013-204627.
7. Thierry S, Fautrel B, Lemelle I, Guillemin F. Prevalence and incidence of juvenile idiopathic arthritis: a systematic review. Joint Bone Spine. 2014;81(2):112–117. doi:10.1016/j.jbspin.2013.09.003.
8. Rees F, Doherty M, Grainge MJ, Lanyon P, Zhang W. The worldwide incidence and prevalence of systemic lupus erythematosus: a systematic review of epidemiological studies. Rheumatology. 2017;56(11):1945–1961. doi:10.1093/rheumatology/kex260.
9. Lee YH, Choi SJ, Ji JD, Song GC. Overall and cause-specific mortality in systemic lupus erythematosus: an updated meta-analysis. Lupus. 2016;25(7):727–734. doi:10.1177/0961203315627202.
10. Crowson CS, Matteson EL. Contemporary prevalence estimates for giant cell arteritis and polymyalgia rheumatica. 2015. Semin Arthritis Rheum. 2017;47(2):253–256. doi:10.1016/j.semarthrit.2017.04.001.
11. Tande AJ, Patel R. Prosthetic joint infection. Clin Microbiol Rev. 2014;27(2):302–345. doi:10.1128/CMR.00111-13.
12. Hellmich B, Agueda A, Monti S, Buttgereit F, de Boysson H, et al. 2018 update of the EULAR recommendations for the management of large vessel vasculitis. Ann Rheum Dis. 2020;79(1):19–30. doi:10.1136/annrheumdis-2019-215672.
13. Dejaco C, Singh YP, Perel P, Hutchings A, Camellino D, et al. 2015 recommendations for the management of polymyalgia rheumatica: a European League Against Rheumatism/American College of Rheumatology collaborative initiative. Ann Rheum Dis. 2015;74(10):1799–1807. doi:10.1136/annrheumdis-2015-207492.
14. Armstrong MD, Carli AV, Abdelbary H, Poitras S, Lapner P, et al. Tertiary care centre adherence to unified guidelines for management of periprosthetic joint infections: a gap analysis. Can J Surg. 2018;61(1):34–41. doi:10.1503/cjs.008617.
15. Lapic I, Padoan A, Bozzato D, Plebani M. Erythrocyte sedimentation rate and C-reactive protein in acute inflammation. Am J Clin Pathol. 2020;153(1):14–29. doi:10.1093/ajcp/aqz142.
16. Carli AV, Abdelbary H, Ahmadzai N, Cheng W, Shea B, et al. Diagnostic accuracy of serum, synovial, and tissue testing for chronic periprosthetic joint infection after hip and knee replacements: a systematic review. J Bone Joint Surg Am. 2019;101(7):635–649. doi:10.2106/JBJS.18.00632.

17. Berbari E, Mabry T, Tsaras G, Spangehl M, Erwin PJ, et al. Inflammatory blood laboratory levels as markers of prosthetic joint infection: a systematic review and meta-analysis. J Bone Joint Surg Am. 2010;92(11):2102–2109. doi:10.2106/JBJS.I.01199.

18. Gwinnutt JM, Sharp CA, Symmons DP, Lunt M, Verstappen SM. Baseline patient reported outcomes are more consistent predictors of long-term functional disability than laboratory, imaging or joint count data in patients with early inflammatory arthritis: a systematic review. Semin Arthritis Rheum. 2018;48(3): 384–398. doi:10.1016/j.semarthrit.2018.03.004.

19. Dejaco C, Singh YP, Perel P, Hutchings A, Camellino D, et al. Current evidence for therapeutic interventions and prognostic factors in polymyalgia rheumatica: a systematic literature review informing the 2015 European League Against Rheumatism/American College of Rheumatology recommendations for the management of polymyalgia rheumatica. Ann Rheum Dis. 2015;74(10):1808–1817. doi:10.1136/annrheumdis-2015-207578.

20. Feldman M, Aziz B, Kang GN, Opondo MA, Belz RK, et al. C-reactive protein and erythrocyte sedimentation rate discordance: frequency and causes in adults. Transl Res. 2013;161(1):37–43. doi:10.1016/j.trsl.2012.07.006.

21. Stojan G, Fang H, Magder L, Petri M. Erythrocyte sedimentation rate is a predictor of renal and overall SLE disease activity. Lupus. 2013;22(8):827–834. doi:10.1177/0961203313492578.

22. Assasi N, Blackhouse G, Campbell K, Hopkins RB, Levine M, et al. Comparative value of erythrocyte sedimentation rate (ESR) and c-reactive protein (CRP) testing in combination versus individually for the diagnosis of undifferentiated patients with suspected inflammatory disease or serious infection: a systematic review and economic analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2015.

23. Diaz-Ledezma C, Lichstein PM, Dolan JG, Parvizi J. Diagnosis of periprosthetic joint infection in Medicare patients: multicriteria decision analysis. Clin Orthop Relat Res. 2014;472(11):3275–3284. doi:10.1007/s11999-014-3492-2.

EDL Section I.b: Disease-specific IVDs for community settings and health facilities without laboratories

CORONAVIRUS DISEASE

3.2. SARS-CoV-2 antigen

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a SARS-CoV-2 antigen test to the EDL as an aid in the diagnosis of COVID-19 infection in symptomatic and asymptomatic individuals with known close contact with a confirmed case or to aid in the identification and investigation of outbreaks and community spread of COVID-19.

B. Applicant

Foundation for Innovative New Diagnostics (FIND)

C. WHO technical department

N/A

D. Background (from application)

Note: This submission was made in response to a call from WHO for late-breaker submissions for SARS-CoV-2 diagnostic tests.

Disease condition and impact on patients

COVID-19 is a result of infection caused by the SARS-CoV-2 virus. COVID-19 typically presents with flu-like symptoms of respiratory disease from mild to moderate to severe. The common symptoms at early onset include fever, cough, headache, myalgia and fatigue. More severe cases present with symptoms of pneumonia and acute respiratory distress syndrome (ARDS), including shortness of breath, confusion, low blood pressure, persistent pain or pressure in the chest, and lethargy. Diarrhoea and bloody sputum are indications of progressive severity, as are patchy shadows and ground-glass opacity observed in chest x-ray and tomography scans. Severe complications include sepsis, respiratory failure, heart failure and septic shock. COVID-19 infection presents a challenge for clinical diagnosis, as many infected (and infectious) patients may present as asymptomatic (*1*–*4*).

Does the test meet a medical need?

The clinical utility of SARS-CoV-2 infection testing lies in early identification and isolation of cases, but also in choosing the right therapeutic approach in a clinical picture that can mimic several other entities (*5*, *6*). WHO guidance on the use of rapid antigen tests recommends use in settings where "NAT is unavailable or where prolonged turnaround times preclude clinical utility" (*7*). Given the generally lower sensitivity of these tests compared to reverse transcription PCR (RT-PCR), they should only be used to identify COVID-19 infection in patients who are within 5–7 days of the onset of symptoms.

Identifying COVID-19 may also limit further diagnostic investigations for other etiologies and eventually limit antibiotic use for empiric treatment of assumed bacterial pneumonia (*8*).

In health care settings, early identification of infected individuals (both patients and health care workers) can prevent unrecognized spread of the virus within an institution. Early identification allows for proper isolation of infected patients and appropriate use of personal protective equipment (PPE) for health care workers.

How the test is used

Given the typically high specificity of rapid antigen tests, a positive result is considered indicative of SARS-CoV-2 infection. However, the sensitivity of these tests is variable, and therefore a negative test cannot be taken to rule out the presence of infection (*9*). Negative tests should be confirmed by an RT-PCR test or by a repeat antigen test when RT-PCR is not available.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

SARS-CoV-2 infection was first observed in Wuhan Province, China, in December 2019 as a new form of respiratory infection. From there it rapidly spread around the world, resulting in almost 60 million reported cases by November and claiming the lives of more than 1.3 million people. The infection has placed health care systems under severe strain around the world and resulted in serious global economic hardship, as countries restricted the movement of their populations and the conduct of daily business activities. Age appears to play a significant role in morbidity and mortality risk for COVID-19 (*8*). As seen with SARS-CoV, there have been very few SARS-CoV-2 deaths reported for the paediatric population, in stark contrast to the 14.8% mortality rate observed for patients > 80 years and 8% mortality rate for patients 70–79 years (*10–12*).

F. WHO or other clinical guidelines relevant to the test

In September 2020, WHO released guidelines for the use of SARS-CoV-2 rapid antigen tests (7). The guidelines stated that "SARS-CoV-2 Ag-RDTs [antigen rapid diagnostic tests] can be used to diagnose SARS-CoV-2 infection in a range of settings where NAAT [nucleic acid amplification test] is unavailable or where prolonged turnaround times preclude clinical utility provided that they meet minimum performance requirements of ≥ 80% sensitivity and ≥ 97% specificity compared to a NAAT reference assay." Several scenarios were described for the appropriate use of these tests. National public health entities around the world have rapidly developed and published similar guidelines, including the US Centers for Disease Control and Prevention (CDC), European Centre for Disease Prevention and Control (ECDC), and Health Canada, to name only a very few.

G. Basic test characteristics (from application)

Test formats available	**RDT**
Specimen types	Nasal swab, nasopharyngeal swab
Equipment required	No instrument, or small handheld or benchtop immunoassay reader

Table *continued*

Regulatory status	WHO Emergency Use Listing, FDA Emergency Use Authorization, CEIVD, SFDA, ANVISA and many others
Availability	Worldwide
Price per test range	From US$ 5/test to ~US$ 15/test
Instrument price range	~US$ 1000 where applicable

H. Evidence for diagnostic accuracy (from application)

The most significant review to date is a Cochrane review published in May 2020 (*9*) that reported eight evaluations in five studies of five different rapid antigen tests. While specificity was found to be consistently high at an average of 98.9% (95% CI: 97.3–99.5), sensitivity varied widely across the studies from 0 to 94%, with an average of 56.2% (95% CI: 29.5–79.8). However, the authors had low confidence in the results because several studies were poorly designed or reported, and many tests were done without following the manufacturer's instructions. The authors stressed the urgent need for further evaluations and stated their intention to update their review soon. The update was published after this application was submitted.

A meta-analysis of commercial assays registered in Brazil (*13*) included only two antigen direct assays to date in April, and showed a sensitivity range of 70–86% and specificity of 95–97%.

Several primary studies have been done to assess the performance of antigen tests for SARS CoV-2. A PubMed search yielded seven independent evaluation studies (*14–20*) and two manufacturer-sponsored studies (*21–22*). All of the assays evaluated in these studies were rapid point-of-care (POC) tests. No published studies were found for laboratory-based high-throughput automated SARS CoV-2 antigen tests.

All studies reviewed made similar observations, namely that the tests evaluated showed consistently high specificity when compared to an RT-PCR test, and hence strong positive predictive value. However, the tests also showed variable sensitivity depending on the study population, meaning that their negative predictive values are unreliable. Sensitivities were typically higher in patients with high viral loads (low RT-PCR cycle threshold values), i.e. those in the first 5–10 days after onset of symptoms.

These observations led to universally similar conclusions and recommendations that rapid antigen tests may be useful when testing patients within the first week after onset of symptoms, but not thereafter. Given that the first week after onset of symptoms (and a few days prior to the onset of symptoms) is the period when these patients are at their most infectious (high viral load), and given the ease of use, low cost and amenability of these tests

to mass production and broad distribution, they have an important role to play in infection control despite their limited sensitivity. Many more infectious individuals will be identified because of broad availability than will be missed because of lower sensitivity. The key caveat would be that negative test results would need to be managed carefully, and different suggestions have been made to address this issue. For example, to diagnose disease, patients with a negative antigen test and clinically suspected infection or known prior exposure should be retested with the antigen test if a PCR test is not available; ideally the result should be confirmed with an RT-PCR test.

I. Evidence for clinical usefulness and impact (from application)

Although no primary studies or reviews were found that demonstrated the clinical utility of the SARS CoV-2 rapid antigen tests, several of the clinical accuracy studies touched on the utility of this test in their conclusions. For diagnostic purposes, these tests may be best suited for use as a first-line test in acute patients within a few days of the onset of symptoms due to the fact that sensitivity decreases significantly as viral load declines. Negative tests in patients suspected of COVID-19 infection or with known prior exposure should be confirmed with an RT-PCR test. This approach may reduce the use of more costly RT-PCR tests and, thus, the burden on laboratories where they are performed.

The ease of use, relatively low cost (US$5 has been announced by two suppliers), and potential for mass production and wide distribution make this test a useful tool for broad-based, rapid identification of individuals infected with SARS CoV-2 in settings where RT-PCR may not be available as well as for screening in settings where the pretest probability of a positive result is low.

J. Evidence for economic impact and/or cost–effectiveness (from application)

No studies were found to demonstrate the cost–effectiveness of using rapid antigen tests for SARS-CoV-2.

K. Ethics, equity and human rights issues (from application)

No ethical issues were identified. Given that COVID-19 has disproportionately affected resource-constrained settings, the expanded availability of testing offered by the rapid antigen test may help to alleviate the disease burden in these populations (*23*, *24*).

L. Summary of evidence evaluation

The first field studies of these tests are now emerging. The studies show that whereas the specificity of the tests is generally very high (97–99%), their sensitivity may be compromised. Evaluations in different settings have shown

that sensitivity varies among them. There are no convincing data of the sensitivity of these tests in asymptomatic people. Although high specificity suggests that positive results in symptomatic patients indicate SARS-Cov-2 infection, the poor and variable sensitivity indicates that negative test results do not rule out disease, and a second test is required before infection can be ruled out.

M. Summary of SAGE IVD deliberations

Several concerns were raised about the unreliable sensitivity of the SARS CoV-2 antigen tests as shown in the various studies conducted to date. However, the test was considered essential especially to provide access to testing in settings where RT-PCR is not available. Nonetheless, it should be listed with clear caveats. SAGE IVD also decided not to recommend the test for use in asymptomatic individuals without known close contact with confirmed cases, because of concerns about the variable sensitivity and lack of test performance data in these individuals.

Other concerns expressed related to the rapidly evolving evidence base and ongoing updates to guidelines, and how these factors may affect the stated test purpose. It was decided that these concerns should also be included in the caveats to the test purpose for this test.

N. SAGE IVD recommendations

SAGE IVD recommended including the SARS-CoV-2 antigen test in the third EDL:

- as a disease-specific IVD for use in community settings and health facilities without laboratories (EDL 3, Section I.b, Coronavirus disease (COVID-19));
- using an RDT format, handheld or small benchtop instrument for POC use;
- to aid in the diagnosis of COVID-19 in settings where NAT is unavailable or where prolonged turnaround times preclude clinical utility; and
- to aid in the diagnosis of COVID-19 in the early symptomatic phases of illness, or in asymptomatic individuals with known close contact with a confirmed case.

The test was recommended for inclusion provided that the following caveats are clearly stated:

- A negative test does not rule out infection and should not determine clinical care.
- Listing was based on available evidence and interim WHO guidelines and is subject to change.

- Regulatory oversight of the majority of commercially available tests was limited to emergency use authorizations at the time of listing.

O. References

1. Pan American Health Organization, WHO Regional Office for the Americas. Laboratory guidelines for detection and diagnosis of the novel coronavirus (2019-nCoV) infection. Washington (DC): Pan American Health Organization; 2020.
2. Hu Z, Song C, Xu C, Jin G, Chen Y, et al. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. Sci China Life Sci. 2020;53(5):706–711. doi:10.1007/s11427-020-1661.4.
3. Wang Y, Kang H, Liu X, Tong Z. Combination of RT-qPCR testing and clinical features for diagnosis of COVID-19 facilitates management of SARS-CoV-2 outbreak. J Med Virol. 2020;92(6):538–539. doi:10.1002/jmv.25721.
4. Zhou F, Yu T, Du R, Fan G, Liu Y, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054–1062. doi:10.1016/S0140-6736(20)30566-3.
5. Huang C, Wang Y, Li X, Ren L, Zhao J, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497–506. doi:10.1016/S0140-6736(20)30183-5.
6. Vetter P, Vu DL, L'Huillier AG, Schibler M, Kaiser L, et al. Clinical features of covid-19. BMJ. 2020;369:m1470. doi:10.1136/bmj.m1470.
7. Antigen-detection in the diagnosis of SARS-CoV-2 infection using rapid immunoassays. Interim guidance. Geneva: World Health Organization; 2020. (https://www.who.int/publications/i/item/antigen-detection-in-the-diagnosis-of-sars-cov-2infection-using-rapid-immunoassays, accessed 30 November 2020).
8. Lai CC, Wang CY, Hsueh P-R. Co-infections among patients with COVID-19: the need for combination therapy with non-anti-SARS-CoV-2 agents? J Microbiol Immunol Infect. 2020;53(4):505–512. doi:10.1016/j.jmii.2020.05.013.
9. Dinnes J, Deeks JJ, Adriano A, Berhane S, Davenport C, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. Cochrane Database Syst Rev. 2020 Aug 26;8:CD013705. doi:10.1002/14651858.CD013705.
10. Yang X, Yu Y, Xu J, Shu H, Xia J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. Lancet Respir Med. 2020;8(5):475–481. doi:10.1016/S2213-2600(20)30079-5.
11. Li Q, Guan X, Wu P, Wang X, Zhou L, et al. 2020. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med. 2020;382:1199–1207. doi:10.1056/NEJMoa2001316.
12. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA. 2020;323(13):1239–1242. doi:10.1001/jama.2020.2648.
13. Castro R, Luz PM, Wakimoto MD, Veloso VG, Grinsztejn B, Perazzo H. COVID-19: a meta-analysis of diagnostic test accuracy of commercial assays registered in Brazil. Braz J Infect Dis. 2020;24(2):180–187. doi:10.1016/j.bjid.2020.04.003.
14. Liotti FM, Menchinelli G, Lalle E, Palucci I, Marchetti S, et al. Performance of a novel diagnostic assay for rapid SARS-CoV-2 antigen detection in nasopharynx samples. Clin Microbiol Infect. 2020 Sep 23; S1198-743X(20)30583-8. Advance online publication. doi:10.1016/j.cmi.2020.09.030.

15. Cerutti F, Burdinop E, Milia MG, Allice T, Gregori G, et al. Urgent need of rapid tests for SARS CoV-2 antigen detection: evaluation of the SD-Biosensor antigen test for SARS-CoV-2. J Clin Virol. 2020;132:104654. doi:10.1016/j.jcv.2020.104654.

16. Mak GC, Cheng PK, Lau SS, Wong KK, Lau CS, et al. Evaluation of rapid antigen test for detection of SARS-CoV-2 virus. J Clin Virol. 2020;129:104500. doi:10.1016/j.jcv.2020.104500.

17. Porte L, Legarraga P, Vollrath V, Aguilera X, Munita JM, et al. Evaluation of a novel antigen-based rapid detection test for the diagnosis of SARS-CoV-2 in respiratory samples. Int J Infect Dis. 2020;99:328–333. doi:10.1016/j.ijid.2020.05.098.

18. Lambert-Niclot S, Cuffel A, Le Pape S, Vauloup-Fellous C, Morand-Joubert L, et al., on behalf of the AP-HP/Universities/INSERM COVID-19 Research Collaboration. Evaluation of a rapid diagnostic assay for detection of SARS CoV-2 antigen in nasopharyngeal swabs. J Clin Microbiol. 2020;58:e00977-20. doi:10.1128/JCM.00799-20.

19. Diao B, Wen K, Zhang J, Chen J, Han C, et al. Accuracy of a nucleocapsid protein antigen rapid test in the diagnosis of SARS-CoV-2 infection. Clin. Microbiol. Infect. 2020 Oct 5. doi:10.1016/j.cmi.2020.09.057.

20. Scohy A, Anantharajah A, Bodéus M, Kabamba-Mukadi B, Verroken A, et al. Low performance of rapid antigen detection test as frontline testing for COVID-19 diagnosis. J Clin Virol. 2020;129:104455. doi10.1016/j.jcv.2020.104455.

21. Hirotsu Y, Maejima M, Shibusawa M, Nagakubo Y, Hosaka K, et al. Comparison of automated SARS-CoV-2 antigen test for COVID-19 infection with quantitative RT-PCR using 313 nasopharyngeal swabs, including from seven serially followed patients. Int J Infect Dis. 2020;99:397–402. doi:10.1016/j.ijid.2020.08.029.

22. Young S, Taylor SN, Cammarata CL, Varnado KG, Roger-Dalbert C, et al. Clinical evaluation of BD Veritor SARS-CoV-2 point-of-care test performance compared to PCR-based testing and versus the Sofia 2 SARS Antigen point-of-care test. J Clin Microbiol. 2020 Oct 6. doi:10.1128/JCM.02338-20.

23. Tai DBG, Shah A, Doubeni CA, Sia IG, Wieland ML. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. Clin. Infect. Dis. 2020 Jun 20;ciaa815. doi:10.1093/cid/ciaa815.

24. Covid-19: increased risk among ethnic minorities is largely due to poverty and social disparities, review finds. BMJ. 2020 Oct 22:371. doi:10.1136/bmj.m4099.

SICKLING DISORDERS

3.3. Sickle cell testing

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a POC diagnostic test to the EDL as an IVD for screening for sickle cell disease (SCD) and trait in newborns, adults and children.

B. Applicant

Silver Lake Research Corporation and the Centre of Excellence for Sickle Cell Disease Research & Training, University of Abuja (joint application)

C. WHO technical department

Not applicable

D. Background (from application)

Disease condition and impact on patients

SCD is a widely prevalent haemoglobinopathy in sub-Saharan Africa that is frequently deadly in early life, killing the majority of afflicted, undiagnosed children before their fifth birthday (*1*). In many high-income countries, universal newborn screening programmes coupled with prophylactic interventions and inexpensive treatment have dramatically reduced the mortality and morbidity of SCD during the first 20 years of life (*2*, *3*). But in sub-Saharan Africa and central India, where more than 90% of annual SCD births occur, newborn screening programmes have not been universally implemented, if at all, largely because of the cost and logistical burden of laboratory diagnostic tests (*4*).

Does the test meet a medical need?

WHO estimates that early diagnosis and intervention would prevent 70% of existing SCD deaths. The main barriers to implementing newborn screening programmes at scale include the cost of diagnostic tests, lack of adequately distributed laboratory infrastructure and lack of adequate, sustained funding. Standard clinical laboratory methods to identify Hb variants include gel-based or capillary electrophoresis, isoelectric focusing (IEF), and high-performance liquid chromatography (HPLC). These methods require collecting a relatively large volume of whole blood, as well as uninterrupted electrical supply, highly trained and dedicated operating personnel, and the transport of blood samples from POC to possibly distant testing facilities (*5*). With emphasis on UHC, POC diagnostic tests can be deployed to rapidly evaluate acutely ill babies and children in high-prevalence regions and offer appropriate life-saving interventions once an SCD diagnosis is made.

How the test is used

For children born with SCD to survive into adulthood, they must be diagnosed as early in life as possible. Once the disease is diagnosed, the patient can be clinically managed with inexpensive and widely available drugs such as hydroxyurea (to elevate expression of fetal haemoglobin) and penicillin (to prevent SCD-related infections).

The POC diagnostic test for sickle cell disease and trait provides a reliable result that supports a final diagnosis in newborns, children and adults without confirmation by a laboratory method. It can determine whether an individual is healthy, homozygous for SCD or is carrying the trait (heterozygous for SCD).

It requires just a small droplet of blood as the sample input and, with minimal training, returns an accurate diagnostic result within 10–15 minutes. This test has been validated with high accuracy in multiple published field studies, conducted at both rural clinics and developed health centres. It requires no electricity or cold chain, and is stable for up to 2 years in temperatures of 15–40 °C and high humidity.

In addition to widespread SCD diagnostic screening initiatives in maternity wards, enhanced population coverage can be obtained by screening children during visits to vaccination clinics, as well as screening all adults of marriage age. This three-pronged approach would be expected to increase the life expectancy of SCD patients, but also to decrease future disease incidence by allowing young adults to make informed parental and family planning choices based on the probability of passing on the sickle haemoglobin gene.

The test is innovative and has the potential to address a clinical need that is not met by existing technologies. Existing diagnostic methods for SCD include large machines for HPLC and electrophoresis which require electricity, infrastructure, equipment, reagents and trained personnel. These technologies are also relatively expensive, cannot be easily scaled and are not easily deployed in rural environments. By contrast, the POC diagnostic test for SCD does not need electricity or refrigeration and can be run just as easily in a hospital ward or waiting room as in an outdoor rural clinic. Also, since the test delivers results in just 10–15 minutes, its use could significantly reduce the number of patients lost to follow-up.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Up to 90% of children with SCD in sub-Saharan Africa are thought to die, undiagnosed, before their fifth birthday, making SCD one of the leading causes of childhood deaths in the region (*6*–*8*). Individuals with SCD in these regions are commonly identified only after hospitalization for severe pain or other overt or life-threatening manifestations of the disease. Its effects on mortality and quality of life and its economic burden on regional health care systems have led SCD to be declared both a disease of public concern by the UN General Assembly and a priority noncommunicable disease by WHO.

In 2015, McGann et al. called SCD a "tremendously under-recognized public health challenge" that may contribute up to 5% of under-5 deaths in Africa (*9*). The authors stated that universal public health efforts such as pneumococcal

immunization would likely improve the survival of infants with SCD, but suggested that SCD's contribution to childhood mortality remains largely unaddressed because the scale of the problem remains unrecognized. McGann et al. concluded that newborn diagnostic screening and preventive care for SCD in Africa is both feasible and highly cost-effective. And they recommended that both should be considered in the development of national health care strategies in the region.

In 2019, Simpson et al. suggested that the impact of SCD on infectious disease deaths, poverty and economic growth is slowing sub-Saharan Africa's progress towards key development indicators (*10*). The authors stated that technological, medical and systems innovation has made SCD more detectable, treatable and curable than ever before.

F. WHO or other clinical guidelines relevant to the test

Guidelines on the widespread clinical use of the POC test for SCD are being developed with reference to several key publications (*4*, *9–11*).

G. Basic test characteristics (from application)

Test formats available	POC lateral flow assay (LFA)
Specimen types	Fresh whole blood, EDTA whole blood, dried blood spot (DBS)
Equipment required	Not applicable
Regulatory status	CE-marked IVD; approved in several African countries; pending approval in India
Availability	In countries accepting CE-marked IVD and in Ghana, Kenya and Nigeria; pending approval in India
Price per test range	US$ 2–12, depending on supplier and country
Instrument price range	Not applicable

H. Evidence for diagnostic accuracy (from application)

No systematic reviews of the test's clinical accuracy exist yet, as the primary clinical studies are still very new (most were published in 2019). It is anticipated that such reviews will begin to be published in early 2020.

A primary study by Mukherjee et al. evaluated the diagnostic accuracy of a POC test against automated HPLC in a newborn screening programme (*12*). It found the test to have a sensitivity of 98.1% and specificity of 99.1% for all possible phenotypes (HbAA, HbAS and HbSS) detected.

Nankanja et al. designed a blinded, prospective diagnostic accuracy trial of a POC test for SCD as an investigational test compared with using capillary zone electrophoresis and found a sensitivity of 99.8% and specificity of 99.9% (*13*).

The first multicentre, real-world assessment of a low-cost POC device was conducted in a low- income country (*14*). Between September and November 2017, 1121 babies were screened using both the POC test and HPLC, with discordant samples confirmed by molecular diagnosis. The sensitivity and specificity of the POC test were found to be 93.4% and 99.9%, respectively.

Three further studies evaluated different products; all showed similar results, with the lowest sensitivity being 94.9% and specificity over 99% (*15–17*).

I. Evidence for clinical usefulness and impact (from application)

No systematic reviews of the test's clinical usefulness or impact exist yet, as the primary study reports are still very new (most were published in 2019). It is anticipated that systematic reviews will begin to be published in early 2020.

There are, however, two published reports with interesting findings about the test's impact on patient care. The first is the Nankanja et al. study in south-eastern Uganda, which reported that the POC test detected a trace level of haemoglobin A in a recently transfused patient's blood that the reference laboratory failed to detect (*13*). This makes the POC test more sensitive and more accurate than the laboratory test. This report was chosen to be presented as a late-breaking abstract at the 2018 American Society of Hematology conference, as one of only seven abstracts selected globally.

The second study, by Steele et al., showed that the POC test could test whole blood samples obtained via heel prick from low-birthweight, premature babies, whereas the laboratory reference test could not be performed for these patients because insufficient blood volume was available (*15*). This means that the POC test enables even the youngest of patients to be screened, with much less pain and discomfort (to the child and parents) during the blood collection process.

J. Evidence for economic impact and/or cost–effectiveness (from application)

The cost of the POC diagnostic test for sickle cell disease and trait is projected to be approximately US$ 2 per test strip (delivered price to the end user), with additional cost savings possible when manufactured at a large scale. Early diagnosis and intervention programmes for SCD are projected to be cost-effective in sub-Saharan Africa and India (*9*, *18*).

K. Ethics, equity and human rights issues (from application)

Potential ethical issues that must be considered involve the negative impact on reputation and societal value of individuals who are diagnosed as SCD positive. These individuals may experience stigma from peers or potential marriage partners; but they would also have increased quality of life and life expectancy.

Use of the POC SCD test is projected to heavily reduce health care inequities and increase health care accessibility, as even individuals with very few resources who live in inaccessible and rural areas will receive enhanced clinical care relating to SCD. The test enables diagnosis early in life and in virtually any testing environment: even midwives working in rural villages could conceivably administer the test. All SCD-positive individuals could then be subject to early interventions with pneumococcal prophylaxis, oral penicillin, folic acid and hydroxyurea, which is also being made available to rural populations via drone delivery in Ghana (*19*). Rapid diagnosis and inexpensive treatment of rural and city populations alike will dramatically improve the penetration of quality health care throughout Africa and India.

L. Summary of evidence evaluation

The evidence base for these tests is in its infancy, with no systematic review available nor any studies of test impact. The primary studies provided indicate that the tests have very high sensitivity and specificity. The studies have generally been well done, recruiting appropriate populations. Many used discrepant analysis to resolve differences between the reference standard and the POC test; but the possible bias this will have introduced is small, as there have been very low numbers of discrepant cells.

Several large, well-designed studies have shown the POC test for sickle cell disease and trait to have high accuracy, with sensitivities and specificities > 97% across newborns, children and adults. There is no evidence as yet of the impact of using the test.

M. Summary of SAGE IVD deliberations

SCD is a global health problem that contributes to excess mortality in children under 5 years of age; early recognition in infants is critical. Last year, sickle cell testing using electrophoresis was added to the EDL (EDL 2), and SAGE IVD requested the submission of a POC test.

The type of POC test under consideration is a simple, accessible test that can be used in rural settings. It is cheaper compared with gold standard tests, such as electrophoresis. While there is no systematic review available for the products under this test category, there are sufficient large, well-designed studies to show that they are very accurate. SAGE IVD noted that no studies for implementing the tests were submitted but acknowledged that, as new tests, it is unreasonable

to expect evidence on clinical impact. The group also noted that since the test category was submitted for consideration, the findings of a pilot implementation of one of these products in Nigeria have been published.

Some concerns were raised about the tests' lack of international regulatory approval, despite one of the commercially available tests holding a CE mark (in this case the CE mark is self-declared without any independent evaluation since the test is considered a low-risk IVD in the European context). Given the geographic distribution of the sickle cell gene, the test is unlikely to ever get stringent oversight from a regulatory authority in any of the founding members of the Global Harmonization Task Force (Australia, Canada, European Union, Japan or United States of America). One WHO expert also raised concerns about the potential for the addition of a POC RDT to the EDL to divert much-needed resources away from gold standard tests. But the recent report from Nigeria shows the tests can be procured for less than US$ 2.

Some concerns were also raised about how the test could and should be used. Clarity of purpose was emphasized as important, especially given that some countries have high prevalence of both SCD and thalassaemia and need to be able to differentiate between the two. And SAGE IVD cautioned against using DBS as a specimen type because of a lack of evidence and regulatory oversight of the DBS protocol.

N. SAGE IVD recommendations

SAGE IVD recommended including the sickle cell testing test category in the third EDL:

- as a disease-specific IVD for use in community settings and health facilities without laboratories (EDL 3, Section I.b, Sickling disorders);
- using a rapid diagnostic test format;
- to screen for or to aid in the diagnosis of sickle cell disease, sickle cell trait and other sickling disorders.

O. References

1. Grosse SD, Odame I, Atrash HK, Amendah DD, Piel FB, et al. Sickle cell disease in Africa: a neglected cause of early childhood mortality. Am J Prev Med. 2011;41(6 Suppl 4):S398–405. doi:10.1016/j.amepre.2011.09.013.
2. Wong TE, Brandow AM, Lim W, Lottenberg R. Update on the use of hydroxyurea therapy in sickle cell disease. Blood. 2014;24(26):3850–3857. doi:10.1182/blood-2014-08-435768.
3. Gardner K, Douri A, Drasar E, Allman M, Mwirigi A, et al. Survival in adults with sickle cell disease in a high-income setting. Blood. 2016;128(10):1436–1438. doi:10.1182/blood-2016-05-716910.
4. Aygun B, Odame I. A global perspective on sickle cell disease. Pediatr Blood Cancer. 2012;59(2):386–390. doi:10.1002/pbc.24175.

5. Patra PK, Khodiar PK, Hambleton IR, Serjeant GR. The Chhattisgarh state screening programme for the sickle cell gene: a cost-effective approach to a public health problem. J Community Genet. 2015;6(4):361–368. doi:10.1007/s12687-015-0222-8.

6. Piel FB, Patil AP, Howes RE, Nyangiri OA, Gething PW, et al. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. Lancet. 2013;381(9861):142–151. doi:10.1016/S0140-6736(12)61229-X.

7. Piel FB, Patil AP, Howes RE, Nyangiri OA, Gething PW, et al. Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. Nat Commun. 2010;1:104. doi:10.1038/ncomms1104.

8. Weatherall DJ. The inherited diseases of hemoglobin are an emerging global health burden. Blood. 2010;115(22):4331–4336. doi:10.1182/blood-2010-01-251348.

9. McGann PT, Grosse SD, Santos B, de Oliveira V, Bernardino L, et al. A cost-effectiveness analysis of a pilot neonatal screening program for sickle cell anemia in the Republic of Angola. J Pediatr. 2015;167(6):1314–1319. doi:10.1016/j.jpeds.2015.08.068.

10. Simpson S. Sickle cell disease: a new era 2019. Lancet Haematol. 2019;6(8):e393–e394. doi:10.1016/S2352-3026(19)30111-5.

11. Mvundura M, Kiyaga C, Metzler M, Kamya C, Lim JM, et al. Cost for sickle cell disease screening using isoelectric focusing with dried blood spot samples and estimation of price thresholds for a point-of-care test in Uganda. J Blood Med. 2019 Feb 5;10:59–67. doi:10.2147/JBM.S186528.

12. Mukherjee MB, Colah RB, Mehta PR, Shinde N, Jain D, et al. Multicenter evaluation of HemoTypeSC as a point-of-care sickle cell disease rapid diagnostic test for newborns and adults across India. Am J Clin Pathol. 2019;153(1):82–87. doi:10.1093/ajcp/aqz108.

13. Nankanja R, Kadhumbula S, Tagoola A, Geisberg M, Serrao E, Balyegyusa S. HemoTypeSC demonstrates > 99% field accuracy in a sickle cell disease screening initiative in children of southeastern Uganda. Am J Hematol. 2019;94(6):E164–E166. doi:10.1002/ajh.25458.

14. Nnodu O, Isa H, Nwegbu, M, Ohiaeri C, Adegoke S, Chianumba R. HemoTypeSC, a low-cost point-of-care testing device for sickle cell disease: promises and challenges. Blood Cells Mol Dis. 2019;78:22–28. doi:10.1016/j.bcmd.2019.01.007.

15. Steele C, Sinski A, Asibey J, Hardy-Dessources MD, Elana G, Brennan C. Point-of-care screening for sickle cell disease in low-resource settings: a multi-center evaluation of HemoTypeSC, a novel rapid test. Am J Hematol. 2019;94(1):39–45. doi:10.1002/ajh.25305.

16. Nwegbu MM, Isa HA, Nwankwo BB, Okeke CC, Edet-Offong UJ, et al. Preliminary evaluation of a point-of-care testing device (SickleSCANTM) in screening for sickle cell disease. Hemoglobin. 2017;41(2):77–82. doi:10.1080/03630269.2017.1329151.

17. Segbena AY, Guindo A, Buono R, Kueviakoe I, Diallo DA, et al. Diagnostic accuracy in field conditions of the sickle SCAN® rapid test for sickle cell disease among children and adults in two West African settings: the DREPATEST study. BMC Hematol. 2018;18:26. doi:10.1186/s12878-018-0120-5.

18. Ryan K, Bain BJ, Worthington D, James J, Plews D, et al. Significant haemoglobinopathies: guidelines for screening and diagnosis. Br J Haematol. 2010;149(1):35–49. doi:10.1111/j.1365-2141.2009.08054.x.

19. Improving the lives of people with sickle cell disease in Ghana. In: Devex [website]. Washington (DC): Devex; 2019 (https://devex.shorthandstories.com/sickle-cell-disease-in-Ghana/index.html, accessed 25 March 2020).

STREPTOCOCCAL PHARYNGITIS

3.4. Group A *Streptococcus* antigen

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a Group A *Streptococcus* (GAS) (*Streptococcus pyogenes*) rapid antigen test to the EDL as an IVD to aid in the diagnosis of streptococcal pharyngitis.

B. Applicant

World Health Organization

C. WHO technical department

Antimicrobial Resistance

D. Background (from application)

Disease condition and impact on patients

Streptococcal pharyngitis, also known as strep throat, is an infection of the back of the throat, including the tonsils, caused by GAS (*Streptococcus pyogenes*). Common symptoms include fever, sore throat, red tonsils and enlarged lymph nodes in the neck. A headache, abdominal pain, and nausea or vomiting may also occur. Symptoms typically begin 1–3 days after exposure and last 7–10 days. Complications that may arise if strep throat is not correctly diagnosed and managed include rheumatic fever with potential long-term heart damage or death, kidney inflammation and peri-tonsillar abscess, sepsis and necrotizing fasciitis.

Does the test meet a medical need?

A rapid antigen detection test (RADT) for GAS pharyngitis enables rapid administration of antibiotics to patients (mostly children between the ages of 5 and 15) to prevent progression to suppurative and non-suppurative complications. Antibiotic treatment does not affect the incidence or outcomes related to some non-suppurative complications such as post-streptococcal glomerulonephritis and arthritis.

The test could also potentially reduce antibiotic prescription in children and adults presenting with signs and symptoms of pharyngitis, by

offering selective treatment to those testing positive and those at higher risk of complications.

How the test is used

The GAS RADT comes in two formats: a lateral flow immunoassay rapid test and an instrument-based digital immunoassay. Both have high specificity and relatively low sensitivity. Guidelines state that a positive result is diagnostic of GAS pharyngitis. In children, a negative result must be confirmed with culture to prevent the risk of infection progressing from strep throat to more serious conditions such as rheumatic fever.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Strep throat is a common bacterial infection in children 5–15 years of age and a rare one in children under 3 years of age. It causes 15–40% of sore throats among children and 5–15% among adults. Cases are more common in late winter and early spring.

A lack of reliable data makes accurately estimating global disease burden difficult. A 2005 review estimated that at least 18.1 million people suffered from invasive GAS diseases, with another 1.78 million incident cases occurring each year (*1*). These estimates did not include the more than 111 million cases of streptococcal pyoderma and 616 million cases of GAS pharyngitis each year.

The health, economic and social burden of GAS pharyngitis (and skin infections) can be significant despite the fact that these diseases are relatively benign. Direct health costs come from antibiotic use, as well as missed schooldays for children and workdays for their parents; but their causal association with invasive infection also has clinical and public health implications. A small 2008 study in Boston, USA, estimated that the total cost of GAS pharyngitis among children in the USA ranges from US$ 224 to US$ 539 million per year, half of which is non-medical costs related to school- and workdays missed (*2*).

By far the heaviest burden of GAS infections falls on low- and middle-income countries (LMICs), because children living in crowded or unsanitary conditions are at much higher risk not only of contracting the infections but also of these not being diagnosed correctly and progressing to invasive disease (*3*).

F. WHO or other clinical guidelines relevant to the test

The Infectious Diseases Society of America (IDSA) in its 2012 guidelines recommends the GAS RADT or culture of throat swabs to diagnose GAS pharyngitis in children older than 3 years of age with a suspicion of bacterial pharyngitis. The recommendations reiterate that a positive test result does not

require culture; but that a negative RADT result should be backed up with a throat swab culture, especially in children and adolescents (*4*).

G. Basic test characteristics (from application)

Test formats available	Lateral flow immunochromatographic assay (LFIA), digital immunoassay (DIA)
Specimen types	Throat swab
Equipment required	Fluorescent immunoassay reader, where applicable
Regulatory status	CE-marked or FDA-approved tests available
Availability	Global
Price per test range	US$ 2–10
Instrument price range	US$ 300–1000, where applicable

H. Evidence for diagnostic accuracy (from application)

In 2018, the Canadian Agency for Drugs and Technologies in Health (CADTH) compiled a report to assess the clinical accuracy and utility and economic benefit of RADTs and molecular tests (*5*). The report examined three systematic reviews, one UK randomized clinical trial (RCT) that included an economic analysis, and 23 primary studies across 11 countries and a broad range of income levels. It found that the sensitivity and specificity of immunoassay-based tests was 55–94% and 81–100%, respectively, for children and mixed populations of children and adults tested using immunoassays with culture assays as reference tests.

Lean et al. (*6*) evaluated LFIAs and reported sensitivity and specificity at 84% (95% CI: 80–88) and 96% (95% CI: 94–97), respectively. A 2014 meta-analysis by Stewart et al. (*7*) evaluated both LFIAs and enzyme immunoassays (EIAs) and found these to be accurate in adults but not in children. It reported pooled sensitivity of LFIAs among children ranging from less than 80% (6 of 28 studies) to 90%, and pooled specificity ranging from less than 80% to more than 95%. For EIA-based methods in children, there was less heterogeneity: pooled sensitivity was 86% (95% CI: 79–92) and pooled specificity was 92% (95% CI: 88–95). In adults, pooled sensitivity for LFIAs was 91% (95% CI: 87–94) with pooled specificity of 93% (95% CI: 92–95); and pooled sensitivity for EIAs was 86% (95% CI: 81–91) with pooled specificity of 97% (95% CI: 96–99).

A 2016 systematic review by Cohen et al. (*8*) compared EIAs and OIAs among children with throat culture standard and found a summary sensitivity of 85.6% (95% CI: 83.3–87.6) and summary specificity of 95.4% (95% CI: 94.5–96.2).

I. Evidence for clinical usefulness and impact (from application)

Reduction in antibiotic use

GAS RADTs have a low sensitivity, and thus a low negative predictive value for diagnosing GAS pharyngitis in children. The sensitivity is acceptable for adults. Recent studies have shown that use of GAS RADTs in emergency departments along with clinical scoring reduced antibiotic use in children (*9*) and in hospital settings (*10*).

In adult populations, primary studies recently published show that the test has potential for application in pharmacies and family health clinics to reduce antibiotic prescription and subsequent nucleic acid testing (*11*, *12*).

Prevention of suppurative complications:

No evidence provided or available, although suppurative complications are thought to result from untreated GAS tonsillitis.

Prevention of acute rheumatic fever:

The evidence on GAS pharyngitis diagnosis and treatment preventing acute rheumatic fever has recently been called into question. Developed countries that withhold antibiotics have not seen a resurgence of acute rheumatic fever. And it has been suggested that acute rheumatic fever results from specific M types of GAS rather than all types, so diagnosis with RADT may not directly prevent this burden.

But incidence may vary widely among populations, and data from LMICs are lacking, with the sample size of existing studies called too small.

J. Evidence for economic impact and/or cost–effectiveness (from application)

The Primary Care Streptococcal Management (PRISM) study in the United Kingdom was an RCT done in primary care practices to study the effect of RADT on antibiotic use, compared with clinical scoring (*13*). It found no difference in the two arms and favoured clinical scoring over RADT. The cost/quality-adjusted life-years (QALYs) and cost/change in symptom severity analyses showed clinical scoring alone to be more cost-effective. Cost impact was not the primary outcome in this study. A study done by Kose et al. in Turkey (*10*) investigated the effect of an RADT on the diagnosis of streptococcal pharyngitis. Performing RADT in children with pharyngitis has an important impact on treatment decisions of clinicians, reduction of unnecessary antibiotic prescriptions and antibiotic costs.

K. Ethics, equity and human rights issues (from application)

None identified.

L. Summary of evidence evaluation

The evidence of the diagnostic accuracy of GAS RADT to aid in the diagnosis of GAS pharyngitis has been summarized in three comprehensive systematic reviews, all with similar findings. The evidence base is substantial, and average sensitivity and specificity of the RADTs is 85% and 95%, respectively; this has not been found to vary systematically with test technology – LFIAs, EIAs and OIAs (optical immunoassays) – or age group (adults and children). Although the reviews have revealed methodological weaknesses in the primary studies, these do not appear to influence the observed accuracy. Unexplained heterogeneity in accuracy has been noted, and there may be variations between test brands or versions; but no clear explanations have been given.

Use in patients where the prevalence of GAS is 30% yields positive predictive values of 88% and negative predictive values of 94%. Guidelines suggest that the positive predictive value is high enough to recommend antibiotics. But guidelines also judge that the risk of missing infection in children is high and recommend that those testing negative also have culture to confirm diagnosis within 24–48 h. The negative RADT result is seen as adequate to withhold antibiotics in adults.

Evidence of the impact of RADTs is mixed, and likely to depend on current local practice, although it has been shown to reduce antibiotic use.

M. Summary of SAGE IVD deliberations

GAS pharyngitis is a common and severe condition, especially among children, for which there is no test currently listed in the EDL. Because prompt and adequate treatment improves the prognosis of GAS pharyngitis, rapid diagnosis is a requisite of children's primary care worldwide. The proposed IVDs have enough potential for adequate diagnosis and are both convenient and cost-effective. The current gold standard (throat culture) takes too long to deliver results and is often unavailable in facilities in LMICs.

Importantly, a POC test for GAS pharyngitis could help avoid unnecessary use of antibiotics, thereby contributing to lower health care costs, fewer adverse events and reduced risk of antibiotic resistance. SAGE IVD noted the very recent publication of a systematic review on the impact of GAS RADTs which shows that they reduce antibiotic use by 8% and antibiotic prescription by 25% (*14*).

N. SAGE IVD recommendations

SAGE IVD recommended including the Group A *Streptococcus* antigen test category in the third EDL:

- as a disease-specific IVD for use in community settings and health facilities without laboratories and as a disease-specific IVD for use in clinical laboratories (EDL 3, Sections I.b and II.b);
- using an RDT or immunoassay format;
- to aid in the diagnosis of Group A Streptococcal pharyngitis.

The group clarified the need for the dual listing in Sections I.b and II.b of the EDL, saying that while the RADT can be used in primary care settings, most guidelines require culture confirmation of a negative test (because rapid tests have lower specificities than immunoassays).

SAGE IVD also highlighted the potential for this test to be grossly overused, and it emphasized the value of adding a note on rational use to this entry in the EDL. To this end, while acknowledging that the EDL is designed to be a policy support document rather than a guide to medical use, SAGE IVD recommended considering the inclusion of an additional column in future editions of the EDL, where comments or qualifications of importance can be made.

O. References

1. The current evidence for the burden of group A streptococcal diseases. Geneva: World Health Organization; 2005 (WHO/FCH/CAH/05.07).
2. Pfoh E, Wessels MR, Goldmann D, Lee GM. Burden and economic cost of group A streptococcal pharyngitis. Pediatrics. 2008;121(2):229–234. doi:10.1542/peds.2007-0484.
3. Sanyahumbi AS, Colquhoun S, Wyber R, Carapetis JR. Global disease burden of group A streptococcus. In: Ferretti JJ, Stevens DL, Fischetti VA, editors. Streptococcus pyogenes: basic biology to clinical manifestations. Oklahoma City: University of Oklahoma Health Sciences Center; 2016.
4. Shulman ST, Bisno AL, Clegg HW, Gerber MA, Kaplan EL, et al. Clinical practice guideline for the diagnosis and management of group A streptococcal pharyngitis: 2012 update by the Infectious Diseases Society of America. Clin Infect Dis. 2012;55(10):e86–e102. doi:10.1093/cid/cis629.
5. Banerjee S, Ford C. Rapid tests for the diagnosis of group A streptococcal infection: a review of diagnostic test accuracy, clinical utility, safety, and cost effectiveness. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2018.
6. Lean WL, Arnup S, Danchin M, Steer AC. Rapid diagnostic tests for group A streptococcal pharyngitis: a meta-analysis. Pediatrics. 2014;134(4):771–781. doi:10.1542/peds.2014-1094.
7. Stewart EH, Davis B, Clemans-Taylor BL, Littenberg B, Estrada CA, et al. Rapid antigen group A streptococcus test to diagnose pharyngitis: a systematic review and meta-analysis. PLoS One. 2014;9(11):e111727. doi:10.1371/journal.pone.0111727.
8. Cohen JF, Bertille N, Cohen R, Chalumeau M. Rapid antigen detection test for group A streptococcus in children with pharyngitis. Cochrane Database Syst Rev. 2016;7:CD010502. doi:10.1002/14651858.CD010502.pub2.
9. Bird C, Winzor G, Lemon K, Moffat A, Newton T, et al. A pragmatic study to evaluate the use of a rapid diagnostic test to detect group A streptococcal pharyngitis in children with the aim

of reducing antibiotic use in a UK emergency department. Pediatr Emerg Care. 2018 Jul 24. doi:10.1097/PEC.0000000000001560.

10. Kose E, Kose SS, Akca D, Yildiz K, Elmas C, et al. The effect of rapid antigen detection test on antibiotic prescription decision of clinicians and reducing antibiotic costs in children with acute pharyngitis. J Trop Pediatr. 2016;62(4):308–315. doi:10.1093/tropej/fmw014.
11. Dodd M, Adolphe A, Parada A, Brett M, Culbreath K, et al. Clinical impact of a rapid streptococcal antigen test on antibiotic use in adult patients. Diagn Microbiol Infect Dis. 2018;91(4):339–344. doi:10.1016/j.diagmicrobio.2018.03.018.
12. Demoré B, Tebano G, Gravoulet J, Wilcke C, Ruspini E, et al. Rapid antigen test use for the management of group A streptococcal pharyngitis in community pharmacies. Eur J Clin Microbiol Infect Dis. 2018;37(9):1637–1645. doi:10.1007/s10096-018-3293-8.
13. Little P, Hobbs FDR, Moore M, Mant D, Williamson I, et al. Primary care Streptococcal Management (PRISM) study: in vitro study, diagnostic cohorts and a pragmatic adaptive randomised controlled trial with nested qualitative study and cost-effectiveness study. Health Technol Assess. 2014;18(6):vii–xxv. doi:10.3310/hta18060.
14. Cohen JF, Pauchard JY, Hjelm N, Cohen R, Chalumeau M. Efficacy and safety of rapid tests to guide antibiotic prescriptions for sore throat. Cochrane Database Syst Rev. 2020;6(6):CD012431. doi:10.1002/14651858.CD012431.pub2.

EDL Section II.a: general IVDs for use in clinical laboratories

HAEMATOLOGY

3.5. Cerebrospinal fluid profile: leukocyte count, differential leukocyte count, protein and glucose

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a cerebrospinal fluid (CSF) profile – red blood cells (RBCs), white blood cells (WBCs), glucose protein – by manual and automated analysers to the EDL as an IVD to aid in the diagnosis of bacterial, mycobacterial, fungal and viral meningitis in adults, children and infants.

B. Applicant

Community and Global Neurology Program, Columbia University Irving Medical Center

C. WHO technical department

None

D. Background (from application)

Disease condition and impact on patients

Infectious neurologic disease remains a significant threat to public health, particularly in resource-limited settings. Globally, there are an estimated 2.8 million cases each year, many in sub-Saharan Africa (*1*). Nearly half of these are bacterial meningitis (1.2 million cases). Moreover, although the number of viral meningitis cases has significantly decreased over the past decade, this is also significant at approximately 400 000 annual cases. The prevalence of fungal meningitis often correlates with the prevalence of immunocompromised populations. For example, a recent study of 5500 lumbar punctures in South Africa showed 63% (n = 514) of the abnormal studies were attributable to cryptococcal meningitis (*2*). Other data suggest there are approximately 1 million cases of cryptococcal meningitis yearly among persons living with HIV, with more than 180 000 deaths, 75% of which occur in sub-Saharan Africa (*3*).

All these diseases can have enduring impacts on patients. For example, patients who recover from bacterial meningitis often suffer long-term cognitive deficit, bilateral hearing loss, motor deficit, seizures, visual impairment and hydrocephalus. Less severe problems include behavioural disorders, learning difficulties, unilateral hearing loss, hypotonia and diplopia. Similarly, most patients who recover from mycobacterial meningitis are left with cognitive impairment, motor deficits, optic atrophy and cranial nerve palsies (*4*).

Does the test meet a medical need?

The basic CSF profile – evaluation of CSF WBC count, RBC count, protein and glucose – plays an integral role in diagnosing and managing a broad range of infectious and non-infectious diseases of the central and peripheral nervous system. The test is required to diagnose, and aid in the diagnosis of, bacterial, mycobacterial, fungal and viral meningitis with the help of CSF gram stain, microscopy, polymerase chain reaction (PCR)-based testing and culture.

Incorporating the basic CSF profile into the EDL not only will enable more rapid, accurate and effective management of individual patients with suspected meningoencephalitis but also has the potential to make a major impact on public health internationally. Specifically, it could reduce antimicrobial resistance caused by antibiotic misuse, improve public health monitoring and reduce overall health care costs related to patients with central nervous system (CNS) infections.

How the test is used

Although no single measure is diagnostic, the basic CSF profile (RBCs, WBCs, glucose, protein) in addition to CSF gram stain is used to determine the initial likelihood of bacterial, fungal or viral meningitis and so guide the decision to

begin or continue antibiotic therapy. A combination of pleocytosis, decreased glucose and elevated protein suggest a bacterial etiology for meningitis.

Lymphocytic pleocytosis, normal glucose and elevated protein suggest a viral etiology but may also reflect fungal and mycobacterial meningitis. The results of the CSF profile are therefore not confirmatory in isolation. Several organizational guidelines include algorithms of varying detail for diagnosing and treating meningitis: these begin with evaluating for contraindications to lumbar puncture, followed by lumbar puncture and/or empiric antibiotics.

Results of the basic CSF profile and gram staining consistent with bacterial meningitis indicate continued antibiotic use. Other confirmatory testing, including latex agglutination assays, PCR-based tests and CSF culture, can be used to confirm specific viral and bacterial etiologies.

In children, a meningitis score has been validated to help predict bacterial meningitis (*5*). This score includes five components: positive CSF gram stain, CSF absolute neutrophil count (> 1000 cells/μL), CSF protein (> 80 mg/L), serum WBC (> 10 000 cells/μL) and presence of seizures. The presence of one or more of these components demonstrates a 99.8% sensitivity for bacterial meningitis. The score is not yet used in paediatric guidelines, but it reiterates the importance of the basic CSF profile in predicting bacterial meningitis among children.

Mycobacterial and fungal etiologies require other confirmatory laboratory testing. For cryptococcal meningitis, cryptococcal antigen testing in serum and/or CSF is recommended as well as specialized microscopy using India ink staining to confirm diagnosis. Diagnostic challenges exist around mycobacterial meningitis, but the first step remains the basic CSF profile analysis, followed by acid-fast bacterial culture (Ziehl–Neelsen) and imaging and tissue pathology where possible.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

In the most recent global burden of disease (GBD) study, neurologic disease is the number one group cause of DALYs, causing 12% of global DALYs (*1*). Overall, meningitis contributed 8% of the neurologic-disease DALYs, although this figure varies across income categories and age ranges. For, example, meningitis ranks fourth overall but second in sub-Saharan Africa. Among children 5 years old and younger, meningitis is the primary cause of neurological DALYs, contributing at least 25% of neurologic years of life lost from birth to 25 years of age. Meningitis also causes deaths. All-cause, global mortality attributed to meningitis approaches 320 000 deaths per year. Importantly, mortality estimates more than double in resource-limited settings compared with resource-rich countries (*1*).

Bacterial meningitis affects 1.2 million people each year (*6*). In the 2016 GBD study, bacterial meningitis caused by *Haemophilus influenzae* and meningococcus accounted for 600 cases of meningitis, approximately 200 deaths and 20 years lived with disability per 100 000 in the world (*1*). The GBD study identified bacterial meningitis as the 27th most burdensome condition; but a sub-analysis of rural Burkina Faso ranked it sixth (*7*). Around half (54%) of all cases and 46% of meningitis-associated deaths occur in children under 5 years of age (*1*).

Mortality is ~50% in LMICs compared with 20% in high-income settings (*8*). This is partly because *H. influenzae*, which has a relatively high mortality rate, is rarely seen in regions with good vaccination programmes. The risk of sequelae similarly varies across income settings, being almost double among Africans compared with Europeans (*9*). In terms of specific organisms, major sequelae occur in 25% of patients with pneumococcal infection, 10% with *H. influenzae* and 8% with meningococcal infections.

Viral meningitis similarly varies across income settings. A prospective registry of suspect meningitis in the United Kingdom found that a third of adult patients had viral meningitis, with 55% caused by enteroviruses, 24% herpes simplex virus (HSV), and 19% varicella (*10*). Data on HSV type 1 CNS infections in Africa indicates that the condition overall is more common and virulent than in the industrialized world (*11*). But even in relatively well-resourced settings, the mortality of HSV infections remains high: a recent report across 47 intensive care units (ICUs) in France found a 17% mortality rate among patients with HSV encephalitis, with 54% of survivors having significant sequelae with a modified Rankin score of ≤ 2.6 (*12*).

Mycobacterial meningitis includes CNS tuberculosis (TB), which is the most severe manifestation of TB and accounts for around 1% of all TB cases and 5–10% of all extrapulmonary TB cases. An estimated 100 000 cases of extrapulmonary TB involving CNS occur globally each year. Among HIV-negative individuals, CNS TB is associated with a 20–55% mortality rate, with more advanced disease at the time of diagnosis resulting in higher mortality. For those with HIV infection, mortality is higher at 40–75%. Among survivors of CNS TB, delayed diagnosis contributes to the development of hydrocephalus and infarctions, particularly in the basal ganglia and internal capsule. A US-based study identified sequelae in 54% of survivors, including stroke in 15% and epilepsy in 12% (*13*). Studies from Mexico have shown that clinical outcomes are improved for patients with a bacteriologically confirmed diagnosis (*14*).

Fungal meningitis is relatively rare and is caused by CNS fungal infections. People with weakened immune systems are at higher risk of getting fungal meningitis, and recovery is possible only with timely diagnosis and treatment;

definitive diagnosis requires CSF analyses or biopsy. Delayed diagnoses increase the risk of stroke, hydrocephalus and death. The annual incidence of CNS fungal infections has been rising as the number of immunocompromised people continues to grow, driven first by the HIV epidemic and then by the increase in noncommunicable diseases, including diabetes. The greatest burden of CNS fungal infections is in HIV-infected individuals, among whom cryptococcal meningitis remains one of the most common causes of death, with a mortality rate of 15–39% (*15*, *16*). Among survivors of CNS fungal infections, strokes that occurred during the acute infection or during delayed diagnosis contribute substantially to long-term morbidity (*15*).

F. WHO or other clinical guidelines relevant to the test

Many national and international guidelines incorporate the basic CSF profile into the diagnosis and management of patients with diseases of the central and peripheral nervous system, and even certain systemic diseases. Examples include guidelines from the IDSA and the American Society for Microbiology (*17–19*), the European Federation of Neurological Societies (*20*, *21*), the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) (*22*), and the British Infection Society (*23*, *24*).

In patients with suspected acute bacterial meningitis, practice guidelines recommend or strongly recommend a lumbar puncture as soon as safely possible, a CSF basic profile, and a CSF gram stain or culture (level of evidence: I; grade of recommendation: A) (*18–23*).

Guideline consensus for typical CSF profiles in acute bacterial meningitis includes a WBC count of > 1000 cells/mm^3 (ranging from 1 to 10 000, with neutrophilic predominance), decreased CSF glucose with a CSF : serum glucose ratio of less than 0.4–0.6, and elevated CSF protein (*18–20*).

Guideline consensus for typical CSF profiles in viral meningitis includes a WBC count of 5–1000 cells/mm^3, low to normal glucose with a CSF : serum ratio of > 0.4–0.5 and normal to slightly elevated protein. A normal CSF profile can be used to rule out infectious causes of meningoencephalitis, though this can be seen in rare circumstances, for example in immunocompromised patients and neonates (*8*, *10*). In the case of bacterial meningitis, it is recommended that CSF gram stain and culture (included in the current EDL) be used to help identify specific causative bacterial organisms to help tailor the antibiotic regimen and duration of steroid therapy.

Similarly, practice guidelines recommend CSF analysis to diagnose and treat CNS TB when there is high clinical suspicion based on history, examination or imaging (*23*). While not the gold standard for diagnosis, the basic CSF profile analysis plays an important part in guiding treatment decisions for CNS TB, as the diagnostic yield of CSF microscopy or culture is often low, requires large

volumes of CSF and may take days to weeks to return positive. Guidelines outline expected findings of CSF leukocytosis (lymphocytic predominance, which can be falsely low in HIV-infected patients), elevated protein and a CSF : serum glucose ratio of less than 0.5 (*23*).

Finally, clinical practice guidelines recommend getting a CSF basic profile and CSF culture or gram stain when clinically concerned for health care-associated causes of ventriculitis or shunt infections and CNS abscesses (*19*).

G. Basic test characteristics (from application)

Test format	Manual counting chamber, automated (flow cytometry, photometry or immunoassay)
Specimen types	CSF
Equipment required	Manual: light microscope, counting grid Automated: analyser, refrigerator, cuvettes
Regulatory status	Many FDA-approved
Availability	Global
Price per test range	US$ 9–10
Instrument price range	US$ 500–30 000

H. Evidence for diagnostic accuracy (from application)

No systematic reviews were available at the time of this review.

Khatib et al. (*25*) describe the accuracy of manual analysis of CSF in meningitis. Among a number of clinical and laboratory markers, CSF pleocytosis best differentiated between bacterial meningitis and other etiologies (area under curve > 0.95).

A total of 48 published studies were available comparing manual and automated methods or evaluating automated methods for cellular and biochemical analysis of CSF. Studies comparing manual and automated methods found that semi-automated or fully automated methods:

- are comparable to manual methods for cell count analysis (RBCs and nucleated cells);
- have good agreement with manual methods (kappa of 0.8 or above, and most studies report r^2 of > 0.98) (*26–28*);
- have good inter-assay agreement; and
- save time and costs (*29*).

Studies noted, however, that diagnostic agreement varies with different cell counts (*27*). At low nucleated cell/WBC counts, the agreement was low (< 20 cells/μL in most studies) (*30*, *31*). Furthermore, at low or abnormal differential counts, manual review of cell counts should follow automated counts (*32–34*). Recent updates in analyser technology and software may have improved detection at lower counts (*35*, *36*). But a 2019 study by Zelazowska-Rutkowska et al. (*37*) in children reiterates the need for manual review of all abnormal counts reported on automated analysers.

An initial 1980 study by Warren et al. (*38*) evaluated automated analyser thresholds for CSF glucose and protein and reported very high thresholds. In 2017, Lefrere et al. (*39*) reported CSF protein assessment on analysers with a kappa of 0.93 with index methods; glucose testing also correlates well with reference methods (r^2 of > 0.98). Londeree et al. (*40*) evaluated the POC utility of two biochemical analysers for CSF protein and glucose and concluded that this use correlated well with in-hospital use.

I. Evidence for clinical usefulness and impact (from application)

The basic CSF profile remains the reference standard for initial management of meningitis. Its established diagnostic accuracy allows for immediate decisions regarding antibiotic and antiviral treatment. Consequently, studies focusing on clinical utility are futile, as obtaining a basic CSF profile is established standard practice. Some data do, however, exist regarding the impact of delayed analysis and further research is exploring the prognosis potential of the CSF basic profile after the diagnosis is made.

Evidence of prognostic value

A 2004 nationwide analysis of meningitis patients (696 episodes) in Dutch centres by van de Beek et al. (*41*) found low CSF WBC count to be an independent predictor of poor outcome. Auburtin et al. (*42*) published a multicentre analysis also indicating the prognostic value of CSF cell counts in meningitis (pleocytosis of > 1000/mL correlated with poor outcomes). Julian-Jimenez et al. (*43*) performed a non-systematic review of data from 59 articles and concluded that CSF pleocytosis and low glucose are independent predictors of poor outcome in meningitis.

Evidence of impact in antibiotic administration delay:

Michael et al. (*44*) performed a single-centre analysis of 92 patient records and concluded that delay in lumbar puncture (and therefore subsequent CSF analysis) delays the administration of antibiotics. Proulx et al. (*45*) similarly concluded from a single-centre retrospective study that diagnostic-treatment algorithms

(and, therefore, delay in performing lumbar puncture and CSF analysis) impact antibiotic treatment and mortality.

J. Evidence for economic impact and/or cost–effectiveness (from application)

The cost for basic CSF profile varies globally, as resource expenditures change with different health systems, cultures and sociopolitical contexts. Though there are limited global data, estimates suggest a complete profile in sub-Saharan Africa costs as little as US$ 9 (*46*). As an approximate estimate for monetary cost, the Centers for Medicare and Medicaid Services in the USA cites the cost of the total basic CSF profile at approximately US$ 28.74, with individual components costing: US$ 5.25 for each RBC and WBC count; US$ 4.37 for glucose in body fluid other than blood; and US$ 19.12 for CSF protein assay (*47*). But private laboratories charge approximately US$ 109.50 for the full CSF profile in the USA (*48*). Documented costs do not specify whether costs reflect manual or automatic cell counting.

The most cost-effective approach to caring for CNS infections is vaccination against threatening organisms such as *H. influenzae* (*49*, *50*). Most cost analyses on CNS infections do not provide specifics in terms of pathogen but address meningoencephalitis overall. One exception is a 2006 US-based study, which showed that delayed lumbar puncture was a strong predictor of higher cost of care for meningitis (*51*). In a 2016 study in China, the incremental cost–effectiveness ratio of the Hib vaccine compared with no vaccination was US$ 13 640 at market price, which was less than three times the GDP per capita of China (*52*).

Data on the cost–effectiveness of manual microscopy compared with automated microscopy for assessing CSF cell count are sparse, possibly because manual microscopy remains the standard of care and there is limited use of automated analysers for CSF cell count worldwide. Still, existing data suggest that while automated methods require larger initial investment, the decreased labour and test timing make a quick return). A study in the *International Journal of Laboratory Hematology* (*53*) compared the Fuchs–Rosenthal manual counting method with the Sysmex XE-5000 and found that the average time for each manual cell count was 635 seconds, which was significantly greater than the 85 seconds required for the automated method. The same study reported that the total analytical performance cost (including personnel costs, material expenses, laboratory equipment) for the counting chamber was €6.74 for the manual counting method compared with €1.22 for the XE-5000 analyser (*53*).

A study from sub-Saharan Africa shows that diagnostic algorithms that start with the basic CSF profile to determine the need for further testing can provide more efficient care through better antibiotic stewardship and less expensive upfront testing (*47*).

K. Ethics, equity and human rights issues (from application)

Laboratory capabilities vary widely, depending on country and location; quality of laboratory facilities and clinical care also varies between high- and low-income settings. While there is an ethical issue with the variability of hospital-based infrastructure, the proposed CSF test itself does not pose significant ethical challenges.

There are no current competing interests attached to the use of this test, and if any arise these shall be discussed openly and officially declared. The use of this test is not sponsored by any pharmaceutical company and shall not be to the financial benefit of a sponsor.

This test shall not in any way increase risk of loss of confidentiality, and the samples will be handled using appropriate procedures in the management of human samples as per requirements.

CSF analysis is an essential diagnostic laboratory test that is low cost and fundamental to the diagnosis and treatment of neurological illnesses, including infectious diseases and non-infectious disorders.

L. Summary of evidence evaluation

No systematic reviews of relevant evidence for these tests are available, and there is little direct evidence for their accuracy (relating to disease) and impact.

The basic CSF profile is part of standard practice in evaluating patients suspected of having meningitis. The tests are done as an early part of a strategy and may be used to decide initial therapy. There is strong evidence that early treatment affects outcomes and the case for early diagnosis is strongly made, although it is not clear how the CSF profile affects timing of treatment. There is some evidence (one study) that the CSF profile can help give a first indication of the most likely cause, but accuracy is not high. Further microbiological testing (gram stain and culture) is essential to obtain definitive diagnoses.

M. Summary of SAGE IVD deliberations

CNS infections are life-threatening conditions of public health importance that require rapid and accurate diagnosis and early treatment to decrease morbidity, mortality and late sequelae. There may be a lack of prospective studies on the diagnostic accuracy of using the basic CSF profile in CNS infections; but the test has long been embedded in clinical practice and is included in almost every relevant guideline as an aid to diagnosis.

There are three relevant tests: CSF leukocyte count, CSF differential count, and CSF protein and glucose. Each one uses a different assay format; but in each case the test is usually performed as the first step in a sequence of tests and cannot provide a definitive diagnosis on its own. Nevertheless, SAGE IVD members confirmed that the test has a clear role in clinical practice as there is

no other test available for acute meningitis and it provides an immediate result that can influence patient management decisions.

N. SAGE IVD recommendations

SAGE IVD recommended including the CSF profile (leukocyte count, differential leukocyte count, protein and glucose) test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.a Haematology,);
- using a haemocytometer or automated haematology analyser with body fluid mode (CSF leukocyte count), Wright–Giemsa-stained smears or automated haematology analyser with body fluid mode (CSF differential leukocyte count), automated or semi-automated chemistry analyser (CSF protein and glucose);
- to aid in the diagnosis of bacterial, mycobacterial, fungal and viral meningitis.

SAGE IVD requested the addition of a note to the test category entry in the EDL stating that definitive diagnosis requires microbiological confirmation, including through gram staining, antigen testing, nucleic acid testing and culture.

The group further recommended reviewing all entries in EDL 2 and adding CSF as a specimen type as and where relevant in the next edition of the EDL (EDL 4).

O. References

1. GBD 2016 Meningitis Collaborators. Global, regional, and national burden of meningitis, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. 2018;17(12):1061–1082. doi:10.1016/S1474-4422(18)30387-9.
2. Jarvis JN, Meintjes G, Williams A, Brown Y, Crede T, et al. Adult meningitis in a setting of high HIV and TB prevalence: findings from 4961 suspected cases. BMC Infect Dis. 2010;10(1):67. doi:10.1186/1471-2334-10-67.
3. Park BJ, Wannemuehler KA, Marston BJ, Govender N, Pappas PG, et al. Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. AIDS. 2009;23(4):525–530. doi:10.1097/QAD.0b013e328322ffac.
4. Kalita J, Misra UK, Ranjan P. Predictors of long-term neurological sequelae of tuberculous meningitis: a multivariate analysis. Eur J Neurol. 2007;14(1):33–37. doi:10.1111/j.1468-1331.2006.01534.x.
5. Nigrovic LE, Malley R, Kuppermann N. Meta-analysis of bacterial meningitis score validation studies. Arch Dis Child. 2012;97(9):799–805. doi:10.1136/archdischild-2012-301798.
6. John CC, Carabin H, Montano SM, Bangirana P, Zunt JR, et al. Global research priorities for infections that affect the nervous system. Nature. 2015;527(7578):S178–S186. doi:10.1038/nature16033.
7. Würthwein R, Gbangou A, Sauerborn R, Schmidt CM. Measuring the local burden of disease. A study of years of life lost in sub-Saharan Africa. Int J Epidemiol. 2001;30(3):501–508. doi:10.1093/ije/30.3.501.

8. Kasanmoentalib ES, Brouwer MC, van de Beek D. Update on bacterial meningitis: epidemiology, trials and genetic association studies. Curr Opin Neurol. 2013;26(3):282–288. doi:10.1097/WCO.0b013e328360415c.

9. Edmond K, Clark A, Korczak VS, Sanderson C, Griffiths UK, et al. Global and regional risk of disabling sequelae from bacterial meningitis: a systematic review and meta-analysis. Lancet Infect Dis. 2010;10(5):317–328. doi:10.1016/S1473-3099(10)70048-7.

10. McGill F, Griffiths MJ, Bonnett LJ, Geretti AM, Michael BD, et al. Incidence, aetiology, and sequelae of viral meningitis in UK adults: a multicenter prospective observational cohort study. Lancet Infect Dis. 2018;18(9):992–1003. doi:10.1016/S1473-3099(18)30245-7.

11. Harfouche M, Chemaitelly H, Abu-Raddad LJ. Herpes simplex virus type 1 epidemiology in Africa: systematic review, meta-analyses, and meta-regressions. J Infect. 2019;79(4):289–299. doi:10.1016/j.jinf.2019.07.012.

12. Jaquet P, de Montmollin E, Dupuis C, Sazio C, Conrad M, et al. Functional outcomes in adult patients with herpes simplex encephalitis admitted to the ICU: a multicenter cohort study. Intensive Care Med. 2019;45(8):1103–1111. doi:10.1007/s00134-019-05684-0.

13. Chiang SS, Khan FA, Milstein MB, Tolman AW, Benedetti A, et al. Treatment outcomes of childhood tuberculous meningitis: a systematic review and meta-analysis. Lancet Infect Dis. 2014;14(10):947–957. doi:10.1016/S1473-3099(14)70852-7.

14. Zurcher K, Ballif M, Kiertiburanakul S, Chenal H, Yotebieng M, et al. Diagnosis and clinical outcomes of extrapulmonary tuberculosis in antiretroviral therapy programmes in low- and middle-income countries: a multicohort study. J Int AIDS Soc. 2019;22(9):e25392. doi:10.1002/jia2.25392.

15. Góralska K, Blaszkowska J, Dzikowiec M. Neuroinfections caused by fungi. Infection. 2018;46(4):443–459. doi:10.1007/s15010-018-1152-2.

16. Rajasingham R, Smith RM, Park BJ, Jarvis JN, Govender NP, et al. Global burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. Lancet Infect Dis. 2017;17(8):873–881. doi:10.1016/S1473-3099(17)30243-8.

17. Miller JM, Binnicker MJ, Campbell S, Carroll KC, Chapin KC, et al. A guide to utilization of the microbiology laboratory for diagnosis of infectious diseases: 2018 update by the Infectious Diseases Society of America and the American Society for Microbiology. Clin Infect Dis. 2018;67(6):e1–e94. doi:10.1093/cid/ciy381.

18. Tunkel AR, Hartman BJ, Kaplan SL, Kaufman BA, Roos KL, et al. Practice guidelines for the management of bacterial meningitis. Clin Infect Dis. 2004;39(9):1267–1284. doi:10.1086/425368.

19. Tunkel AR, Rodrigo H, Bhimraj A, Byers K, Kaplan SL, et al. 2017 Infectious Diseases Society of America's clinical practice guidelines for healthcare-associated ventriculitis and meningitis. Clin Infect Dis. 2017;64(6):e34–e65. doi:10.1093/cid/ciw861.

20. Chaudhuri A, Martinez-Martin P, Kennedy PG, Seaton RA, Portegies P, et al. EFNS guideline on the management of community-acquired bacterial meningitis: report of an EFNS task force on acute bacterial meningitis in older children and adults. Eur J Neurol. 2008;15:649–659. doi:10.1111/j.1468-1331.2008.02193.x.

21. Deisenhammer F, Bartos A, Egg R, Gilhus NE, Giovannoni G, et al. Guidelines on routine cerebrospinal fluid analysis. Report from an EFNS task force. Eur J Neurol. 2006;13(9):913–922. doi:10.1111/j.1468-1331.2006.01493.x.

22. Van de Beek D, Cabellos C, Dzupova O, Esposito S, Klein M, et al. ESCMID guideline: diagnosis and treatment of acute bacterial meningitis. Clin Microbiol Infect. 2016;22(S3):S37–S62. doi:10.1016/j.cmi.2016.01.007.

23. Thwaites G, Fisher M, Hemingway C, Scott G, Solomon T, et al. British Infection Society guidelines for the diagnosis and treatment of tuberculosis of the central nervous system in adults and children. J Infect. 2009;59(3):167–187. doi:10.1016/j.jinf.2009.06.011.
24. Connolly ES Jr, Rabinstein AA, Carhuapoma JR, Derdyn CP, Dion J, et al. Guidelines for the management of aneurysmal subarachnoid haemorrhage: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 43(6):1711–1137. doi:10.1161/STR.0b013e3182587839.
25. Khatib U, van de Beek D, Lees JA, Brouwer MC. Adults with suspected central nervous system infection: a prospective study of diagnostic accuracy. J Infect. 2017;74(1):1–9. doi:10.1016/j.jinf.2016.09.007.
26. Aulesa C, Mainar I, Prieto M, Cobos N, Galimany R. Use of the Advia 120 hematology analyzer in the differential cytologic analysis of biological fluids (cerebrospinal, peritoneal, pleural, pericardial, synovial, and others). Lab Hematol. 2003;9:214–224.
27. Buoro S, Apassiti Esposito S, Alessio M, Crippa A, Ottomano C, et al. Automated cerebrospinal fluid cell counts using the new body fluid mode of Sysmex UF-1000i. J Clin Lab Anal. 2016;30:381–391. doi:10.1002/jcla.21866.
28. Hod EA, Brugnara C, Pilichowska M, Sandhaus LM, Luu HS, et al. Automated cell counts on CSF samples: a multicenter performance evaluation of the GloCyte system. Int J Lab Hematol. 2018;40(1):56–65. doi:10.1111/ijlh.12728.
29. Genc S, Dervisoglu E, Omer D, Kucukates E, Omer B, et al. Evaluation of cell counting in body fluids: comparison of two automated hematology analyzers with manual microscopy. Clin Lab. 2016;62:2449–2453. doi:10.7754/Clin.Lab.2016.160529.
30. Boer K, Deufel T, Reinhoefer M. Evaluation of the XE-5000 for the automated analysis of blood cells in cerebrospinal fluid. Clin Biochem. 2009;42:684–691. doi:10.1016/j.clinbiochem.2009.01.025.
31. Glasser L, Murphy CA, Machan JT. The clinical reliability of automated cerebrospinal fluid cell counts on the Beckman-Coulter LH750 and Iris iQ200. Am J Clin Pathol. 2009;131(1):58–63. doi:10.1309/AJCPCHHTZ6UDQ6WF.
32. Perné A, Hainfellner JA, Womastek I, Haushofer A, Szekeres T, et al. Performance evaluation of the Sysmex XE-5000 hematology analyzer for white blood cell analysis in cerebrospinal fluid. Arch Pathol Lab Med. 2012;136(2):194–198. doi:10.5858/arpa.2011-0030-OA.
33. Buoro S, Seghezzi M, Mecca T, Vavassori M, Crippa A, et al. Evaluation of Mindray BC-6800 body fluid mode for automated cerebrospinal fluid cell counting. Clin Chem Lab Med. 2016;54:1799–1810. doi:10.1515/cclm-2015-1092.
34. Seghezzi M, Manenti B, Previtali G, Alessio MG, Dominoni P, et al. Preliminary evaluation of UF-5000 Body Fluid Mode for automated cerebrospinal fluid cell counting. Clin Chim Acta. 2017;473:133–138. doi:10.1016/j.cca.2017.08.026.
35. Fleming C, Russcher H, Brouwer R, Lindemans J, de Jonge R. Evaluation of Sysmex XN-1000 high-sensitive analysis (hsA) research mode for counting and differentiating cells in cerebrospinal fluid. Am J Clin Pathol. 2016;145:299–307. doi:10.1093/ajcp/aqv093.
36. Takemura H, Ai T, Kimura K, Nagasaka K, Takahashi T, et al. Evaluation of cell count and classification capabilities in body fluids using a fully automated Sysmex XN equipped with high-sensitive analysis (hsA) mode and DI-60 hematology analyzer system. PloS One. 2018;13(4):e0195923. doi:10.1371/journal.pone.0195923.
37. Zelazowska-Rutkowska B, Zak J, Wojtkowska M, Zaworonek J, Cylwik B. Use of the Sysmex XT-4000i hematology analyzer in the differentiation of cerebrospinal fluid cells in children. J Clin Lab Anal. 2019;33(3):e22822. doi:10.1002/jcla.22822.

38. Warren K, Kubasik NP, Brody BB, Sine HE, D'Souza JP. The multilayered film analyzer: glucose in serum, plasma, cerebrospinal fluid, and urine; and urea nitrogen in serum and plasma. Clin Chem. 1980;26:133–137.
39. Lefrere B, Plantamura J, Renard C, Ceppa F, Delacour H. Biochemical analysis of cerebrospinal fluid in the laboratories of deployed medical treatment facilities: are Multistix 10 SG strip and iSTAT useful? J R Army Med Corps. 2017;163:397–400. doi:10.1136/jramc-2017-000778.
40. Londeree W, Davis K, Helman D, Abadie J. Bodily fluid analysis of non-serum samples using point-of-care testing with iSTAT and Piccolo analyzers versus a fixed hospital chemistry analytical platform. Hawaii J Med Public Health. 2014;73(9S1):3–8.
41. van de Beek D, de Gans J, Spanjaard L, Weisfelt M, Reitsma JB, et al. Clinical features and prognostic factors in adults with bacterial meningitis. N Engl J Med. 2004;351(18):1849–1859. doi:10.1056/NEJMoa040845.
42. Auburtin M, Wolff M, Charpentier J, Varon E, Le Tulzo Y, et al. Detrimental role of delayed antibiotic administration and penicillin-nonsusceptible strains in adult intensive care unit patients with pneumococcal meningitis: the PNEUMOREA prospective multicenter study. Crit Care Med. 2006;34(11):2758–2765. doi:10.1097/01.CCM.0000239434.26669.65.
43. Julián-Jiménez A, Morales-Casado MI. Usefulness of blood and cerebrospinal fluid laboratory testing to predict bacterial meningitis in the emergency department. Neurologia. 2019;34(2):105–113. doi:10.1016/j.nrl.2016.05.009.
44. Michael B, Menezes BF, Cunniffe J, Miller A, Kneen R, et al. Effect of delayed lumbar punctures on the diagnosis of acute bacterial meningitis in adults. Emerg Med J. 2010;27(6):433–438. doi:10.1136/emj.2009.075598.
45. Proulx N, Fréchette D, Toye B, Chan J, Kravcik S. Delays in the administration of antibiotics are associated with mortality from adult acute bacterial meningitis. QJM. 2005;98(4):291–298. doi:10.1093/qjmed/hci047.
46. Durski KN, Kuntz KM, Yasukawa K, Virnig BA, Meya DB, et al. Cost-effective diagnostic checklists for meningitis in resource-limited settings. J Acquir Immune Defic Syndr. 2013;63(3):e101–e108. doi:10.1097/QAI.0b013e31828e1e56.
47. Current procedural terminology, 4th edition. Baltimore: Centers for Medicare and Medicaid Services; 2019 (https://www.cms.gov/apps/ama/license.asp?file=/Medicare/Medicare-Fee-for-Service-Payment/ClinicalLabFeeSched/Downloads/19CLABQ3.zip, accessed 12 March 2020).
48. LabCorp net fee schedule. Burlington: LabCorp; 2017 (http://www.dbhds.virginia.gov/library/administration/ADM%20-%2004093%20LabCorp%20Pricing%20Sheet%202015-2017.pdf, accessed 12 March 2020).
49. Mangen MJ, Rozenbaum MH, Huijts SM, van Werkhoven CH, Postma DF, et al. Cost-effectiveness of adult pneumococcal conjugate vaccination in the Netherlands. Eur Respir J. 2015;46(5):1407–1416. doi:10.1183/13993003.00325-2015.
50. Ning G, Yin Z, Li Y, Wang H, Yang W. Cost-effectiveness of the Haemophilus influenzae type B vaccine for infants in mainland China. Hum Vaccin Immunother. 2018;14(1):36–44. doi:10.1080/21645515.2017.1385687.
51. Balada-Llasat JM, Rosenthal N, Hasbun R, Zimmer L, Bozzette S, et al. Cost of managing meningitis and encephalitis among infants and children in the United States. Diagn Microbiol Infect Dis. 2019;93(4):349–354. doi:10.1016/j.diagmicrobio.2018.10.012.
52. Cadena-Herrera D, Esparza-De Lara JE, Ramirez-Ibañez ND, López-Morales CA, Pérez NO, et al. Validation of three viable-cell counting methods: manual, semi-automated, and automated. Biotechnol Rep (Amst). 2015;7:9–16. doi:10.1016/j.btre.2015.04.004.

53. Zimmermann M, Ruprecht K, Kainzinger F, Heppner FL, Weimann A. Automated vs. manual cerebrospinal fluid cell counts: a work and cost analysis comparing the Sysmex XE-5000 and the Fuchs-Rosenthal manual counting chamber. Int J Lab Hematol. 2011;33:629–637. doi:10.1111/j.1751-553X.2011.01339.x.

EDL Section II.b: Disease-specific IVDs for clinical laboratories

ASPERGILLOSIS

3.6. *Aspergillus* antigen

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding an Aspergillus-specific antigen detection assay to the EDL as an IVD to diagnose invasive aspergillosis (IA).

B. Applicant

Global Action Fund for Fungal Infections

C. WHO technical department

None

D. Background (from application)

Disease condition and impact on patients

IA is almost always fatal unless diagnosed and treated promptly. It is silent in its initial clinical manifestations, and microbiological cultures are very insensitive. Unless clinical suspicion is very high, and appropriate non-culture-based tests are done, most patients do not survive. IA is frequently an incidental finding at autopsy. For example, it was the most common major error in missed infections in ICU deaths in the United Kingdom (*1*); in Italy, only 11% of cases in AIDS patients were diagnosed in life (*2*). Major improvements in diagnosis are required to improve outcomes, and these are now available for LMICs.

Does the test meet a medical need?

Access to *Aspergillus* antigen is critical to diagnosing IA. As LMICs establish cancer treatment programmes and intensive care, they will face an increasing number of IA cases. These will only be effectively diagnosed through *Aspergillus*

antigen detection, as culture is insensitive, biopsy with histopathology requires high levels of skill and radiology is often non-specific.

How the test is used

IA is a multimodal diagnosis and galactomannan (GM) is a critical component of this, although not the only one. It is therefore challenging to separate out GM's specific role from other tests, particularly imaging, except in ICU/ventilated patients, where it is pivotal.

E. Public health relevance (from application)

Prevalence

The clinical presentation of IA is relatively rapid, ranging from days to a few weeks. Fever is uncommon and symptoms often absent until late in the course. While neutropenia is the best known risk group, IA is also increasingly recognized in other patients, notably those hospitalized with chronic obstructive pulmonary disease (COPD), lung cancer and occasionally other solid tumours, or liver failure (4%), as well as those receiving corticosteroid therapy and transplant recipients (notably allogeneic stem cell and lung, but also renal and heart recipients). In lung transplant recipients, tracheobronchial disease is the earliest and most common manifestation of IA, usually diagnosed when ulcerations or pseudomembranes are noted on surveillance bronchoscopy. IA is present in 5% of renal and hepatic transplant recipients.

In medical ICUs 2–5% of patients develop IA. The widespread use of steroids for COPD exacerbations may contribute to the increased rate of IA, but COPD is an independent risk factor. The IA rate in patients hospitalized for COPD is 1.3% in Spain (*3*) and 3.9% in Southern China (*3*). In patients with lung cancer, the IA attack rate is 2.6% based on data from China (*4*). IA is found in 4% of all patients who die of AIDS, based on a mean of autopsy studies, and in 19% of cases with complicating severe influenza (*5*). All other immunocompromising conditions and the rare non-immunocompromised patients (such as post-influenza) comprise 5% of cases.

According to conservative burden estimates, IA occurs in 13% of acute myeloid leukaemia cases and 50% of all cases in haematology patients.

Aspergillus contamination of air filters in ICUs is linked to nosocomial IA.

Socioeconomic impact

Data from HICs reveal that IA has a high economic cost in ICU patients. Patients with IA have high death rates and longer hospital stays. Early diagnosis is expected to improve outcomes and reduce lengths of hospitalization (*6*).

F. WHO or other clinical guidelines relevant to the test

Guidelines were published by IDSA in 2016 (*7*); ESCMID, the European Confederation of Medical Mycology (ECMM) and the European Respiratory Society (ERS) in 2017 (*8*); and the American Thoracic Society in 2019 (*9*). All recommend the use of serum or bronchoalveolar lavage (BAL) GM antigen – EIA based on rat-derived EB-A2 monoclonal immunoglobulin M (IgM) capture and detector antibody, and EIA as well as LFIAs based on glycoprotein/GM-like antigens, respectively – to screen for and diagnose IA in patients with haematologic malignancies (both neutropenic and non-neutropenic; the sensitivity is lower in the latter group) and haematopoietic stem cell transplant recipients (except those on anti-*Aspergillus* prophylaxis). With a lower strength of recommendation, the guidelines also recommend use in solid organ transplant recipients, and in other patient groups. Optical density index cut-offs recommended for these purposes are 0.5–1.

IDSA and ESCMID-ECMM-ERS both recommend that this test should not be performed on patients receiving anti-mould-prophylaxis.

G. Basic test characteristics (from application)

Test formats available	Enzyme-linked immunosorbent assay (ELISA), LFA
Specimen types	Serum, BAL fluid
Equipment required	ELISA reader, centrifuge, water bath
Regulatory status	CE-marked, FDA-approved
Availability	Global (some commercial IVDs)
Price per test range	US$ 12–50
Instrument price range	~US$ 5000

H. Evidence for diagnostic accuracy (from application)

The GM EIA tests yield quantitative results, with a linear range of measurement that varies with each batch of testing and with the spectrophotometer used. Diagnostic thresholds well within this linear range have been proposed, but there are no universally agreed thresholds for predicting the outcome.

False negative results can be seen in patients receiving antifungal treatment. False positives have also been reported in patients with severe mucositis and in the presence of *Histoplasma* spp. antigen and other fungal antigens in serum or BAL.

Two Cochrane systematic reviews of diagnostic test accuracy have studied GM detection in serum and in BAL using EIA as the index test and

against elaborate EORTC/MSG criteria as reference standard. The first, by Leeflang et al. (*10*) looked at GM detection in serum and included 50 studies, containing 5660 neutropenic patients, of whom 586 had proven or probable IA. When using Oswestry Disability Index (ODI) cut-offs of 0.5, 1.0 and 1.5, the sensitivities of the test were 78% (95% CI: 70–85), 71% (95% CI: 63–78) and 63% (95% CI: 49–78), respectively; the specificities were 85% (95% CI: 78–91), 90% (95% CI: 86–93) and 93% (95% CI: 89–97), respectively.

The second Cochrane systematic review, by de Heer et al. (*11*), looked at 17 studies on GM detection in BAL. The diagnostic performance at ODI cut-off of 0.5 showed sensitivity of 88% (95% CI: 75–100) and specificity 81% (95% CI: 71–91). An ODI cut-off of 1.0 yielded better specificity, i.e. sensitivity of 78% (95% CI: 61–95) and specificity of 93% (95% CI: 87–98). At ODI cut-offs of 1.5 or higher, the heterogeneity in specificity decreased significantly and was invariably > 90%.

LFA tests yield qualitative results. These tests detect different fungal targets, using either the GM antigen or the mouse monoclonal antibody JF5 to detect *Aspergillus*-specific antigens. Systematic reviews have been done on the new LFA tests (*12*), and some landscape reviews are also published (*13*). The systematic review by Pan et al. (*12*) included seven primary studies and showed that pooled sensitivity, specificity and diagnostic odds ratio for the proven/probable versus no IA cases were 86% (95% CI: 76–93), 93% (95% CI: 89–96) and 65.94 (95% CI: 27.21–159.81) in the lateral flow device test using BAL fluid. The performance in BAL was better than in serum. Zhang et al. (*14*) describe performance of the *Aspergillus* glycoprotein LFA in combination with other tests.

I. Evidence for clinical usefulness and impact (from application)

GM is a critical component of the methods used to diagnose IA (clinical signs, imaging signs) and is best used in combination with other diagnostic modalities.

GM testing has a high positive predictive value for diagnosing IA. In immunocompromised patients suspected of having IA, at a cut-off of 1.0, serum GM testing had positive and negative likelihood ratios of 6.6 (95% CI: 3.4–12.5) and 0.24 (95% CI: 0.11–0.5), respectively, with a diagnostic odds ratio (DOR) of 28 (95% CI: 9–83). In BAL, at a cut-off of 1.0, GM positive and negative likelihood ratios were 14.3 (95% CI: 7.2–28.5) and 0.11 (95% CI: 0.04–0.26), respectively, with a DOR of 134 (95% CI: 43–420) (*9*). These data show further evidence of clinical accuracy, but also indicate, in the absence of outcome studies, the role of these IVDs in prompting more diagnosis and thus earlier treatment.

Dabas et al. (*15*) studied serum GM in a heterogenous group of 235 critically ill patients and found that for 37% of patients, IA diagnoses were made earlier with serum GM than with radiology. The review of 13 studies by Zhang

et al. (*14*) also concluded that adding Aspergillus GM or LFA to the diagnostic algorithm confirmed more IA cases.

Serial screening for serum GM in prolonged neutropenia and in allogeneic stem cell transplant recipients (those not receiving mould-active prophylaxis) during the early engraftment phase has a high sensitivity and negative predictive value for IA. But the inability of a negative GM to exclude non-*Aspergillus* mould infections is well known.

Thresholds are not well established to predict outcome; for GM as a prognostic marker, some evidence points to GM levels falling with effective antifungal treatment.

J. Evidence for economic impact and/or cost–effectiveness (from application)

In the absence of a screening test for IA, patients at risk receive antifungal prophylaxis and/or empiric treatment. GM tests can be used to trigger pre-emptive treatment with the potential to influence cost as well as antifungal exposure. In a meta-analysis of testing to screen vs empiric treatment (*16*), using the test resulted in minimal cost savings but significantly reduced exposure to antifungal agents.

In another comparative study (*17*), pre-emptive treatment after GM testing was 5% less cost-effective than empiric treatment and reduced antifungal treatment days.

K. Ethics, equity and human rights issues (from application)

The test targets immunocompromised patients and generates no ethical, equity or human rights concerns.

L. Summary of evidence evaluation

There is substantial evidence of the accuracy of both serum and BAL GM ELISA tests for diagnosing IA in patients receiving intensive chemotherapy or stem cell transplants with high likelihood of neutropenia. Choice of threshold is critical to balance risks of missing cases against false positives. Evidence is from both using the test for diagnosis in those suspected of having infection (for BAL and serum) and using it for surveillance for infection in those at high risk (serum).

There has been little evaluation of the benefits of using GM ELISA tests to improve patient care and outcomes, or of using the test in formal screening or monitoring programmes.

Guideline panels have judged the accuracy and evidence as adequate to recommend using the test across several different patient groups.

There are few published studies of the accuracy of the newer LFA versions of the antigen tests. Available evidence suggests similar performance to the ELISA tests in BAL samples, but estimates of accuracy remain uncertain.

M. Summary of SAGE IVD deliberations

IA is a WHO priority and a very severe disease. Its incidence may not be high, but it is an almost uniformly fatal disease in LMICs, where the mortality rate is mainly affected by the timing and initiation of antifungal therapy. It is a difficult disease to diagnose, and *Aspergillus* antigen testing is imperfect. For example, it has a highly variable sensitivity in serum testing in LFA test formats and the antigen detected (GM) can sometimes cross-react with other species. In all cases, *Aspergillus* antigen tests should only be used when accompanied by a strong clinical suspicion of IA.

The cost per test seems prohibitive, but it likely comes at the benefit of better or more selective use of voriconazole, which is on the EML specifically to treat IA.

The submission contains convincing evidence on the public health relevance of the conditions and the usefulness, performance and accessibility of the test. And clinical practice guidelines around the world already include *Aspergillus* antigen testing in the diagnostic workup for IA.

The test is easy to perform compared with other methods, although availability in LMICs is doubtful.

N. SAGE IVD recommendations

SAGE IVD recommended the inclusion of the *Aspergillus* antigen test category in the third EDL:

- as a disease-specific IVD for clinical laboratories (EDL 3; Section II.b; Aspergillosis);
- using an RDT or immunoassay format;
- to aid in the diagnosis of invasive aspergillosis in immunocompromised patients.

In considering the test formats available, SAGE IVD noted that while both LFA and ELISA devices should be listed in the EDL, the use of LFA devices should be limited to BAL specimens until more data can be obtained on the accuracy of these devices using serum.

SAGE IVD further recommended comparing LFA devices with PCR for EDL 4 as an alternative, because many tertiary institutions already have PCR technologies available.

The group also noted that in the case of this IVD, it is particularly important that the EDL is clearly linked to relevant guidelines so that users of the test are aware of the potential cross-reactivity issues associated with GM assays.

O. References

1. Winters B, Custer J, Galvagno SM, Colantuoni E, Kapoor SG, et al. Diagnostic errors in the intensive care unit: a systematic review of autopsy studies. BMJ Qual Saf. 2012;21(11):894–902. doi:10.1136/bmjqs-2012-000803.
2. Antinori S, Corbellino M, Necchi A, Corradini P, Vismara C, et al. Immune reconstitution inflammatory syndrome associated with Aspergillus terreus pulmonary infection in an autologous stem cell transplant recipient. Transpl Infect Dis. 2010;12(1):64–68. doi:10.1111/j.1399-3062.2009.00460.x.
3. Xu H, Li L, Huang WJ, Wang LX, Li WF, et al. Invasive pulmonary aspergillosis in patients with chronic obstructive pulmonary disease: a case control study from China. Clin Microbiol Infect. 2012;18(4):403–408. doi:10.1111/j.1469-0691.2011.03503.x.
4. Yan X, Li M, Jiang M, Zou LQ, Luo F, Jiang Y. Clinical characteristics of 45 patients with invasive pulmonary aspergillosis: retrospective analysis of 1711 lung cancer cases. Cancer. 2009;115(21):5018–5025. doi:10.1002/cncr.24559.
5. Schauwvlieghe AF, Rijnders BJ, Philips N, Verwijs R, Vanderbeke L, et al. Invasive aspergillosis in patients admitted to the intensive care unit with severe influenza: a retrospective cohort study. Lancet Respir Med. 2018;6(10):782–792. doi:10.1016/S2213-2600(18)30274-1.
6. Baddley JW, Stephens JM, Ji X, Gao X, Schlamm HT, et al. Aspergillosis in intensive care unit (ICU) patients: epidemiology and economic outcomes. BMC Infect Dis. 2013;13:29. doi:10.1186/1471-2334-13-29.
7. Patterson TF, Thompson III GR, Denning DW, Fishman JA, Hadley S, et al. Practice guidelines for the diagnosis and management of aspergillosis: 2016 update by the Infectious Diseases Society of America. Clin Infect Dis. 2016;63(4):e1–e60. doi:10.1093/cid/ciw326.
8. Ullmann AJ, Aguado JM, Arikan-Akdagli S, Denning DW, Groll AH, et al. Diagnosis and management of Aspergillus diseases: executive summary of the 2017 ESCMID-ECMM-ERS guideline. Clin Microbiol Infect. 2018;24:e1–e38. doi:10.1016/j.cmi.2018.01.002.
9. Hage CA, Carmona EM, Epelbaum O, Evans SE, Gabe LM, et al. Microbiological laboratory testing in the diagnosis of fungal infections in pulmonary and critical care practice. An official American Thoracic Society clinical practice guideline. Am J Respir Crit Care Med. 2019;200(5):535–550. doi:10.1164/rccm.201906-1185ST.
10. Leeflang MM, Debets-Ossenkopp YJ, Visser CE, Scholten RJ, Hooft L, et al. Galactomannan detection for invasive aspergillosis in immunocompromised patients. Cochrane Database Syst Rev. 2008;4:CD007394. doi:10.1002/14651858.CD007394.
11. de Heer K, Gerritsen MG, Visser CE, Leeflang MM. Galactomannan detection in broncho-alveolar lavage fluid for invasive aspergillosis in immunocompromised patients. Cochrane Database Syst Rev. 2019;5:CD012399. doi:10.1002/14651858.CD012399.pub2.
12. Pan Z, Fu M, Zhang J, Zhou H, Fu Y, et al. Diagnostic accuracy of a novel lateral-flow device in invasive aspergillosis: a meta-analysis. J Med Microbiol. 2015;64(7):702–707. doi:10.1099/jmm.0.000092.
13. Heldt S, Hoenigl M. Lateral flow assays for the diagnosis of invasive aspergillosis: current status. Current fungal infection reports. 2017;11(2):45–51. doi:10.1007/s12281-017-0275-8.
14. Zhang L, Guo Z, Xie S, Zhou J, Chen G, et al. The performance of galactomannan in combination with 1,3-β-D-glucan or aspergillus-lateral flow device for the diagnosis of invasive aspergillosis: evidences from 13 studies. Diagn Microbiol Infect Dis. 2019;93(1):44–53. doi:10.1016/j.diagmicrobio.2018.08.005.

15. Dabas Y, Mohan A, Xess I. Serum galactomannan antigen as a prognostic and diagnostic marker for invasive aspergillosis in heterogeneous medicine ICU patient population. PloS One. 2018;13(4). doi:10.1371/journal.pone.0196196.
16. Fung M, Kim J, Marty FM, Schwarzinger M, Koo S. Meta-analysis and cost comparison of empirical versus pre-emptive antifungal strategies in hematologic malignancy patients with high-risk febrile neutropenia. PLoS One. 2015;10(11). doi:10.1371/journal.pone.0140930.
17. Walker BS, Schmidt RL, Tantravahi S, Kim K, Hanson KE. Cost-effectiveness of antifungal prophylaxis, preemptive therapy, or empiric treatment following allogeneic hematopoietic stem cell transplant. Transpl Infect Dis. 2019;21(5):e13148. doi:10.1111/tid.13148.

3.7. *Aspergillus* immunoglobulin G antibody

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding an *Aspergillus* immunoglobulin G (IgG) antibody test to the EDL as an IVD for diagnosing and monitoring chronic pulmonary aspergillosis.

B. Applicant

Global Action Fund for Fungal Infections

C. WHO technical department

None

D. Background (from application)

Disease condition and impact on patients

Chronic pulmonary aspergillosis (CPA) is a slowly destructive lung infection, with marked systemic (weight loss, fatigue) and pulmonary features (productive cough, haemoptysis, breathlessness) almost indistinguishable from TB. Chest imaging may reveal nodules, cavities or fungal balls. CPA often presents as "smear-negative" or "GeneXpert-negative" TB (*1*). It usually occurs after a pulmonary insult or chronic problem including TB, sarcoidosis, non-tuberculous mycobacterial infection, pneumothorax, emphysema, COPD, asthma and RA. Occasionally (5–10% of cases), it is observed in people with no history of lung disease and is often mistaken for TB.

Some subtle immune defects, including low CD4 cells (not HIV related), low NK cells, gamma interferon or interleukin-12 production defects, and some

poorly defined genetic defects also predispose to CPA. It is also found in HIV-infected people with slightly different radiological manifestations.

Does the test meet a medical need?

The test is important in the differential diagnosis of chronic pulmonary infections, especially in patients with underlying lung disease.

Aspergillus antibody serology and chest radiograph and/or computed tomography (CT) scanning with accurate interpretation is critical to establishing the correct diagnosis. CPA progression rates vary, but worsening symptoms and lung destruction or fibrosis occur over many months or years. The key diagnostic features are cavitary lung lesions on radiology, sometimes containing a fungal ball (aspergilloma), and *Aspergillus* IgG antibodies.

How the test is used

The diagnosis of CPA requires presence of symptoms for more than 3 months, consistent radiographic features, and microbiological or immunological evidence of *Aspergillus* infection. The last is achieved by performing an *Aspergillus* IgG test on serum.

E. Public health relevance (from application)

Prevalence

There are an estimated 3 million CPA cases worldwide (*2*).

While CPA is regarded as a rare disease in HICs, its burden in LMICs with high incidences of TB is considerable. One estimate for India (*3*) put the 5-year prevalence at 290 147 cases of complicated pulmonary TB, or 24 per 100 000 population; an estimate for Pakistan (*4*) put it at 72 438 cases, or 39 per 100 000 population, including all underlying diseases. A prospective study in Uganda in both HIV-infected and uninfected patients found an annual rate of CPA development of 6.5% in those with a residual cavity at the end of TB treatment, which is typically found in 22–35% of cases (*5*).

Socioeconomic impact

The economic impact of aspergillosis has been measured in US hospitals (*6*) but not in LMICs, where most of the CPA burden can be expected. In US hospitals it has been estimated that costs can be saved through prevention, improved diagnosis and management.

CPA is associated with severe morbidity and mortality; but outcomes can be improved with long-term antifungal therapy or surgery.

F. WHO or other clinical guidelines relevant to the test

ESCMID, ERS and ECMM published clinical guidelines in 2018 (*7*). These recommend using *Aspergillus* IgG antibody testing to diagnose or exclude CPA, in combination with other diagnostic criteria, that is in non-immunocompromised patients with cavitary or nodular pulmonary infiltrate (level of evidence: II; grade of recommendation: A).

G. Basic test characteristics (from application)

Test formats available	Quantitative EIA, ELISA, qualitative LFA
Specimen types	Serum, plasma
Equipment required	ELISA reader
Regulatory status	Many CE-marked
Availability	Not provided
Price per test range	US$ 4–9
Instrument price range	US$ 3000–75 000

H. Evidence for diagnostic accuracy (from application)

The sensitivity and specificity of quantitative *Aspergillus* IgG tests in diagnosing CPA in various primary studies ranged from 71% to 98% and 81% to 100%, respectively. The variation was chiefly due to different cut-off values. Receiver operating characteristic (ROC) curve analyses of commercial assays identified optimal cut-offs that were different from manufacturer-recommended cut-offs, but optimal cut-offs have been proposed in one study (*8*).

The commercial tests available also have variable reproducibility and precision in these primary studies. The tests have a coefficient of variation ranging from 3.4% to 43.7% (*9*), and acceptable precision levels have not been defined.

The tests are designed to detect *Aspergillus fumigatus* antibody, and may have lower sensitivities for detecting CPA caused by non-*fumigatus* species.

I. Evidence for clinical usefulness and impact (from application)

Guidelines recommend using *Aspergillus* IgG to diagnose CPA in non-immunocompromised patients. Mycological cultures may be used in immunocompromised patients, and a comparison of the utility of culture with aspergillus IgG is unavailable. It is, however, generally accepted that cultures can remain negative in patients with CPA. According to one primary study, *Aspergillus* IgG tests were superior to GM in diagnosing CPA (*10*).

Aspergillus IgG levels also decrease in patients with CPA on effective antifungal treatment, and the test's use has been proposed to monitor antifungal effectiveness. But this use has not been validated in reviews or recommended in guidelines.

J. Evidence for economic impact and/or cost–effectiveness (from application)

No data or sensitivity analyses were available.

K. Ethics, equity and human rights issues (from application)

Diagnosis of CPA with the *Aspergillus* IgG antibody test is targeted at several vulnerable populations, including those with past or present TB or underlying pulmonary disease.

L. Summary of evidence evaluation

Diagnosing CPA requires evidence of infection, and there is evidence that measuring IgG response related to aspergillosis infection provides this in a high proportion of cases.

No evidence was provided to support the suitability of IgG measures for monitoring disease progression and treatment success.

M. Summary of SAGE IVD deliberations

There are no systematic reviews on the clinical utility or impact of *Aspergillus* antibody diagnosis. But the IgG antibody test is already in use and is the subject of several studies. The *Aspergillus* IgG assay result also forms part of the Global Action Fund for Fungal Infections' definition of CPA and as such is critical for a CPA diagnosis.

All ICT, plate ELISA, and auto ELISA for *Aspergillus*-specific IgG have sufficient sensitivity and specificity for a differential diagnosis between CPA and other chronic pulmonary disease (including TB).

The data submitted show it to be a simple, robust and easy-to-use test that is particularly adaptable to LMIC contexts, where there are no realistic diagnostic alternatives.

N. SAGE IVD recommendations

SAGE IVD recommended including the *Aspergillus* IgG antibody test category in the third EDL:

- as a disease-specific IVD for clinical laboratories (EDL 3, Section II.b, Aspergillosis);
- using an RDT or immunoassay format;
- to aid in the diagnosis of chronic pulmonary aspergillosis;

The group highlighted the need for more information on RDTs and other tests that are currently in the pipeline, for which existing data are scarce.

O. References

1. Denning DW, Page ID, Chakaya J, Jabeen K, Jude CM, et al. Case definition of chronic pulmonary aspergillosis in resource-constrained settings. Emerg Infect Dis. 2018;24(8). doi:10.3201/eid2408.171312.
2. Denning DW, Pleuvry A, Cole DC. Global burden of chronic pulmonary aspergillosis as a sequel to tuberculosis. Bull World Health Organ. 2011;89:864–872. doi:10.2471/BLT.11.089441
3. Agarwal R, Denning DW, Chakrabarti A. Estimation of the burden of chronic and allergic pulmonary aspergillosis in India. PLoS One. 2014;9:e114745. doi:10.1371/journal.pone.0114745.
4. Jabeen K, Farooqi J, Mirza M, Denning DW, Zafar A. Serious fungal infections in Pakistan. Eur J Clin Microbiol Infect Dis. 2017;36:949–956. doi:10.1007/s10096-017-2919-6.
5. Page ID, Byanyima R, Hosmane S, Onyachi N, Opira C, et al. Chronic pulmonary aspergillosis commonly complicates treated pulmonary tuberculosis with residual cavitation. Eur Resp J. 2019;53. doi:10.1183/13993003.01184-2018.
6. Tong KB, Lau CJ, Murtagh K, Layton AJ, Seifeldin R. The economic impact of aspergillosis: analysis of hospital expenditures across patient subgroups. Int J Infect Dis. 2009;13(1):24–36. doi:10.1016/j.ijid.2008.02.013.
7. Ullmann AJ, Aguado JM, Arikan-Akdagli S, Denning DW, Groll AH, et al. Diagnosis and management of aspergillus diseases: executive summary of the 2017 ESCMID-ECMM-ERS guideline. Clin Microbiol Infect. 2018;24:e1–38. doi:10.1016/j.cmi.2018.01.002.
8. Sehgal IS, Choudhary H, Dhooria S, Aggarwal AN, Garg M, et al. Diagnostic cut-off of aspergillus fumigatus-specific IgG in the diagnosis of chronic pulmonary aspergillosis. Mycoses. 2018;61(10):770–776. doi:10.1111/myc.12815.
9. Page ID, Richardson MD, Denning DW. Comparison of six aspergillus-specific IgG assays for the diagnosis of chronic pulmonary aspergillosis (CPA). J Infect. 2016;72(2):240–249. doi:10.1016/j.jinf.2015.11.003.
10. Kitasato Y, Tao Y, Hoshino T, Tachibana K, Inoshima N, et al. Comparison of aspergillus galactomannan antigen testing with a new cut-off index and aspergillus precipitating antibody testing for the diagnosis of chronic pulmonary aspergillosis. Respirology. 2009;14(5):701–708. doi:10.1111/j.1440-1843.2009.01548.x.

CANCER

3.8. Serum and urine protein electrophoresis for multiple myeloma

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a serum and urine protein electrophoresis (PEP) to the EDL as an IVD for diagnosing multiple myeloma.

B. Applicant

International Patient Organisation for Primary Immunodeficiencies

C. WHO technical department

Noncommunicable Diseases

D. Background (from application)

Note: This submission was made in response to a recommendation by SAGE IVD during its second annual meeting in March 2019. SAGE IVD recommended that a previous submission be supported by further evidence.

Disease condition and impact on patients

According to the International Myeloma Foundation, multiple myeloma is a cancer of the bone marrow plasma cells. It is also called "myeloma" and "plasma cell myeloma." Plasma cells normally make antibodies against infectious agents such as viruses and bacteria. A cancerous or malignant plasma cell is called a myeloma cell. Myeloma is called "multiple" because there are frequently multiple patches or areas in bone marrow where it grows.

Diagnosing multiple myeloma as early as possible is important to reduce the number of potential complications with more advanced myeloma. Myeloma can be slow-moving or more aggressive (*1*).

Does the test meet a medical need?

Myeloma is unique as a cancer because basic diagnostic testing only requires a complete blood cell count with a differential, basic metabolic panel; serum calcium, serum and urine protein electrophoresis; and osseous survey. All of these should be accessible in LMICs (*2*, *3*).

How the test is used

When gammopathy is suspected, and in particular multiple myeloma, the clinician aims to:

- check for monoclonal protein;
- determine the isotype and the concentration of the immunoglobulin;
- assess the response to the therapy and monitor the progression of disease or relapse; and
- consider any progress that has been made in the diagnostic criteria, diagnostic workup, prognosis, and treatment of patients with multiple myeloma (*4*).

PEP is an inexpensive and easy-to-perform medical technique that allows proteins to be separated and identified based on their electrical charges. PEP may display monoclonal, polyclonal or oligoclonal bands, which are usually seen in the γ zone, but which may be seen in proximity of the β band or, rarely, in the α2 region (*5*). PEP aims to screen for gammopathy, hypergammaglobulinaemia, and hypogammaglobulinaemia (*6*). It is performed for a presumptive diagnosis of monoclonal or polyclonal gammopathies or hypogammaglobulinaemia.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Every year, 159 985 new cases of multiple myeloma are diagnosed worldwide. Multiple myeloma is twice as likely to occur in people of African descent. Its incidence is highly variable among countries but has increased uniformly since 1990, with the largest increase in LMICs (*7*–*9*).

F. WHO or other clinical guidelines relevant to the test

The European Society for Medical Oncology (SMO) (*10*), the UK National Institute for Health and Care Excellence (NICE) (*11*), and the Medical Scientific Advisory Group to the Myeloma Foundation of Australia (*12*) have all published relevant guidelines.

The European Myeloma Network (*13*) issued recommendations on what tools to use and when for the diagnosis and monitoring of multiple myeloma. Developed by a panel of clinical experts on multiple myeloma, these recommendations are based on evidence of published data through August 2017. The recommendations, which are aligned with International Myeloma Working Group guidelines, endorse the use of a serum-free light-chain (SFLC) assay to diagnose and monitor patients with oligosecretory disease (grade of recommendation: 2B). Patients with measurable urinary M proteins should, however, be monitored for multiple myeloma through 24 h urine collections.

When albumin is the dominant protein found in the urine, a glomerulopathy (such as amyloid light-chain amyloidosis or light-chain deposition disease) should be excluded. The 24 h urine collection remains important when results are discordant.

The NICE guideline on myeloma diagnosis and management (*11*) was last updated in 2018 and recommends using serum PEP and SFLC assay to confirm the presence of a paraprotein indicating possible myeloma or monoclonal gammopathy of undetermined significance (MGUS). If serum PEP is abnormal, it further recommends using serum immunofixation to confirm the presence of a paraprotein indicating possible myeloma or MGUS. The NICE guideline advises against using serum PEP, serum immunofixation, SFLC assay or urine Bence–Jones protein assessment alone to exclude a diagnosis of myeloma. This is because the sensitivity is not high enough for a negative test result to safely rule out myeloma.

G. Basic test characteristics (from application)

Test formats available	Gel electrophoresis
Specimen types	Serum
Equipment required	Electrophoresis instrument
Regulatory status	None
Availability	Worldwide
Price per test range	~£5 in the United Kingdom; ~US$ 1.66 in India
Instrument price range	Not provided

H. Evidence for diagnostic accuracy (from application)

No specific evidence provided.

I. Evidence for clinical usefulness and impact (from application)

The references listed in Section F (on guidelines) were provided as evidence of the clinical utility of using serum or urine PEP to diagnose myeloma (*10–12*).

In general, early diagnosis limits distress to the family and prevents damage to the patient and waste of health care resources (*14*).

J. Evidence for economic impact and/or cost–effectiveness (from application)

Fatima et al. (*15*), Anand et al. (*16*) and McTaggart et al. (*17*) were quoted in support of the economic impact and/or cost–effectiveness of serum or urine PEP testing to diagnose.

K. Ethics, equity and human rights issues (from application)

None quoted.

L. Summary of evidence evaluation

There is little evidence provided in this portfolio on the diagnostic accuracy measures of serum PEP. Guidelines recommend not using serum PEP, serum immunofixation, SFLC assay or urine electrophoresis (urine Bence–Jones protein assessment) alone to exclude a diagnosis of myeloma. This indicates that sensitivity is not high enough to safely rule out myeloma when the test result is negative.

M. Summary of SAGE IVD deliberations

In 2019, SAGE IVD requested a submission of serum and urine PEP as an IVD for diagnosing multiple myeloma. Multiple myeloma is a common and increasingly prevalent condition that was added as a syndrome to the EML in 2019. The serum and urine PEP is an easy-to-use, robust test that is already being used as part of the work up for multiple myeloma. WHO endorses use of the test as first step for diagnosing the disease; and WHO experts present at the SAGE IVD meeting also suggested that the test has a role in monitoring disease progression.

But SAGE IVD members warned that while serum and urine PEP may be a good test for detecting multiple myeloma, it cannot be used to rule out the disease. Guidelines recommend not using the test alone to exclude a diagnosis of multiple myeloma, with international criteria requiring bone marrow evidence as the primary evidence. Arriving at the correct diagnosis of multiple myeloma requires a combination of investigations, including flow cytometry, bone marrow biopsy and molecular tests such as fluorescence in situ hybridization (FISH) for prognostication. The lack of reference to these other tests in the submission was raised as a major gap by SAGE IVD.

The group also highlighted the lack of evidence submitted on accuracy, use and cost. No diagnostic accuracy measures were provided. Because serum and urine PEP is a general test, it is especially important to get more evidence on its diagnostic accuracy for multiple myeloma. And while it may be clear that the test is of value in the diagnostic pathway, there remains insufficient evidence to indicate how it should be used, or when it is of greatest value.

A detailed cost analysis is also required along with cost comparisons with other commonly used methods.

N. SAGE IVD recommendations

Based on the evidence submitted, SAGE IVD recommended excluding serum and urine PEP as an IVD to diagnose multiple myeloma from the third EDL.

The group requested that the test be resubmitted by a specialist group next year, with additional evidence; it further requested additional submissions for immunofixation and serum free light chain test.

O. References

1. Ramsenthaler C, Osborne TR, Gao W, Siegert RJ, Edmonds PM, et al. The impact of disease-related symptoms and palliative care concerns on health-related quality of life in multiple myeloma: a multi-centre study. BMC Cancer. 2016;16:427. doi:10.1186/s12885-016-2410-2.
2. Fleming KA, Naidoo M, Wilson M, Flanigan J, Horton S, et al. An essential pathology package for low- and middle-income countries. Am J Clin Pathol. 2017;147(1):15–32. doi:10.1093/ajcp/aqw143.
3. Tan D, Chng WJ, Chou T, Nawarawong W, Hwang SY, et al. Management of multiple myeloma in Asia: resource-stratified guidelines. Lancet Oncol. 2013;14(12):e571–e581. doi:10.1016/S1470-2045(13)70404-2.
4. Ting HY, Sthaneshwar P, Bee PC, Shanmugam H, Lim M. Heavy/light chain assay in the monitoring of multiple myeloma. Pathology. 2019;51(5):507–511. doi:10.1016/j.pathol.2019.04.002.
5. Wahed A, Dasgupta A. Monoclonal gammopathy and its detection. In: Wahed A, Dasgupta A. Hematology and coagulation. Amsterdam: Elsevier; 2015.
6. Lassoued K, Coppo P. Conduite à tenir devant une hypogammaglobulinémie. Med Ther. 2001;7(9): 694–9.
7. Cook R. Economic and clinical impact of multiple myeloma to managed care. J Manag Care Pharm. 2008;14(7 Suppl):19–25. doi:10.18553/jmcp.2008.14.S7-A.19.
8. Pineli M, Pinho W, Alvarado CA, Figueroa R, Bucheli E. Multiple myeloma: epidemiology and burden of disease analysis in Latin America. Value in Health. 2017;20(9): A872. doi:10.1016/j.jval.2017.08.2558.
9. Cowan A, Allen C, Barac A, Basaleem H, Bensenor I, et al. Global burden of multiple myeloma. a systematic analysis for the global burden of disease study 2016. JAMA Oncol. 2018;4(9):1221–1227. doi:10.1001/jamaoncol.2018.2128.
10. Moreau P, San Miguel J, Sonneveld P, Mateos MV, Zamagni E, et al. Multiple myeloma: ESMO clinical practice guidelines for diagnosis treatment and follow-up. Ann Oncol. 20171;28(suppl_4):iv52–iv61. doi:10.1093/annonc/mdx096.
11. NICE guideline [NG35]. Myeloma: diagnosis and management. In: NICE/Guidance [website]. London: National Institute for Health and Care Excellence; 2018 (https://www.nice.org.uk/guidance/ng35/chapter/recommendations, accessed 4 December 2020).
12. Quach H, Prince HM (on behalf of the Medical Scientific Advisory Group). Clinical practice guideline on multiple myeloma. Richmond: Myeloma Foundation of Australia; 2017.
13. Caers J, Garderet L, Kortüm KM, O'Dwyer ME, van de Donk NWCJ, et al. European Myeloma Network recommendations on tools for the diagnosis and monitoring of multiple myeloma: what to use and when. Haematologica. 2018;103(11):1772–1784. doi:10.3324/haematol.2018.189159.
14. Gardulf A, Björvell H, Gustafson R, Hammarström L, Smith CI. The life situations of patients with primary antibody deficiency untreated or treated with subcutaneous gammaglobulin infusions. Clin Exp Immunol. 1993;92(2):200–4. doi:10.1111/j.1365-2249.1993.tb03380.x.
15. Fatima S, Ijaz A, Khan DA, Fatima F, Asif N, et al. Cost-effectiveness of screening and confirmatory tests for multiple myeloma in Pakistani population: an audit report. J Coll Physicians Surg Pak. 2018;28(7):560–563. doi:10.29271/jcpsp.2018.07.560.

16. Anand M, Kumar R, Sharma OD. Economical choices in setting up a diagnostic myeloma laboratory. Indian J Pathol Microbiol. 2004 Oct;47(4):506–508.

17. McTaggart MP, Lindsay J, Kearney EM. Replacing urine protein electrophoresis with serum free light chain analysis as a first-line test for detecting plasma cell disorders offers increased diagnostic accuracy and potential health benefit to patients. Am J Clin Pathol. 2013;140(6):890–897. doi:10.1309/AJCP25IHYLEWCAHJ.

3.9. Epidermal growth factor receptor gene mutation

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding an epidermal growth factor receptor (EGFR) PCR to the EDL as an IVD to aid the diagnosis and treatment of non-squamous non-small cell lung carcinoma.

B. Applicant

World Health Organization

C. WHO technical department

Noncommunicable Diseases

D. Background (from application)

Disease condition and impact on patients

Lung cancer is the most commonly diagnosed cancer worldwide and is a major cause of cancer-related morbidity, disability and mortality, with an estimated 2 million new cases and 1.8 million related deaths in 2018 (*1*). The most commonly diagnosed histology type is non-small cell lung carcinoma (NSCLC), followed by small-cell lung carcinoma (SCLC). Various molecular alterations have been detected in NSCLC, which are known to drive tumour progression and clinical pattern. NSCLC can occur in either oncogene-addicted or non-addicted forms, according to the presence of specific genetic mutations driving tumorigenesis and their pharmacological targetability. The molecular characteristics of NSCLC generally orient therapeutic approaches, especially in the advanced setting. The EGFR gene can present a somatic acquired pathogenetic mutation in a significant portion of NSCLC.

EGFR-addicted NSCLC, given its incidence, comprises a high burden and leads to a high death rate. But advances in cancer oncoprotein-directed

treatments means survival and quality of life have both improved. The use of anti-EGFR treatments, including tyrosine kinase inhibitors (TKIs) doubles objective response rates (ORRs) compared with chemotherapy and improves progression-free survival (PFS). The median survival time is nearly 3 years if the patient receives both targeted medicines and chemotherapy, compared with a median survival time of around 10 months for patients receiving chemotherapy alone (*2*).

Does the test meet a medical need?

The main international guidelines for treating advanced NSCLC confirm that testing EGFR to assess its mutational status has crucial therapeutic implications because patients with mEGFR NSCLC can derive significant benefit from treatment with an anti-EGFR TKI.

ESMO (*2*) recommends using EGFR TKIs as the standard of care for first-line treatment of advanced EGFR-mutated NSCLC (level of evidence: I; grade of recommendation: A). EGFR mutation as an oncogenic target has proven a predictive role in NSCLC from multiple phase 3 trials of EGFR-TKIs versus platinum-based chemotherapy. The improvement in ORR and PFS is consistent across all age groups, genders, smoking status and performance status.

According to the WHO cancer medicines working group, the ESMO Magnitude of Clinical Benefit Scale (MCBS) is a useful tool for selecting medicines that are eligible for priority screening in the WHO EML. Based on the results of the LUX-Lung 3 study (*3*), afatinib scores 4/5 on the MCBS v 1.1 for first-line use in metastatic mEGFR-positive NSCLC. The OPTIMAL (*4*) and EURTAC (*5*) studies give erlotinib an MCBS v 1.1 score of 4/5; and the IPASS study (*6*) similarly gives gefitinib an MCBS v 1.1 score of 4/5 for first-line use in metastatic mEGFR NSCLC.

The US National Comprehensive Cancer Network (NCCN) guidelines also recognize the value of the EGFR TKIs gefitinib, erlotinib and afatinib, giving these TKIs scores of 3/5 for efficacy, 3/5 for safety, 5/5 for quality and consistency of evidence, and 2/5 for affordability.

In 2019, based on data of clinical benefit, the WHO Expert Committee on the Selection and Use of Essential Medicines added erlotinib to the complementary list of the EML for first-line treatment of mEGFR advanced NSCLC, identifying afatinib and gefitinib as therapeutically equivalent alternatives (*7*). The committee noted that these medicines show relevant survival benefits for patients and have better toxicity profiles and quality of life compared with chemotherapy. The committee also noted that since these medicines were considered for inclusion on the EML in 2015, generic versions of these medicines are more widely available, as are quality-assured diagnostic molecular tests for EGFR mutations.

How the test is used

The test (EGFR PCR) is used as a stand-alone test, in patients diagnosed with NSCLC, adenocarcinoma or mixed histology with adenocarcinoma components (non-squamous NSCLC). There is currently no role for a pre-screening test with EGFR immunohistochemistry (IHC); nor can EGFR IHC substitute for the molecular assay. Although novel applications of PCR and next-generation sequencing methods can have various applications in this setting, this submission addresses EGFR testing in tumour tissue samples, and not cell-free DNA or other types of "liquid biopsies"; the test can be performed either from samples of the primary tumour or from metastatic sites, as clinically appropriate.

Although some clinical-pathologic features can enrich the population more likely to harbour mEGFR (e.g. women, non- or light smokers), EGFR molecular testing should be offered to all-comer advanced non-squamous NSCLC patients to indicate EGFR-targeted TKI therapy, regardless of specific clinical characteristics (*8*, *2*).

According to ESMO, routine EGFR somatic mutation testing is recommended for all non-squamous tumours in patients with advanced/recurrent disease (level of evidence:1; grade of recommendation: A). ESMO encourages a wide coverage of mutations in exons 18–21, as some are associated with sensitivity to therapies. At a minimum, when resources or material are limited, the most common activating mutations (Exon19del, L858R) should be determined (level of evidence:1; grade of recommendation: A) (*9*).

In 2018, the College of American Pathologists (CAP), the International Association for the Study of Lung Cancer (IASCL), and the Association for Molecular Pathology updated their joint standards for the molecular analysis of lung cancers to guide treatment decisions with targeted inhibitors (*10*). The panel recommends that pathologists use cell blocks or other cytologic preparations as specimens for lung cancer EGFR molecular testing. It further recommends laboratories to use, or to have available at an external reference laboratory, clinical lung cancer biomarker molecular testing assays that can detect molecular alterations in specimens with as little as 20% cancer cells. It clarifies that laboratories should not use total EGFR expression by IHC testing to select patients for mEGFR-targeted TKI therapy as there is no role for IHC against total EGFR protein as a determinant of treatment with an EGFR kinase inhibitor in NSCLC. Mechanistically, the targetable mutations lead to constitutive and uncontrolled activation of the protein kinase EGFR; however, these molecular alterations have no bearing on the extent of expression at the cell surface of EGFR, which is what is detected by the total EGFR immuno-stain. Although EGFR expression by IHC has been performed in some early studies of EGFR TKI, clinical responses were observed across a wide range of EGFR IHC expression, including in NSCLC with absent/weak IHC expression.

The joint standards also make clear that while early studies showed that EGFR-mutated lung cancers responded to treatment with EGFR inhibitors using unmodified Sanger sequencing with a sensitivity limit of 50% tumour cellularity, in practice this is insufficient because many lung cancer samples are small and mostly comprise benign stromal cells. Most of the larger phase 3 clinical trials that confirmed the clinical utility of mEGFR testing used PCR-based methods that were more sensitive than unmodified Sanger sequencing. Given the widespread availability of technology capable of reliably detecting lower-frequency mutational events in small samples, it is no longer appropriate to offer a low-sensitivity test that cannot test tumours with 20–50% tumour content and requires patients to undergo more invasive procedures to procure a suitable tissue sample (*11*).

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Lung cancer is the most commonly diagnosed cancer worldwide; and NSCLC is the most commonly diagnosed histology type. The disease has a huge economic impact, estimated at around US$ 8 billion in lost productivity in Brazil, Russia, India, China and South Africa (*12*). The absence of effective screening programmes combined with the large number of tobacco smokers means that lung cancer diagnoses are made at the advanced stage in nearly two-thirds of cases (*13*). The mutational pattern of NSCLC relative to EGFR varies across regions, with the highest prevalence in the Asia-Pacific region, where up to 76% of patients present with mEGFR, and the lowest prevalence in Oceania (12%). Africa, Europe and North America registered the same rate of EGFR-mutated NSCLC, at around 20%. Regardless of geographic region, prevalence is higher among never-smokers, women, and the adenocarcinoma subtype (*8*).

The recognition of a pathogenetic mutation of EGFR in lung cancer patients implicates a possible change in the algorithm of treatment of 15–75% of patients, according to the region of origin. Implementing EGFR testing in Asia, where the incidence of mEGFR in lung tumours is highest, could considerably address the health issues of 500 000 to a million patients. Enhancing access to TKI could substantially improve the quality of life and increase life expectancy in patients with generally poor outcomes.

F. WHO or other clinical guidelines relevant to the test

The EGFR molecular test is recommended in the clinical algorithm for NSCLC in the WHO 21st EML (*14*). It is also mentioned in 2015 WHO Classification of Tumors of the Lung, Pleura, Thymus and Heart (*15*).

International guidelines for treating advanced NSCLC have been issued by the American Society of Clinical Oncology (ASCO) (*11*), CAP (*10*), ESMO (*9*) and NCCN (*16*). All recommend testing EGFR to assess its mutational

status (see "Does the test meet a medical need?" and "How the test is used" in Section D for details).

G. Basic test characteristics (from application)

Test formats available	Real-time PCR
Specimen types	Formalin-fixed paraffin-embedded lung tumour specimen
Equipment required	Thermal cycler
Regulatory status	FDA-approved, CE-marked, others
Availability	Worldwide
Price per test range	Can$ 170–200
Instrument price range	Around US$ 30 000–50 000

H. Evidence for diagnostic accuracy (from application)

The panel of experts behind the 2018 CAP/IASCL/AMP guidelines (*10*) confirmed the clinical utility of mEGFR testing methods.

The joint panel, endorsed by ASCO and aligned with ESMO and other national and international societies of oncology, also developed a systematic revision of EGFR testing to address clinical validity, utility, diagnostic performance and quality assurance (*10*, *11*).

An FDA non-clinical performance evaluation from one of the commercially available EGFR PCR tests (*17*) showed that the test could detect mutations in EGFR exons 18, 19, 20 and 21 with at least 5% mutation level using the standard input of 50 ng per reaction well.

Rosell et al. (*5*) assessed the clinical performance of the test by comparing it to two reference methods (bidirectional Sanger sequencing and quantitative next-generation sequencing), using 487 specimens from patients with advanced NSCLC. The positive percentage agreement between the test and Sanger sequencing was 96.6% (95% CI: 91.5–98.7); the negative percentage agreement was 88.3% (95% CI: 84.1–91.5), in the detection of exon 19 deletions and L858R mutations in aggregate as presented.

The FDA (*18*) has presented similar clinical studies for other commercially available tests.

I. Evidence for clinical usefulness and impact (from application)

A systematic review was performed and published in 2015 by the Italian medical oncology team of the National Cancer Institute and the Mario Negri Institute for Pharmacological Research IRCCS (*19*). The authors analysed reports of

randomized controlled trials in the peer-reviewed literature and found that all treatments with TKIs had similar efficacy, but differing toxicity. PFS was better for patients receiving TKI compared with chemotherapy; although data on overall survival were inconclusive and no gain was demonstrated for the extensive patients crossing over into the experimental arm.

The role of EGFR TKI first (erlotinib, gefitinib), second (afatinib) and third generation (osimertinib) has been demonstrated in several phase 3 randomized clinical trials, with consistent results. Gefitinib has been tested against platinum-based chemotherapy in four randomized clinical trials (*6*, *20–22*); erlotinib in two trials (*4*, *5*) and afatinib in one trial (*3*) and six studies (*23*). The trials enrolled both Asian and non-Asian populations, providing consistent results across different ethnicities. The use of TKI in the controlled trials was associated with improved ORRs, with 55–85% of patients reaching a partial or complete radiological response. Trials also reported an increase of disease control, with longer PFS rates: for example, the exposure to TKI is associated with a median PFS of 9–13 months, which is up to 8.5 months longer than chemotherapy.

Data on quality of life have also confirmed the benefit of front-line TKI compared with chemotherapy (*2*).

J. Evidence for economic impact and/or cost–effectiveness (from application)

Whiting et al. (*24*) did a systematic review and cost–effectiveness analysis of mEGFR testing in NSCLC, in which two health economists independently assessed studies for inclusion. The long-term costs and QALYs were estimated using a Markov model with a cycle time of 21 days (resembling the duration of one cycle of chemotherapy) and a time horizon of 6 years. The model showed that the PCR-based EGFR testing with the one of the commercially available kits was less costly than direct sequencing of all exons 19–21, with an incremental cost–effectiveness ratio of £32 167.

There are imminent price adjustments that will make the cost of first- and second-generation TKIs comparable in the near future. Data on costs, cost comparisons or cost analyses are highly variable across countries. EGFR TKIs are more expensive than standard chemotherapy agents. But they are oral medicines and thus do not require the same level of infrastructure or personnel support to be administered.

K. Ethics, equity and human rights issues

As included in the WHO EML, access to TKI should be prioritized in the eligible population worldwide. Access to more effective and less toxic treatments is a priority in oncology, especially for mEGFR NSCLC patients presenting with a clinical phenotype of "lower risk" group (e.g. non-smokers), who are less likely to be included in lung cancer screening programmes.

EGFR PCR should be made available when diagnostic capacity and pathology laboratories are available and well functioning. WHO lists anti-EGFR TKIs as essential medicines, and their prescription requires an assessment of the EGFR mutational status. Accordingly, efforts must be made to ensure that there are no inequalities among NSCLC patients. When access to the test is not linked to an assured access to TKI, or is entirely out of pocket, it could challenge household incomes of patients and their families.

L. Summary of evidence evaluation

Diagnosis of EGFR mutations is recommended for treatment selection. There is some evidence that measurement with RT-PCR instead of Sanger sequencing (paired with tumour enrichment) and next-generation sequencing (NGS) is possible in a higher proportion of cases (less tumour content needed).

Two FDA reports were provided, but only on analytical sensitivity. No evidence was provided on accuracy measures of RT-PCR in a prospective cohort of patients.

M. Summary of SAGE IVD deliberations

Lung cancer is the most commonly diagnosed cancer worldwide. Among NSCLCs, there is a high prevalence of EGFR mutations, particularly in the Asia-Pacific region, where up to 76% of patients presenting with cancer have one.

Therapies for EGFR-mutated lung cancers (TKIs) show a 60–70% response rate and are associated with a significant improvement in progression-free and overall survival compared with chemotherapy. They also have significantly fewer side effects and are available as oral medications. The first-generation TKIs are now available in generic form and are much more cost accessible. Some are also now included in the complementary list of the EML.

The EGFR PCR test is recognized as the gold standard for detecting EGFR mutations. It does not form part of any guidelines (yet) but is already included in a WHO list for priority medical devices. WHO experts advised that the quality of the specimen is critical to detection, suggesting that specimens need to be buffered to ensure that the DNA does not degrade and there is a PCR reaction.

SAGE IVD noted a co-dependency between access to the EGFR test and treatment. Patients can only get the more effective medicines for EGFR-mutated NSCLC if they can prove their EGFR mutation status with a positive EGFR test result.

N. SAGE IVD recommendations

SAGE IVD recommended including the epidermal growth factor receptor (EGFR) gene mutation test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, Cancer);
- using a nucleic acid test format;
- to aid in the diagnosis and treatment of non-squamous non-small cell lung carcinoma.

The group requested the addition of a note to the test category EDL entry stating that it is only recommended for use in specialized anatomical pathology laboratories.

O. References

1. New global cancer data: GLOBOCAN 2018. In: UICC/News [website]. Geneva: Union for International Cancer Control; 2018 (https://www.uicc.org/news/new-global-cancer-data-globocan-2018, accessed 25 March 2020).
2. Planchard D, Popat S, Kerr K, Novello S, Smit EF, et al. Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up, Ann Oncol. 2018; 29(Suppl 4):iv192–iv237. doi:10.1093/annonc/mdy275.
3. Yang JCH, Schuler MH, Yamamoto N, O'Bryne J, Hirsh V, et al. LUX-Lung 3: a randomized, open-label, phase III study of afatinib versus pemetrexed and cisplatin as first-line treatment for patients with advanced adenocarcinoma of the lung harboring EGFR-activating mutations. J Clin Oncol. 2017;30(18_suppl):LBA7500. doi:10.1200/jco.2012.30.18_suppl.lba7500.
4. Zhou C, Wu YL, Chen G, Feng J, Liu et al. Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. Lancet Oncol. 2011;12(8):735–742. doi:10.1016/S1470-2045(11)70184-X.
5. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. Lancet Oncol. 2012;13(3):239–246. doi:10.1016/S1470-2045(11)70393-X. Epub 2012.
6. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009;361:947–957. doi:10.1056/NEJMoa0810699.
7. The selection and use of essential medicines: report of the on Selection and Use of Essential Medicines, 2019 (including the 21st WHO Model List of Essential Medicines and the 7th WHO Model List of Essential Medicines for Children). Geneva: World Health Organization; 2019 (WHO Technical Report Series, No. 1021).
8. Mok TS, Carbone DP, Hirsch FR, editors. IASCL atlas of EGFR testing in lung cancer. Aurora: International Association for the Study of Lung Cancer; 2017.
9. Kerr KM, Bubendorf L, Edelman MJ, Marchetti A, Mok T, et al. Second ESMO consensus conference on lung cancer: pathology and molecular biomarkers for non-small-cell lung cancer. Ann Oncol. 2014;25(9):1681–1690. doi:10.1093/annonc/mdu145.
10. Lindeman NI, Cagle PT, Aisner DL, Arcila ME, Beasley MB, et al. Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: guideline from the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. Arch Pathol Lab Med. 2018;142(3):321–346. doi:10.5858/arpa.2017-0388-CP.

11. Kalemkerian GP, Narula N, Kennedy EB, Biermann WA, Donington J, et al. Molecular testing guideline for the selection of patients with lung cancer for treatment with targeted tyrosine kinase inhibitors: American Society of Clinical Oncology Endorsement of the College of American Pathologists/International Association for the Study of Lung Cancer/Association for Molecular Pathology clinical practice guideline update. J Clin Oncol. 2018;36(9):911–919. doi:10.1200/JCO.2017.76.7293.
12. Pearce A, Sharp L, Hanly P, Barchuk A, Bray F, et al. Productivity losses due to premature mortality from cancer in Brazil, Russia, India, China, and South Africa (BRICS): a population-based comparison. Cancer Epidemiol. 2018;53:27–34. doi:10.1016/j.canep.2017.12.013.
13. Cheng TYD, Cramb SM, Baade PD, Youlden DR, et al. The international epidemiology of lung cancer: latest trends, disparities, and tumor characteristics. J Thorac Oncol. 2016;11(10):1653–1671. doi:10.1016/j.tho.2016.05.021.
14. World Health Organization model list of essential medicines. Geneva: World Health Organization; 2019 (https://apps.who.int/iris/bitstream/handle/10665/325771/WHO-MVP-EMP-IAU-2019.06-eng.pdf?ua=1, accessed 25 March 2020).
15. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, et al. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. J Thora Oncol. 2015;10(9):1243–1260. doi:10.1097/JTO.0000000000000630.
16. NCCN clinical practice guidelines in oncology: non-small cell lung cancer. Plymouth Meeting: National Comprehensive Cancer Network; 2020 (https://www.nccn.org/professionals/physician_gls/pdf/nscl_blocks.pdf, accessed 25 March 2020).
17. Cobas® EGFR mutation test v2: for in vitro diagnostic use. Silver Spring: US Food and Drug Administration; 2015 (https:// www.accessdata.fda.gov/ cdrh_docs/ pdf12/ P120019S007c. pdf, accessed 25 March 2020).
18. Summary of safety and effectiveness data (SSED) [Therascreen EGFR RCQ PCR Kit (Qiagen N.V.)]. Silver Spring: US Food and Drug Administration; 2013 (https://www.accessdata.fda.gov/cdrh_docs/pdf12/P120022b.pdf, accessed 25 March 2020).
19. Haspinger ER, Agustoni F, Torri V, Gelsomino F, Platania M, et al. Is there evidence for different effects among EGFR-TKIs? Systematic review and meta-analysis of EGFR tyrosine kinase inhibitors (TKIs) versus chemotherapy as first-line treatment for patients harboring EGFR mutations. Crit Rev Oncol Hematol. 2015;94(2):213–227. doi:10.1016/j.critrevonc.2014.11.005.
20. Han JY, Park K, Kim SW, Lee DH, Kim HY, et al. First-SIGNAL: first-line single-agent iressa versus gemcitabine and cisplatin trial in never-smokers with adenocarcinoma of the lung. J Clin Oncol. 2012;30(10):1122–1128. doi:10.1200/JCO.2011.36.8456.
21. Yoshioka H, Mitsudomi T, Morita S, Yatabe Y, Negoro S, et al. Final overall survival results of WJTOG 3405, a randomized phase 3 trial comparing gefitinib (G) with cisplatin plus docetaxel (CD) as the first-line treatment for patients with non-small cell lung cancer (NSCLC) harboring mutations of the epidermal growth factor receptor (EGFR). J Clin Oncol. 2014;32(15_suppl):8117.
22. Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, et al. Acquired Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib Is Associated with a Second Mutation in the EGFR Kinase domain. PLoS Med. 2005;2(3):e73. doi:10.1371/journal.pmed.0020073.
23. Wu YL, Zhou C, Hu CP, Feng J, Lu S, et al. LUX-Lung 6: a randomized, open-label, phase III study of afatinib (A) vs gemcitabine/cisplatin (GC) as first-line treatment for Asian patients (pts) with EGFR mutation-positive (EGFR M+) advanced adenocarcinoma of the lung. Eur Respir J. 2013;42:371.

24. Whiting P, Al M, Westwood M, Ramos IC, Ryder S, et al. Viscoelastic point-of-care testing to assist with the diagnosis, management and monitoring of haemostasis: a systematic review and cost-effectiveness analysis. Health Technol Assess. 2015:19(58);1–228. doi:10.3310/hta19580.

CORONAVIRUS DISEASE

3.10. SARS-CoV-2 nucleic acid test

A. Proposal

The application proposed adding a SARS-CoV-2 NAT to diagnose infection by SARS-CoV-2 in symptomatic and asymptomatic individuals suspected to have been exposed.

B. Applicant

Foundation for Innovative New Diagnostics

C. WHO Technical Department

None

D. Background (from application)

Disease condition and impact on patients

COVID-19 is a result of infection caused by the SARS-CoV-2 virus. COVID-19 infection presents a challenge for clinical diagnosis, as many infected (and infectious) patients may present as asymptomatic (*1–4*). COVID-19 typically presents with flu-like symptoms of respiratory disease from mild to moderate to severe. The common symptoms at early onset include fever, cough, headache, myalgia and fatigue. More severe cases present with symptoms of pneumonia and ARDS, including shortness of breath, confusion, low blood pressure, persistent pain or pressure in the chest, and lethargy. Diarrhoea and bloody sputum are indications of progressive severity, as are patchy shadows and ground-glass opacity observed in chest x-ray and tomography scans. Severe complications include sepsis, respiratory failure, heart failure and septic shock. The median time from exposure to onset of symptoms is 4 days for fever and cough. While asymptomatic cases do not appear to progress to severe cases, mild to moderate cases may deteriorate, depending on the timing and level of care received. Severe cases typically require hospitalization, with deterioration requiring intensive care and mechanical ventilation. The mortality rate for COVID-19 has been described as 9% of severe cases (*3*). SARS-CoV-2 appears to be more infectious than SARS or MERS (Middle East respiratory syndrome), with a longer period of infectivity (as measured by persistent viral shedding) before the onset of symptoms, with transmissions also noted in the days to weeks after symptoms have disappeared (*4*). Age appears to play a significant role in morbidity and

mortality risk for COVID-19 (*2*, *5–7*). As seen with SARS-CoV, very few SARS-CoV-2 deaths have been reported for the paediatric population, in stark contrast to the 14.8% mortality rate observed for patients > 80 years and 8% mortality rate for patients 70–79 years (*7–9*). Older patients are more likely to present as severe cases and are two to three times more likely to develop ARDS and require mechanical ventilation, with an observed 49% mortality rate for critical cases. The majority of younger and paediatric cases present asymptomatic or mild cases of respiratory infection, even with high levels of detected viraemia, though COVID-19 paediatric fatalities have been documented. COVID-19 also disproportionally impacts patients with co-morbidities such as hypertension, diabetes, cardiovascular disease, and chronic pulmonary or liver disease (*4*, *10*, *11*).

Does the test meet a medical need?

The clinical utility of SARS-CoV-2 infection testing lies in early identification and isolation of cases, but also in choosing the right therapeutic approach in a clinical picture that can mimic several other entities (*12*). Although the treatment options for less severe forms of COVID-19 are limited, there is increasing evidence of beneficial treatment of severe cases with dexamethasone (*13*, *14*). Other therapeutic options are the right supportive care, such as optimized ventilation strategies and prevention of secondary infections, as well as anticipation and prevention of complications specific to COVID-19 (e.g. inflammation, renal failure, cardiovascular and neurological complications) (*14*, *15*). The latter aspect is more relevant in HICs with advanced clinical and intensive care capacity. In low-income countries lacking capacity for advanced treatment options, greater emphasis is placed on preventing nosocomial infections and on public health measures such as isolation and quarantine.

Identification of COVID-19 may also limit further diagnostic investigations for other etiologies and eventually limit antibiotic use for empiric treatment of assumed bacterial pneumonia (*16*). In a health care settings, early identification of infected individuals (both patients and health care workers) can prevent unrecognized spread of the virus within an institution. Early identification allows for proper isolation of infected patients, and appropriate use of PPE for health care workers. As the test is mostly based on conventional RT-PCR technology, it does not represent a new or specifically innovative technology. Nonetheless, it has the advantage of ample experience with the method, and standard reagents and laboratory equipment are sufficient to carry out testing. Because SARS-CoV-2 is a global pandemic pathogen, in most areas the positive predictive value (PPV) of a SARS-CoV-2 diagnostic test based on PCR is high, especially for patients in high-risk groups.

How the test is used

According to current WHO interim guidance, wherever possible, suspected active SARS-CoV-2 infections should be tested using a NAT, such as real-time RT-PCR. A positive NAT result is considered a confirmed case. A negative result does not rule out infection in cases where clinical suspicion is high. An additional sample and NAT test are recommended in these patients. For more information, see the WHO interim guidance on diagnostic testing for SARS-CoV-2 at: https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

It is estimated that most transmissions occur in the pre-syndromic phase, and that viral loads peak before or around symptom onset (*17*, *18*). Thus, early identification of infected individuals is essential to prevent further spread. Widely available testing for SARS-CoV-2 with high sensitivity and specificity makes it possible to identify infected individuals early, isolate them and limit transmission. In addition to individual diagnostics, identifying infected health care workers can prevent nosocomial spread and protect populations at high risk, such as those in institutions (e.g. nursing homes). The extent of circulating virus can also inform policy-makers for efficient planning of health care resources. Data on SARS-CoV-2 virus circulation obtained by direct virus detection can further help to adapt public health measures, such as social distancing, mask use and school closures.

Due to the high infectivity of the virus, increased case numbers may indicate the necessity to shut down parts of official social interactions (e.g. gatherings and public transport) to prevent overwhelming health systems with high numbers of severe cases. Testing also helps to identify superspreading events. Although most participants may present asymptomatically, they may then serve as a reservoir for a new local outbreak (as has been observed in meat processing plants, clubs and fitness studios).

F. WHO or other clinical guidelines relevant to the test (from application)

WHO has published specific guidance on diagnostic testing for SARS-CoV-2 at: https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2 (*19*). These guidelines are updated based on the availability of evidence and are rapidly evolving.

Other guidance documents have been published by the CDC. For example, a SARS-CoV-2 testing strategy for non-health-care workplaces (*20*) describes the five populations for which SARS-CoV-2 testing with viral tests (i.e. nucleic acid or antigen tests) is appropriate:

- individuals with signs or symptoms consistent with COVID-19;
- asymptomatic individuals with recent known or suspected exposure to SARS-CoV-2 to control transmission;
- asymptomatic individuals without known or suspected exposure to SARS-CoV-2 for early identification in special settings;
- individuals being tested to determine resolution of infection (i.e. test-based strategy for discontinuation of transmission-based precautions, health care personnel returning to work and discontinuation of home isolation); and
- individuals being tested for purposes of SARS-CoV-2 public health surveillance.

Other documents advise on strategies where testing resources are limited. For example, IDSA (*21*) advocates focusing on hospitalized patients and patients with respiratory disease, health care workers, outpatients and community surveillance. Similar documents are available from the ECDC on testing strategies for SARS-CoV-2 (*22*) as well as on surveillance by testing (*23*).

G. Basic test characteristics (from application)

Test formats available	Nucleic acid amplification
Specimen types	Nasopharyngeal swabs, oropharyngeal swabs, nasal swabs, sputum, BAL fluid
Equipment required	Real-time PCR equipment, nucleic acid extraction equipment (if applicable), other ancillary equipment (brand dependent)
Regulatory status	CE-marked, FDA EUA, WHO EUL, TGA, HC, ANVISA
Availability	Global
Price per test range	Brand and volume dependent: as low as €6/test
Instrument price range	None provided

H. Evidence for diagnostic accuracy (from application)

Several clinical studies have reported on NAT sensitivity and specificity (*24–30*). As NATs are highly specific by design, false positive results rarely occur. Nonetheless, they have been described with the use of some automated systems, and occurred in the beginning of the pandemic due to contamination in manufacturing units for primers and probes (*31*, *32*). False-negative results are mostly related to pre-analytical quality, but initially negative results have also been observed in hospitalized patients in two early studies from China and

Singapore. Note that most hospitalizations occur after the first week of illness, when viral load is highest; lower RNA levels are observed when patients present with pneumonia. Many fewer false negatives have been reported in a larger cohort (*26*). In particular, improved assays and use of multiple targets have resulted in optimized detection.

I. Evidence for clinical usefulness and impact (from application)

Not available.

J. Evidence of economic impact and/or cost–effectiveness (from application)

No peer-reviewed literature on the economic impact of testing was available at the time of submission. However, a few preprint publications on the subject, such as a report on the clinical and economic impact of five SARS-CoV-2 testing strategies in Massachusetts by Neilan et al. (*33*), did find testing to be cost-effective. It appears highly likely that in the light of substantial economic losses due to COVID-19, testing will pay off as part of a containment strategy.

K. Ethics, equity and human rights issues (from application)

COVID-19 has disproportionally affected racial and ethnical minorities. Consequently, accessible testing and control of the pandemic will at least to some extent alleviate the disease burden on these populations (*34*).

L. Summary of evidence evaluation

Many studies have shown that RT-PCR testing has high analytical sensitivity for detecting SARS-CoV-2 viral infection. Because the swabbing process and late swabbing may miss the virus, clinical sensitivity may be lower than the analytical sensitivity, as has been demonstrated in studies which used repeat swabs in those who were initially RT-PCR negative. Because the possibility of false negatives in those with symptoms or known exposure cannot be ruled out, repeated RT-PCR tests should be considered.

Although clinical specificity has been shown to be exceptionally high, no data on specificity were presented in the submission. However, it may be possible to extract relevant data from large COVID-19 prevalence studies, such as the Real Time Assessment of Community Transmission (REACT)[5] in the United Kingdom, as the false positive rate must be less than the total positive rate. These data suggest a specificity of 99.85%.

[5] Real-time assessment of community transmission findings. In: Imperial College London, Faculty of Medicine, Research and impact [website]. London: Imperial College; 2020 (https://www.imperial.ac.uk/medicine/research-and-impact/groups/react-study/real-time-assessment-of-community-transmission-findings/, accessed 3 November 2020).

M. Summary of SAGE IVD deliberations

Given the seriousness of the global SARS CoV-2 pandemic the SAGE IVD recommended listing a NAT for the diagnosis of SARS CoV-2 infection. Although the evidence was preliminary at the time of submission, SAGE IVD recognized the need to make a SARS CoV-2 NAT available and its role as an essential tool in managing the pandemic. NAT remains the assay of choice for diagnosing infection according to international guidelines.

Although most of the evidence provided was collected for RT-PCR tests, SAGE IVD decided to list the assay format as a NAT, thereby allowing countries to consider other types of amplification – such as transcription-mediated amplification – for selection and procurement based on local quality assessment.

In considering the use of POC NATs, SAGE IVD raised some concerns about the lack of evidence available at the time of discussion. SAGE IVD highlighted the rapid advance in evidence generation for SARS CoV-2 testing and recommended this listing be reviewed as and when additional evidence is published.

N. SAGE IVD recommendations

SAGE IVD recommended including the SARS-CoV-2 NAT category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b);
- using a nucleic acid test format;
- to diagnose infection by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in symptomatic and asymptomatic individuals suspected of having been exposed to the virus and for surveillance and confirmation of outbreaks.

The group requested the addition of a note to the test category entry in the EDL stating the listing was based on evidence for RT-PCR tests and other types of nucleic acid amplification require more evidence and should be subject to further review.

O. References

1. Pan American Health Organization, WHO Regional Office for the Americas. Laboratory guidelines for detection and diagnosis of the novel coronavirus (2019-nCoV) infection. Washington (DC): Pan American Health Organization; 2020.
2. Hu Z, Song C, Xu C, Jin G, Chen Y, et al. 2020. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. Sci China Life Sci. 2020;53(5):706–711. doi:10.1007/s11427-020-1661.4.

3. Wang Y, Kang H, Liu X, Tong Z. Combination of RT-qPCR testing and clinical features for diagnosis of COVID-19 facilitates management of SARS-CoV-2 outbreak. J Med Virol. 2020;92(6):538–539. doi:10.1002/jmv.25721.
4. Zhou F, Yu T, Du R, Fan G, Liu Y, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054–1062. doi:10.1016/S0140-6736(20)30566-3.
5. Ziegler CGK, Allon SJ, Nyquist SK, Mbano IM, Miao VN, et al. SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. Cell. 2020;181(5):1016–1035.e19. doi:10.1016/j.cell.2020.04.035.
6. Kam KQ, Yung CF, Cui L, Tzer Pin Lin R, et al. A well infant with coronavirus disease 2019 with high viral load. Clin Infect Dis. 2020;71(15):847–849. doi:10.1093/cid/ciaa201.
7. Yang X, Yu Y, Xu J, Shu H, Xia J, et al. 2020. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. Lancet Respir Med. 2020;8(5):475–481. doi:10.1016/S2213-2600(20)30079-5.
8. Li Q, Guan X, Wu P, Wang X, Zhou L, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med. 2020;382:1199–1207. doi:10.1056/NEJMoa2001316.
9. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA. 20;323(13):1239–1242. doi:10.1001/jama.2020.2648.
10. Guo L, Huang Y, Tu M, Wang S, Chen S, et al. Confusion and thinking on the diagnosis and treatment of patients with negative RT-PCR results for SARS-CoV-2. SSRN Electron J. 2020 (https://ssrn.com/abstract=3551322, accessed 1 December 2020).
11. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? Lancet Respir Med. 2020;8(4):e21. doi:10.1016/S2213-2600(20)30116-8.
12. Huang C, Wang Y, Li X, Ren L, Zhao J, et al. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497–506. doi:10.1016/S0140-6736(20)30183-5.
13. Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, et al. 2020. Remdesivir for the treatment of Covid-19 – preliminary report. N Engl J Med. 2020;383:1813–1826. doi:10.1056/NEJMoa2007764.
14. RECOVERY Collaborative Group; Horby P, Lim WS, Emberson JR, Mafham M, et al. Dexamethasone in hospitalized patients with Covid-19 – preliminary report. N Engl J Med. 2020 Jul 17;NEJMoa2021436. doi:10.1056/NEJMoa2021436.
15. Chen X, Laurent S, Onur OA, Kleineberg NN, Fink GR, 2020. A systematic review of neurological symptoms and complications of COVID-19. J Neurol. 2020 Jul 20:1–11. doi:10.1007/s00415-020-10067-3.
16. Lai CC, Wang CY, Hsueh PR. 2020. Co-infections among patients with COVID-19: The need for combination therapy with non-anti-SARS-CoV-2 agents? J Microbiol Immunol Infect. 2020; 53(4):505–512. doi:10.1016/jmii.2020.05.013.
17. He X, Lau EHY, Wu P, Deng X, Wang J, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med. 2020;26:672–675. doi:10.1038/s41591-020-0869-5.
18. Woelfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, et al. Clinical presentation and virological assessment of hospitalized cases of coronavirus disease 2019 in a travel-associated transmission cluster. medRxiv. 2020.03.05.20030502.

19. Diagnostic testing for SARS-CoV-2: interim guidance (11 September 2020). Geneva: World Health Organization; 2020 (https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2, accessed 1 December 2020).
20. ARS-CoV-2 testing strategy: considerations for non-healthcare workers. Atlanta: US Centers for Disease Control and Prevention; 2020 [website] (https://www.cdc.gov/coronavirus/2019-ncov/community/organizations/testing-non-healthcare-workplaces.html, accessed 4 December 2020).
21. COVID-19 prioritization of diagnostic testing updated: March 17, 2020. Arlington (VA): Infectious Diseases Society of America; 2020 (https://www.idsociety.org/globalassets/idsa/public-health/covid-19-prioritization-of-dx-testing.pdf, accessed 1 December 2020).
22. Testing strategies for SARS-CoV-2. Solna: European Centre for Disease Prevention and Control; 2020 (https://www.ecdc.europa.eu/en/covid-19/surveillance/testing-strategies, accessed 1 December 2020).
23. Strategies for the surveillance of COVID-19, technical report (9 Apr 2020). Solna: European Centre for Disease Prevention and Control; 2020 (https://www.ecdc.europa.eu/en/publications-data/strategies-surveillance-covid-19, accessed 1 December 2020).
24. Nalla AK, Casto AM, Huang MLW, Perchetti GA, Sampoleo R, et al. Comparative performance of SARS-CoV-2 detection assays using seven different primer-probe sets and one assay kit. J Clin Microbiol. 2020;58(6):e00557-20. doi:10.1128/JCM.00557-20.
25. Lieberman JA, Pepper G, Naccache SN, Huang ML, Jerome KR, Greninger AL. Comparison of commercially available and laboratory-developed assays for in vitro detection of sars-cov-2 in clinical laboratories. J Clin Microbiol. 2020;58(8):e00821-20. doi:10.1128/JCM.00821-20.
26. Weissleder R, Lee H, Ko J, Pittet MJ. 2020. COVID-19 diagnostics in context. Sci Transl Med. 2020;12(546):eabc1931. doi:10.1126/scitranslmed.abc1931.
27. Long DR, Gombar S, Hogan CA, Greninger AL, OReilly Shah V, Bryson-Cahn C, et al. Occurrence and timing of subsequent SARS-CoV-2 RTPCR positivity among initially negative patients. medRxiv. doi:2020.05.03.20089151.
28. Fang Y, Zhang H, Xie J, Lin M, Ying L, et al. Sensitivity of chest CT for COVID-19: Comparison to RT-PCR. Radiology. 2020;296(2). doi:10.1148/radiol.2020200432.
29. Lee TH, Lin RJ, Lin RTP, Barkham T, Rao P, et al. Testing for SARS-CoV-2: Can we stop at two? Clin Infect Dis. 2020 Apr 19:ciaa459. doi:10.1093/cid/ciaa459.
30. Wang W, Xu Y, Gao R, Lu R, Han K, et al. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA. 2020;323(18):1843–1844. doi:10.1001/jama.2020.3786.
31. False positive results with BD SARS-CoV-2 reagents for the BD Max system – letter to clinical laboratory staff and health care providers. Silver Spring (MD): Food and Drug Administration; 2020.
32. Mögling R, Meijer A, Berginc N, Bruisten S, Charrel R, et al. Delayed laboratory response to COVID-19 caused by molecular diagnostic contamination. Emerg Infect Dis. 2020;26(8):1944–1946. doi10.3201/eid2608.201843.
33. Neilan AM, Losina E, Bangs AC, Flanagan C, Panella C, Eskibozkurt GE. Clinical impact, costs, and cost-effectiveness of expanded SARS-CoV-2 testing in Massachusetts. medRxiv. 2020.07.23.20160820. doi:10.1101/2020.07.23.20160820.
34. Tai DBG, Shah A, Doubeni CA, Sia IG, Wieland ML. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. Clin Infect Dis. 2020 Jun 20:ciaa815. doi:10.1093/cid/ciaa815.

ENDOCRINE DISORDERS

3.11. Cortisol (total)

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a cortisol levels (serum) test to the EDL as an IVD for diagnosing central (pituitary) or primary adrenal hypo- or hyperfunction, and for screening for central (pituitary) or primary adrenal hypofunction in asymptomatic patients at high risk of developing adrenal insufficiency.

B. Applicant

British Columbia Children's Hospital
Global Pediatric Endocrinology and Diabetes

C. WHO technical department

Sexual and Reproductive Health and Research
Contraception and Fertility Care

D. Background (from application)

Disease condition and impact on patients

Cortisol is an essential hormone; it helps the body respond to stress, such as surgery and illness, and recover from infections. Cortisol also helps maintain blood pressure and cardiovascular functions and regulates the metabolism of proteins, carbohydrates and fatty acids. Aldosterone plays a key role in sodium and potassium balance.

Cortisol is the main glucocorticoid hormone produced by the adrenal gland. It is secreted in response to stimulation of the adrenal gland by adrenocorticotropic hormone (ACTH), produced by the pituitary.

Adrenal hypofunction describes central and primary adrenal insufficiency and is characterized by insufficient secretion of cortisol. It is caused either by insufficient secretion of ACTH (central adrenal insufficiency, usually associated with deficiency in other pituitary hormones) or by a non-functional adrenal gland (primary adrenal insufficiency, most commonly because of autoimmune destruction of the gland or Addison's disease). Patients with adrenal insufficiency, both central and primary, often present with hypotension, anorexia, vomiting, weight loss, fatigue and recurrent abdominal pain. Reproductive complaints

typically occur in women (amenorrhoea, loss of libido, decreased axillary and pubic hair). In primary adrenal insufficiency, hyperpigmentation and salt craving are also usually present. Patients may also manifest neuropsychiatric signs and symptoms. In children, weight loss with failure to thrive as well as hypoglycaemic crisis with seizures can be seen. Biochemical findings include hyponatraemia, hyperkalaemia (for primary adrenal insufficiency) and hypoglycaemia.

Adrenal insufficiency is a life-threatening disorder, which, if not recognized, leads to high morbidity and mortality. Any type of stress in patients with adrenal insufficiency can precipitate an adrenal crisis; the most frequent precipitating factors are gastrointestinal and other infectious diseases. Patients with adrenal crisis usually present with unexplained shock refractory to vasopressors and fluids. Early identification and treatment of adrenal crisis significantly decrease mortality rates during these episodes (*1*).

Treatment of adrenal insufficiency is relatively easy and affordable, using medicines included in the EML. Although treatment improves quality of life and markedly decreases morbidity, there are still challenges.

Adrenal hyperfunction, or Cushing's syndrome, is characterized by excess production of ACTH by the pituitary or by excess production of cortisol directly by the adrenal gland. Although Cushing's syndrome is clinically unmistakable when fully blown, the spectrum of clinical presentation is broad. It affects numerous systems, such as reproductive, dermatologic, metabolic, cardiovascular, musculoskeletal, neuropsychiatric and infectious. Few, if any, features of Cushing's syndrome are unique, but some are more discriminatory than others, including reddish-purple striae, plethora, proximal muscle weakness, easy bruising and unexplained osteoporosis. Other symptoms, such as fatigue, weight gain, depression, diabetes, hypertension and menstrual irregularity are also common in individuals without the disorder, which makes the diagnosis very challenging. In children, weight gain with decreasing growth velocity is noticeable (*2*).

This potentially lethal disorder is associated with significant comorbidities and significantly impaired quality of life. This seems to improve after remission and appears to be significantly correlated with the degree of disease control (*3*, *4*). Untreated, Cushing's syndrome causes severe illness and death. The earliest reports of mortality documented a median survival of 5 years, with most deaths caused by vascular or infectious complications. With modern-day treatments, however, the standard mortality ratio after normalizing cortisol is similar to that of an age-matched population (*5*, *6*).

Does the test meet a medical need?

Determining serum cortisol is required to diagnose both adrenal hypofunction and hyperfunction. It is also useful in screening for adrenal hypofunction in

asymptomatic patients at high risk of developing adrenal insufficiency, including previous long-term exposure to exogenous corticosteroids, as well as pituitary tumours, pituitary surgery, and history of cranial or total body irradiation. Recommendations for testing for adrenal insufficiency and Cushing's syndrome are based on observational evidence of a large treatment effect on morbidity and mortality in patients diagnosed with the condition.

Adrenal hypofunction. In patients with adrenal insufficiency, any type of infection or stress can precipitate an adrenal crisis, leading to unexplained refractory shock with a high mortality rate (*7*, *8*). Treatment consists of glucocorticoid replacement, with hydrocortisone being the first choice, in two to three daily doses. Prednisone can also be used. Once-daily fludrocortisone is also given to patients with primary adrenal insufficiency (*9*). Generally, appropriately diagnosed and treated patients have a good prognosis and can have a normal lifespan, compared with the high mortality rate seen in untreated patients (*10*).

The benefits of using steroids to treat patients with adrenal insufficiency have long been proven. In the 1930s, many case reports documented miraculous recovery in patients with Addison's disease who were treated with hydrocortisone. When congenital adrenal hyperplasia was discovered, chronic administration of hydrocortisone to these infants was found to dramatically reverse hypotension, hypoglycaemia and salt wasting (*11*). Current research is focused on finding different steroid formulations that can mimic the physiological circadian pattern of cortisol secretion, with less frequent dosing, in order to improve quality of life as well as compliance (*12*).

Preventing adrenal crisis, which has a mortality rate of 0.5 per 100 patient years, is important (*13*). It requires timely diagnosis of at-risk patients as well as education of both patients and health professionals (*1*, *14*). Early treatment with parenteral hydrocortisone is life-saving and is recommended for any patient with even suspected adrenal crisis by all expert guidelines (*9*).

Adrenal hyperfunction. Patients with active Cushing's syndrome have a mortality rate that is 1.7–4.8 times greater than the general population. It is associated with significant comorbidities, including hypertension, diabetes, coagulopathy, cardiovascular disease, infections and fractures (*15*, *16*). Treating patients with moderate to severe Cushing's syndrome clearly reduces illness and death. Because Cushing's syndrome tends to progress and severe hypercortisolism is probably associated with a worse outcome, it is likely that early recognition and treatment of mild disease would also reduce the risk of residual morbidity (*17*).

Even though morbidity and mortality rates decrease with treatment, these may still be higher than the general population, even after hypercortisolism is

cured. A recent meta-analysis of seven studies showed that patients with Cushing's disease in whom initial surgical cure was not obtained had higher death rates than the general population, while patients with initial remission did not (*6*). But a multicentre, retrospective cohort study showed that patients in remission for more than 10 years still had a higher risk of overall mortality compared with the general population, particularly from circulatory disease; median survival from cure was still found to be excellent at about 40 years of remission (*5*).

How the test is used

Adrenal hypofunction. A morning cortisol 140 nmol/L is used as a preliminary test suggestive of adrenal insufficiency. A morning cortisol > 400–500 nmol/L rules out adrenal insufficiency. When corticotropin is available (uncommon in many low-resource settings), intravenous (IV) administration of high dose (250 mcg of cosyntropin in adults, 125 mcg in children > 2 years and 15 mcg/kg of body weight in infants and children < 2 years) or low dose (1 mcg of cosyntropin), followed by determination of serum cortisol 30 or 60 min after the injection should be performed. A peak cortisol level of less than 500 nmol/L (or 250 nmol/L in infants) at this time is commonly used to diagnose adrenal insufficiency (assay dependent) (*9*).

In HICs, the ACTH stimulation test is commonly used because it is independent of the time of the day, cosyntropin is usually readily available and there is no limit on the number of cortisol samples to be assessed. In LMICs, early morning cortisol determination is preferred.

Adrenal hyperfunction. The first step in evaluating patients for hypercortisolism is to exclude any exogenous causes, such as administration of corticosteroid medications. After ruling these out, the two most common options for screening are:

- The so-called low-dose dexamethasone test, which can be performed as an outpatient test. 1 mg dexamethasone per os is administered between 23:00 and 24:00 h, and serum cortisol is then determined between 08:00 and 09:00 h the next morning. A post-dexamethasone serum cortisol of less than 50 nmol/L rules out Cushing's syndrome with a sensitivity rate of greater than 95%. This test is easy to perform and is not expensive; dexamethasone is also listed in the EML, albeit for use in palliative care.
- Midnight cortisol (*2*).

Overall, the evidence in adults indicates that both tests have similar performance, so choosing which one to use depends on feasibility and technical aspects (*2*).

This application includes cortisol to diagnose both adrenal hypo- and hyperfunction. In low-resource settings, implementing one test for both conditions will be beneficial in terms of training, quality control, etc. Given the low incidence of Cushing's disease and the fact that it can be properly diagnosed by measuring cortisol after a dexamethasone suppression test or overnight, we would not recommend including other tests to the EDL at the present time.

Urinary free and salivary cortisol are also mentioned as alternative tests to diagnose Cushing's; but these are not usually readily available in LMICs.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Adrenal hypofunction. The prevalence of central adrenal insufficiency is reported to be 150–280 per million population, but is probably underestimated (*7*). It can be permanent (pituitary tumours, cranial injury from irradiation, surgery, trauma, infections) or transient (for weeks, months or even years, secondary to exogenous glucocorticoid withdrawal). This latter etiology has dramatically increased over the past decades. Glucocorticoids are largely used in the general population worldwide (up to 2%), and adrenal insufficiency after discontinuation is not only common but usually unrecognized. There is no glucocorticoid administration form, dosing, treatment duration or underlying disease that could exclude the risk of transient adrenal insufficiency, although higher doses and longer use give the highest risk (*18*). In a meta-analysis evaluating 3753 participants treated with corticosteroids for various conditions (*19*), the proportion of patients with adrenal insufficiency ranged from 4% for nasal administration to more than 50% for intra-articular administration. Stratified by disease, percentages ranged from slightly below 7% for asthma to 60% for haematological malignancies. The risk also varied according to dose and treatment duration.

The prevalence of Addison's disease is 82–144 cases per million (*7*). Autoimmunity is the most common cause in adults; other insults to the adrenal gland that lead to Addison's disease include adrenal haemorrhage, cancer, infections (HIV, syphilis, TB, bacteria) and some medicines. Genetic causes, especially enzyme defects, are the most common cause in children. About half of paediatric cases can be attributed to congenital adrenal hyperplasia (*8*). These diverse causes mean no distinct group of individuals is at increased risk of disease.

Well-informed patients with Addison's disease undergoing currently accepted replacement therapy are considered to have a normal survival rate. But studies still show that mortality of patients with adrenal insufficiency is 1.5–2 times higher than the general population, particularly for patients diagnosed at a young age. Increased mortality in primary adrenal insufficiency is linked

to adrenal crisis and sudden death, as well as cardiovascular, malignant, and infectious diseases (*10*). There is also significant morbidity and impact on quality of life. A model for measuring health burden in patients with congenital adrenal hyperplasia estimated that adrenal crisis results in an average loss of 7.3 years of life, or 9 QALYs (*20*).

Adrenal hyperfunction. The reported total incidence of endogenous Cushing's syndrome varies from 3 to 7 cases per million each year (*21–23*). But these figures likely underestimate the incidence of iatrogenic Cushing's, undiagnosed mild hypercortisolism, and ectopic ACTH syndrome; the actual incidence of Cushing's disease may be as high as 5–25 per million per year. The gender distribution of Cushing's syndrome varies with the cause. Men used to have a three times greater incidence of the ectopic ACTH syndrome, but the increasing incidence of lung cancer in cigarette-smoking women has narrowed that margin (*24*). Women are more likely than men to develop Cushing's disease, as well as either benign or malignant adrenal tumours. Age at presentation varies depending upon the cause of hypercortisolism (*16*). Incidence of ectopic ACTH syndrome increases rapidly after 50 years of age, as does lung cancer. Cushing's disease occurs mainly in women 25–45 years of age (*25*). Adrenal tumours have a bimodal age distribution, with small peaks in the first decade of life and major peaks at approximately 40–50 years of age (*26*). Adrenal carcinoma accounts for half of all cases of childhood Cushing's syndrome.

F. WHO or other clinical guidelines relevant to the test

Adrenal hypofunction. The Endocrine Society clinical practice guidelines for diagnosing and treating primary adrenal insufficiency (*9*) recommends the standard dose IV corticotropin stimulation test over other existing diagnostics tests to establish the diagnosis of adrenal insufficiency. If a corticotropin stimulation test is not feasible, they suggest using a morning cortisol < 140 nmol/L until confirmatory testing with corticotropin stimulation is available.

The Society for Endocrinology Clinical Committee guidelines on emergency management of adrenal crisis in adults (*27*) recommends that, provided the patient is haemodynamically stable, a short ACTH test (serum cortisol at baseline and 30 min after IV injection of 250 mg of cortrosyn) should be done (*27*). The American Association of Family Physicians (*28*) similarly recommends using the ACTH stimulation test to diagnose primary adrenal insufficiency, which is also suggested as the best diagnostic test for children (*29*).

A consensus statement on the diagnosis, treatment and follow-up of patients with primary adrenal insufficiency (*30*) recommends measuring cortisol and ACTH levels as initial tests to diagnose adrenal insufficiency; but it also states that the cosyntropin test should be used to confirm the diagnosis (*30*).

Adrenal hyperfunction. The Endocrine Society clinical practice guidelines for diagnosing Cushing's syndrome (*2*) recommend various options depending on the patient suitability, including measuring urinary free cortisol (UFC) (at least two measurements) or late-night salivary cortisol (two measurements), or using the 1 mg overnight dexamethasone suppression test or the longer low-dose test (2 mg per day for 48 h). Confirmation of any abnormal results should be made by doing another of these recommended tests. In cases with a high pretest probability of Cushing's syndrome but a normal initial test, an additional alternative test has the potential benefit of identifying those with milder disease. Midnight serum cortisol, though not recommended as an initial test, can be used for further confirmation after performing an initial dexamethasone suppression test.

The Mexican Society consensus (*31*) is to perform two tests, such as the dexamethasone suppression test and UFC measurement; followed by midnight serum cortisol if there is a discrepancy between the two (*31*).

G. Basic test characteristics (from application)

Test formats available	Immunoassays of varying formats, for example electrochemiluminescence or chemiluminescence.
Specimen types	Serum, plasma
Equipment required	Automated analyser
Regulatory status	CE-marked and FDA-approved; also approved by other regulatory agencies, including Health Canada
Availability	Global
Price per test range	Reimbursement of US$ 10–21 (in North America)
Instrument price range	US$ 40 000 to US$ 200 000 depending on instrument size (in North America)

H. Evidence for diagnostic accuracy (from application)

No systematic reviews of serum/plasma cortisol assays were identified. The following summarizes accuracy reviews, test comparisons and relevant non-systematic reviews.

Like all immunoassays, serum cortisol assays have some limitations. Recognition of these has led to some re-formulation of assays, and publications of inter-assay comparisons (*32*, *33*). Inter-assay variation can be somewhat mitigated by using laboratory- or method-specific cut-offs and through appropriate interpretative support of cortisol test results (*34*, *35*). Despite these limitations, serum cortisol measurement remains a key component of

hypothalamic pituitary adrenal (HPA) axis assessment and management of patients with adrenal dysfunction.

Adrenal hypofunction. For the ACTH stimulation tests, a meta-analysis by Kazlauskaite et al. (*36*) evaluating the diagnostic accuracy of cortisol measurement in ambulatory subjects with presumed normal sleep–wake cycle following high-dose corticotropin stimulation (250 mcg dose) to identify HPA insufficiency (defined relative to results of an insulin tolerance test or overnight metyrapone suppression test) showed that 30 min post-dose cortisol values of > 833 nmol/L ruled out adrenal insufficiency, while results of < 440 nmol/L were highly predictive of it, with an area under the curve (AUC) of 0.82 (95% CI: 0.78–0.86). Similarly, cortisol levels 30 min after a low-dose corticotropin stimulation (1 mcg dose) of < 440 nmol/L and > 600 nmol/L predicted adrenal insufficiency or a normal reference test, respectively, with an AUC of 0.94 (95% CI: 0.90–0.94). Using the low-dose protocol in low-resource settings may not prove a disadvantage, as it showed higher AUC or similar performance characteristics when compared with the high-dose protocol (*36*, *37*). For screening, basal, fasting morning cortisol levels collected between 08:00 and 10:00 h of 365 nmol/L predicted normal HPA axis function with an AUC of 0.79 (95% CI: 0.75–82) (*36*).

Adrenal hyperfunction. Chiondini et al. (*38*) reported that measuring serum cortisol after overnight suppression with 1 mg of dexamethasone had sensitivities of > 95% and specificities of 85–90%. Elamin et al. (*39*) did a meta-analysis comparing outcomes of an overnight dexamethasone suppression test with a reference standard for diagnosing Cushing's syndrome and yielded a pooled positive likelihood ratio of 11.6 (95% CI: 5.8–23.1) and a negative likelihood ratio of 0.09 (95% CI: 0.05–0.14). Among the tests evaluated, the authors concluded that UFC and the overnight dexamethasone suppression test have the most evidence supporting their use for detection of Cushing's syndrome.

Hawley et al. (*40*) identify several limitations to immunoassays. These include heterogeneity between assays made by different vendors and cross-reactivity with structurally similar compounds, in the presence of altered serum components or acute illness, and with altered protein binding (e.g. in pregnancy). While reference materials for cortisol exist, there are well documented differences between methods of analysis and immunoassay vendors themselves, attributed to differences in antibodies used as well as the means of dissociating cortisol from its binding. Ortiz-Flores et al. (*41*) find that sex-specific and assay-specific serum cortisol cut-off values may improve the diagnostic accuracy but are not commonly used in practice. And Kline et al. (*42*) suggest that recalibration and reformulation of assay leads by vendors require evaluation of performance of new methods; as experienced with the recent release of an improved Cortisol II

assay, stimulation test cut-offs in the literature may warrant closer examination in the context of new assays and specific patient populations.

Struja et al. (*43*) evaluated patients who underwent ACTH stimulation tests and found basal cortisol levels of ≤ 100 and ≥ 450 nmol/L in almost half of patients tested for possible adrenal insufficiency, with high diagnostic accuracy, abolishing the need for formal ACTH testing. This supports measurement of basal cortisol as the first-line test, particularly in low-resource settings. Lopez Schmidt et al. (*44*) similarly found that using basal cortisol upper (285 nmol/L) and lower (98 nmol/L) cut-off points with high sensitivity and specificity can reduce the number of individuals who need provocative tests. Montes et al. (*45*) present similar results.

A subnormal serum cortisol response 30 min after an ACTH stimulation test is a reliable marker of adrenal dysfunction for primary disease. But Ortiz-Flores et al. show that when central adrenal insufficiency is suspected, 60 min serum cortisol measurement improves the diagnostic accuracy of the test. Peechakara et al. (*46*) found no difference between low- and high-dose ACTH tests performance.

As for Cushing's syndrome, Barrou et al. (47) used a cut-off point of 1.9 nmol/L and found that the sensitivity remained at 100% and the specificity was 94%. They clearly recommended it as one of the screening tests for Cushing's syndrome. Tang et al. (*48*) conclude that overnight low-dose dexamethasone suppression test and late-night plasma total cortisol have similar values in the initial diagnosis of Cushing's syndrome, but the dexamethasone suppression test is more convenient in an outpatient environment.

I. Evidence for clinical usefulness and impact (from application)

Adrenal hypofunction. There are no modern studies comparing replacing therapy with no treatment, due to the proven life-threatening consequences of the latter. Treatment is guided by experts' recommendations, aiming to mimic physiological secretion, both on a usual basis and in times of stress. There are also no randomized controlled studies evaluating glucocorticoid dose requirements in patients with adrenal insufficiency during times of increased cortisol need, and so the doses recommended to treat adrenal crisis are largely set on an empirical basis. Glucocorticoid dose is typically based on the severity and duration of the stressor. Current recommendations place a higher value on preventing underdosage than on reducing potential negative effects of short-term overdosage, as there are no clear data on the potential consequences of the latter, and undertreatment can have significant deleterious effects.

Ho and Druce (*49*) did a systematic review involving patients with either primary or adrenal insufficiency and Cushing's syndrome from any cause; and showed that quality of life is reduced in both groups and that while it improves

with treatment, it is not completely reversed. Al Nofal et al. (*12*) did a systematic review of different glucocorticoids regimens, which showed some preliminary low-quality evidence of improved quality of life with new forms such as extended-/dual-release, and continuous subcutaneous forms. But studies could not show any relationship between glucocorticoid type and dose and bone loss or rates of adrenal crisis.

Non-systematic reviews on the treatment of adrenal insufficiency address the benefits of treatment with steroids; but these reviews also identify the challenges, including finding the right dosing and frequency of administration (*50*), lack of complete resolution of morbidity (*51*), the need for education and availability of emergency treatment for adrenal crisis (*52*), and issues affecting (*53*).

There are few studies evaluating the impact of treatment for adrenal insufficiency and adrenal crisis. It is considered standard practice and recommended by every clinical guideline. Current studies focus on dosage and pharmacological forms that could be associated with fewer adverse effects and improve compliance. A small study by Wichers et al. (*54*) on hydrocortisone doses for the treatment of secondary adrenal insufficiency showed that lower doses (15–20 mg/day) had similar effects on well-being as higher doses, avoiding the risk of adverse bone health effects. Another study by Laureti et al. (*55*) showed benefits of thrice-daily administrations of cortisone acetate in patients with primary adrenal insufficiency compared with twice-daily dosing, with increased total UFC excretion and reduced plasma ACTH levels. Mah et al. (*56*) studied steroid dose adjustment by weight or body surface area in patients with primary adrenal insufficiency and showed benefits in terms of more physiological cortisol levels compared to fixed-dose regimens. Johansson et al. (*57*) studied oral dual-release hydrocortisone, and showed benefits including more circadian-based serum cortisol profile and metabolic improvements in body weight, blood pressure and glucose metabolism.

Adrenal hyperfunction. Broersen et al. (*58*) did a systematic review of endoscopic vs microscopic surgery for Cushing's disease and found remission rates of 80% with either technique (with a tendency towards better results for macroadenomas with endoscopic procedure), and short-term mortality below 0.5%. Another study by Petersen et al. (*59*) showed-low quality evidence for benefits of microscopic transsphenoidal surgery, which provided remission rates of 42–96% (median 77.9%); recurrence was 0–47.4% with a median of 11.5%. Ritzel et al. (*60*) did a review on bilateral adrenalectomy showing adequate success with residual cortisol secretion from 3% to 34% and less than 2% relapse. Surgical morbidity was 18% and mortality 3%, but the latter increased to 17% on the first year after surgery, suggesting the need to improve postoperative care . A systematic review and meta-analysis by Broersen et al. (*61*) on medical treatment

for Cushing's showed effective cortisol normalization in a large percentage of patients, supporting medical treatment as a reasonable option when surgery is not available or is non-curative. Therapy with multiple agents led to normal cortisol levels in up to 65% of patients. Another review by Fleseriu and Castinetti (*62*) focused on new drugs and showed promising results for efficacy and safety of current and emerging adrenal steroidogenesis inhibitors, although these still have to be confirmed in larger-scale phase 3 studies. Finally, van Haalen et al. (*63*) showed that mortality in patients with Cushing's disease remains increased even after initial biochemical cure remission. The hypothesis is that this is because of the metabolic consequences of long-term overexposure to cortisol, which may provide support for early diagnosis and treatment.

Clayton et al. (*64*) did a study on mortality and morbidity in Cushing's disease, which also included a meta-analysis of previous reports, and showed a twofold increase in mortality compared with the general population. Patients in remission, however, fare much better and appear not to have a higher death rate; hypertension and diabetes mellitus are risk factors for worse outcomes. Similarly, Hammer et al. (*65*) studied transsphenoidal surgery for Cushing's disease and showed that successful treatment of Cushing's disease is associated with normal long-term survival, as opposed to initial persistent disease, supporting the need for early and aggressive intervention. Another study by Faggiano et al. (*66*) on cardiovascular risk factors in patients with Cushing's showed improvement of various parameters one year after remission, although these were still abnormal compared with healthy controls. And a small study by Davies et al. (*67*) on children with Cushing disease, all of whom were cured with surgery (with or without radiotherapy), found that most achieved an adult height within target. Excess adiposity improved with treatment but was still greater than in the general population.

J. Evidence for economic impact and/or cost–effectiveness (from application)

None provided.

K. Ethics, equity and human rights issues (from application)

The availability of laboratory studies to diagnose adrenal insufficiency should not prevent prompt therapy in an acutely ill patient with possible adrenal crisis, which is a life-threatening condition. If cortisol levels are measured and ACTH stimulation testing is not available, providing steroid replacement therapy and stress dosing to all patients with low or borderline-normal morning cortisol levels is the safest and recommended approach.

Given the low prevalence of Cushing's disease, testing should not be performed unless it is based on reasonable clinical suspicion. Wider availability of the cortisol levels test should not lead to unnecessary testing; false-positive

results, with their attendant costs, are reduced if case detection is limited to individuals with an increased pretest probability of having the disorder.

Adrenal hypofunction. If the cortisol levels test becomes available, it should reduce inequity, especially in resource-limited settings, by helping diagnose adrenal insufficiency and providing steroid replacement therapy and stress dosing in a timely manner to those who need them, and by preventing their misuse in patients with preserved adrenal function.

Adrenal hyperfunction. Cosyntropin is currently not available in many low-income settings. Should the cortisol levels test be added to the EDL, an application for including cosyntropin in the next EML might need to be considered.

L. Summary of evidence evaluation

Cortisol levels are directly related to the presence or absence of hypo- or hyper-adrenal function disorders such as Cushing's syndrome and Addison's disease. Diagnostic accuracy for these conditions appears to be very good (90% sensitivity and specificity).

It should be noted that cortisol testing is a test that, if positive, might warrant further testing to specify the cause of loss of adrenal function that will guide options for treatment.

There are no direct comparisons for impact on health outcomes, though it seems likely that early detection of loss of function of the adrenal glands will facilitate early detection, diagnosis and treatment. This is supported by guidelines, as it is not recommended to postpone treatment until cortisol measurement has been performed, but to start immediately and adjust later based on the results.

Cost–effectiveness cannot be established based on the current literature.

But there is sufficient evidence in the submission to recommend serum cortisol measurement based on the analytic and diagnostic accuracy and likely benefits of early detection and treatment.

Diagnostic accuracy measures depend on the indication for the test (intended role of the test) and cut-off points used.

M. Summary of SAGE IVD deliberations

Adrenal insufficiency and Cushing's disease are two conditions with high morbidity and mortality. Primary diagnosis of both conditions can usually be done by determining cortisol levels, with very good accuracy. Serum/plasma cortisol concentrations are useful both as a single measurement (morning sample for adrenal insufficiency and midnight cortisol for Cushing's syndrome) or as a dynamic test (cosyntropin stimulation test for adrenal insufficiency and overnight dexamethasone test for Cushing's syndrome).

Early detection of both conditions can lead to early treatment and reversal. And treatment for adrenal insufficiency is also included in the EML, including corticosteroids and fludrocortisone.

Cortisol testing is included in the guidelines of many professional societies and is widely used in clinical practice across the world to both diagnose and follow up adrenal insufficiency and Cushing's disease. SAGE IVD noted that interpretation of cortisol levels varies depending on whether you are measuring total or free cortisol; and also emphasized that the cost–effectiveness of the test cannot be established based on the data submitted.

The group also raised concerns about the potential for misuse of the test if it is introduced in the absence of specialized centres. This is not just about having a specialized laboratory but also about having specialist staff available to interpret results and guide treatment.

SAGE IVD also noted that cortisol testing, if positive, may warrant further testing to specify the cause of loss of adrenal function to guide treatment options. The group further noted that different autoanalysers and assays may give different results, although it acknowledged that this is typical of hormone determinations.

Importantly, SAGE IVD highlighted a number of interpretive issues with cortisol testing and raised concerns about the potential for it to be inappropriately ordered or inaccurately interpreted without guidance. In particular, the timing of sample collection and the need for cosyntropin stimulation (or dexamethasone suppression) is often critical for cortisol testing to work effectively.

N. SAGE IVD recommendations

SAGE IVD recommended including the cortisol (total) test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, within a new subsection for endocrine disorders);
- using an immunoassay format;
- to diagnose central (pituitary) or primary (adrenal, Addison's disease) cortisol deficiency;* and
- to diagnose central (pituitary) or primary (adrenal) hypercortisolism (Cushing's syndrome).**

* Often used with timed collection and stimulation with cosyntropin.
** Often used with timed collection and suppression with dexamethasone.

The group further requested the addition of a note to the test category entry in the EDL stating that it is only recommended for use in specialized health care settings.

O. References

1. Meyer G, Badenhoop K. Addisonian crisis – risk assessment and appropriate treatment. Dtsch Med Wochenschr. 2018;143(6):392–396. doi:10.1055/s-0043-111729.
2. Nieman LK, Biller BMK, Findling JW, Newell-Price J, Savage MO, et al. The diagnosis of Cushing's syndrome: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2008;93(5):1526–1540. doi:10.1210/jc.2008-0125.
3. Milian M, Honegger J, Teufel P, Wolf A, Psaras T. Tuebingen CD-25 is a sensitive tool to investigate health-related quality of life in Cushing's disease patients in the course of the disease. Neuroendocrinology. 2013;98(3):188–199. doi:10.1159/000355622.
4. Webb SM, Ware JE, Forsythe A, Yang M, Badia X, et al. Treatment effectiveness of pasireotide on health-related quality of life in patients with Cushing's disease. Eur J Endocrinol. 2014;171(1):89–98. doi:10.1530/EJE-13-1013.
5. Clayton RN, Jones PW, Reulen RC, Stewart PM, Hassan-Smith ZK, et al. Mortality in patients with Cushing's disease more than 10 years after remission: a multicentre, multinational, retrospective cohort study. Lancet Diabetes Endocrinol. 2016;4(7):569–576. doi:10.1016/S2213-8587(16)30005-5.
6. Graversen D, Vestergaard P, Stochholm K, Gravholt CH, Jørgensen JO. Mortality in Cushing's syndrome: a systematic review and meta-analysis. Eur J Intern Med. 2012;23(3):278–282. doi:10.1016/j.ejim.2011.10.013.
7. Chabre O, Goichot B, Zenaty D, Bertherat J. Epidemiology of primary and secondary adrenal insufficiency: prevalence and incidence, acute adrenal insufficiency, long-term morbidity and mortality. Ann Endocrinol (Paris). 2017;78(6):490–494. doi:10.1016/j.ando.2017.10.010.
8. Rushworth RL, Torpy DJ, Stratakis CA, Falhammar H. Adrenal crises in children: perspectives and research directions. Horm Res Paediatr. 2018;89(5):341–351. doi:10.1159/000481660.
9. Bornstein SR, Allolio B, Arlt W, Barthel A, Don-Wauchope A, et al. Diagnosis and treatment of primary adrenal insufficiency: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2016;101(2):364–389. doi:10.1210/jc.2015-1710.
10. Erichsen MM, Løvås K, Skinningsrud B, Wolff AB, Undlien DE, et al. Clinical, immunological, and genetic features of autoimmune primary adrenal insufficiency: observations from a Norwegian registry. J Clin Endocrinol Metab. 2009;94(12):4882–4890. doi:10.1210/jc.2009-1368.
11. Rogoff M. Addison's disease: further report on treatment with "interrenalin" (adrenal cortical extract). JAMA. 1932;99(16):1309–1315. doi:10.1001/jama.1932.02740680005002.
12. Al Nofal A, Bancos I, Benkhadra K, Ospina NM, Javed A, et al. Glucocorticoid replacement regimens in chronic adrenal insufficiency: a systematic review and meta-analysis. Endocr Pract. 2017;23(1):17–31. doi:10.4158/EP161428.OR.
13. Hahner S, Spinnler C, Fassnacht M, Burger-Stritt S, Lang K, et al. High incidence of adrenal crisis in educated patients with chronic adrenal insufficiency: a prospective study. J Clin Endocrinol Metab. 2015;100(2):407–416. doi:10.1210/jc.2014-3191.
14. White K, Arlt W. Adrenal crisis in treated Addison's disease: a predictable but under-managed event. Eur J Endocrinol. 2010;162(1):115–120. doi:10.1530/EJE-09-0559.
15. Dekkers OM, Horváth-Puhó E, Jørgensen JO, Cannegieter SC, Ehrenstein V, et al. Multisystem morbidity and mortality in Cushing's syndrome: a cohort study. J Clin Endocrinol Metab. 2013;98(6):2277–2284. doi:10.1210/jc.2012-3582.
16. Lindholm J, Juul S, Jørgensen JO, Astrup J, Bjerre P, et al. Incidence and late prognosis of Cushing's syndrome: a population-based study. J Clin Endocrinol Metab. 2001;86 (1):117–123. doi:10.1210/jcem.86.1.7093.

17. Nieman LK, Biller BMK, Findling JW, Hassan Murad M, Newell-Price J, et al. Treatment of Cushing's syndrome: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2015;100(8):2807–2831. doi:10.1210/jc.2015-1818.
18. Dinsen S, Baslund B, Klose M, Rasmussen AK, Friis-Hansen L, et al. Why glucocorticoid withdrawal may sometimes be as dangerous as the treatment itself. Eur J Intern Med. 2013;24(8):714–720. doi:10.1016/j.ejim.2013.05.014.
19. Broersen LH, Pereira AM, Jørgensen JO, Dekkers OM. Adrenal insufficiency in corticosteroids use: systematic review and meta-analysis. J Clin Endocrinol Metab. 2015;100(6):2171–2180. doi:10.1210/jc.2015-1218.
20. Hummel SR, Sadler S, Whitaker MJ, Ara RM, Dixon S, et al. A model for measuring the health burden of classic congenital adrenal hyperplasia in adults. Clin Endocrinol. 2016;85(3):361–398. doi:10.1111/cen.13060.
21. Wengander S, Trimpou P, Papakokkinou E, Ragnarsson O. The incidence of endogenous Cushing's syndrome in the modern era. Clin Endocrinol. 2019;91(2):263–270. doi:10.1111/cen.14014.
22. Broder MS, Neary MP, Chang E, Cherepanov D, Ludlam WH. Incidence of Cushing's syndrome and Cushing's disease in commercially-insured patients under 65 years old in the United States. Pituitary. 2015;18(3):283–289. doi:10.1007/s11102-014-0569-6.
23. Steffensen C, Bak AM, Rubeck KZ, Jørgensen JO. Epidemiology of Cushing's syndrome. Neuroendocrinology. 2010;92(1):1–5. doi:10.1159/000314297.
24. Govindan R, Page N, Morgensztern D, Read W, Tierney R, et al. Changing epidemiology of small-cell lung cancer in the United States over the last 30 years: analysis of the surveillance, epidemiologic, and end results database. J Clin Oncol. 2006;24(28):4539–4544. doi:10.1200/JCO.2005.04.4859.
25. Newell-Price J, Trainer P, Besser M, Grossman A. The diagnosis and differential diagnosis of Cushing's syndrome and pseudo-Cushing's states. Endocr Rev. 1998;19(5):647–672. doi:10.1210/edrv.19.5.0346.
26. Flack MR, Chrousos GP. Neoplasms of the adrenal cortex. In: Holland R, Frei E, editors. Cancer medicine, 4th edition. New York: Lea and Fibinger; 1996:1563–1570.
27. Arlt W, Society for Endocrinology Clinical Committee. Society for Endocrinology endocrine emergency guidance: emergency management of acute adrenal insufficiency (adrenal crisis) in adult patients. Endocr Connect. 2016;5(5):G1–G3. doi:10.1530/EC-16-0054.
28. Michels A, Michels N. Addison disease: early detection and treatment principles. Am Fam Physician. 2014;89(7):563–568.
29. Auron M, Raissouni N. Adrenal insufficiency. Pediatr Rev. 2015;36(3):92–103. doi:10.1542/pir.36-3-92.
30. Husebye ES, Allolio B, Arlt W, Badenhoop K, Bensing S, et al. Consensus statement on the diagnosis, treatment and follow-up of patients with primary adrenal insufficiency. J Intern Med. 2014;275(2):104–15. doi:10.1111/joim.12162.
31. Espinosa de los Monteros-Sánchez AL, Valdivia-López J, Mendoza-Zubieta V, Mercado-Atri M, Gómez-Pérez F, et al. Consenso en el diagnóstico y tratamiento del síndrome de Cushing. Rev Endocrinol Nutr. 2007;15(S20):3–12. [Mexican society consensus]
32. El-Farhan N, Rees DA, Evans C. Measuring cortisol in serum, urine and saliva – are our assays good enough? Ann Clin Biochem. 2017;54(3):308–322. doi:10.1177/0004563216687335.
33. Turpeinen U, Hämäläinen E. Determination of cortisol in serum, saliva and urine. Best Pract Res Clin Endocrinol Metab. 2013;27(6):795–801. doi:10.1016/j.beem.2013.10.008.

34. Dickstein G, Saiegh L. Low-dose and high-dose adrenocorticotropin testing: indications and shortcomings. Curr Opin Endocrinol Diabetes Obes. 2008;15(3):244–249. doi:10.1097/MED.0b013e3282fdf16d.
35. Burgos N, Ghayee HK, Singh-Ospina N. Pitfalls in the interpretation of the cosyntropin stimulation test for the diagnosis of adrenal insufficiency. Curr Opin Endocrinol Diabetes Obes. 2019;26(3):139–145. doi:10.1097/MED.0000000000000473.
36. Kazlauskaite R, Evans AT, Villabona CV, Abdu TA, Ambrosi B, et al. Corticotropin tests for hypothalamic-pituitary-adrenal insufficiency: a meta-analysis. J Clin Endocrinol Metab. 2008;93(11):4245–4253. doi:10.1210/jc.2008-0710.
37. Dorin RI, Qualls CR, Crapo LM. Diagnosis of adrenal insufficiency. Ann Intern Med. 2003;139:194–204. doi:10.7326/0003-4819-139-3-200308050-00009.
38. Chiodini I, Ramos-Rivera A, Marcus AO, Yau H. Adrenal hypercortisolism: a closer look at screening, diagnosis, and important considerations of different testing modalities. J Endocr Soc. 2019;3(5):1097–1109. doi:10.1210/js.2018-00382.
39. Elamin MB, Murad MH, Mullan R, Erickson D, Harris K, et al. Accuracy of diagnostic tests for Cushing's syndrome: a systematic review and metaanalyses. J Clin Endocrinol Metab. 2008;93(5):1553–1562. doi:10.1210/jc.2008-0139.
40. Hawley JM, Owen LJ, Lockhart SJ, Monaghan PJ, Armston A, et al. Serum cortisol: an up-to-date assessment of routine assay performance. Clin Chem. 2016;62(9):1220–1229. doi:10.1373/clinchem.2016.255034.
41. Ortiz-Flores AE, Santacruz E, Jiménez-Mendiguchia L, García-Cano A, Nattero-Chávez L, et al. Role of sampling times and serum cortisol cut-off concentrations on the routine assessment of adrenal function using the standard cosyntropin test in an academic hospital from Spain: a retrospective chart review. BMJ Open. 2018;8(5):e019273. doi:10.1136/bmjopen-2017-019273.
42. Kline GA, Buse, J, Krause RD. Clinical implications for biochemical diagnostic thresholds of adrenal sufficiency using a highly specific cortisol immunoassay. Clin Biochem. 2017;50(9):475–480. doi:10.1016/j.clinbiochem.2017.02.008.
43. Struja T, Briner L, Meier A, Kutz A, Mundwiler E, et al. Diagnostic accuracy of basal cortisol level to predict adrenal insufficiency in cosyntropin testing: results from an observational cohort study with 804 patients. Endocr Pract. 2017;23(8):949–961. doi:10.4158/EP171861.OR.
44. Lopez Schmidt I, Lahner H, Mann K, Petersenn S. Diagnosis of adrenal insufficiency: evaluation of the corticotropin-releasing hormone test and basal serum cortisol in comparison to the insulin tolerance test in patients with hypothalamic-pituitary-adrenal disease. J Clin Endocrinol Metab. 2003;88(9):4193–4198. doi:10.1210/jc.2002-021897.
45. Montes-Villarreal J, Perez-Arredondo LA, Rodriguez-Gutierrez R, Gonzalez-Colmenero AD, Solis-Pacheco RC, et al. Serum morning cortisol as a screening test for adrenal insufficiency. Endocr Pract. 2020;26:30–35. doi:10.4158/EP-2019-0327.
46. Peechakara S, Bena J, Clarke NJ, McPhaul MJ, Reitz RE, et al. Total and free cortisol levels during 1 µg, 25 µg, and 250 µg cosyntropin stimulation tests compared to insulin tolerance test: results of a randomized, prospective, pilot study. Endocrine. 2017;57(3):388–393. doi:10.1007/s12020-017-1371-9.
47. Barrou Z, Guiban D, Maroufi A, Fournier C, Dugué MA, et al. Overnight dexamethasone suppression test: comparison of plasma and salivary cortisol measurement for the screening of Cushing's syndrome. Eur J Endocrinol. 1996;134(1):93–96. doi:10.1530/eje.0.1340093.
48. Tang TJ, Liu YP, Yu YR. Comparing overnight dexamethasone suppression test, urine free cortisol, and midnight serum cortisol for the initial diagnosis of Cushing's syndrome. Sichuan Da Xue Xue Bao Yi Xue Ban. 2013;44(5):764–768.

49. Ho W, Druce M. Quality of life in patients with adrenal disease: a systematic review. Clin Endocrinol. 2018;89(2):119–128. doi:10.1111/cen.13719.

50. Crown A, Lightman S. Why is the management of glucocorticoid deficiency still controversial: a review of the literature. Clin Endocrinol. 2005;63(5):483–492. doi:10.1111/j.1365-2265.2005.02320.x.

51. Johannsson G, Falorni A, Skrtic S, Lennernäs H, Quinkler M, et al. Adrenal insufficiency: review of clinical outcomes with current glucocorticoid replacement therapy. Clin Endocrinol. 2015;82(1):2–11. doi:10.1111/cen.12603.

52. Miller BS, Spencer SP, Geffner ME, Gourgari E, Lahoti A, et al. Emergency management of adrenal insufficiency in children: advocating for treatment options in outpatient and field settings. J Investig Med. 2020;68(1):16–25. doi:10.1136/jim-2019-000999.

53. Khan U, Lakhani OJ. Management of primary adrenal insufficiency: review of current clinical practice in a developed and a developing country. Indian J Endocrinol Metab. 2017;21(5):781–783. doi:10.4103/ijem.IJEM_193_17.

54. Wichers M, Springer W, Bidlingmaier F, Klingmüller D. The influence of hydrocortisone substitution on the quality of life and parameters of bone metabolism in patients with secondary hypocortisolism. Clin Endocrinol. 1999;50(6):759–765. doi:10.1046/j.1365-2265.1999.00723.x.

55. Laureti S, Falorni A, Santeusanio F. Improvement of treatment of primary adrenal insufficiency by administration of cortisone acetate in three daily doses. J Endocrinol Invest. 2003 Nov;26(11):1071–1075. doi:10.1007/BF03345252.

56. Mah PM, Jenkins RC, Rostami-Hodjegan A, Newell-Price J, Doane A, et al. Weight-related dosing, timing and monitoring hydrocortisone replacement therapy in patients with adrenal insufficiency. Clin Endocrinol. 2004;61(3):367–375. doi:10.1111/j.1365-2265.2004.02106.x.

57. Johannsson G, Nilsson AG, Bergthorsdottir R, Burman P, Dahlqvist P, et al. Improved cortisol exposure-time profile and outcome in patients with adrenal insufficiency: a prospective randomized trial of a novel hydrocortisone dual-release formulation. J Clin Endocrinol Metab. 2012;97(2):473–481. doi:10.1210/jc.2011-1926.

58. Broersen LHA, Biermasz NR, van Furth WR, de Vries F, Verstegen MJT, et al. Endoscopic vs. microscopic transsphenoidal surgery for Cushing's disease: a systematic review and meta-analysis. Pituitary. 2018;21(5):524–534. doi:10.1007/s11102-018-0893-3.

59. Petersen S, Beckers A, Ferone D, van der Lely A, Bollerslev J, et al. Therapy of endocrine disease: outcomes in patients with Cushing's disease undergoing transsphenoidal surgery: systematic review assessing criteria used to define remission and recurrence. Eur J Endocrinol. 2015;172(6):R227–R339. doi:10.1530/EJE-14-0883.

60. Ritzel K, Beuschlein F, Mickisch A, Osswald A, Schneider HJ, et al. Clinical review: outcome of bilateral adrenalectomy in Cushing's syndrome: a systematic review. J Clin Endocrinol Metab. 2013;98(10):3939–3948. doi:10.1210/jc.2013-1470.

61. Broersen LHA, Jha M, Biermasz NR, Pereira AM, Dekkers OM. Effectiveness of medical treatment for Cushing's syndrome: a systematic review and meta-analysis. Pituitary. 2018;21(6):631–641. doi:10.1007/s11102-018-0897-z.

62. Fleseriu M, Castinetti F. Updates on the role of adrenal steroidogenesis inhibitors in Cushing's syndrome: a focus on novel therapies. Pituitary. 2016;19(6):643–653. doi:10.1007/s11102-016-0742-1.

63. van Haalen FM, Broersen LH, Jorgensen JO, Pereira AM, Dekkers OM. Management of endocrine disease: mortality remains increased in Cushing's disease despite biochemical remission: a systematic review and meta-analysis. Eur J Endocrinol. 2015;172(4):R143–R149. doi:10.1530/EJE-14-0556.

64. Clayton RN, Raskauskiene D, Reulen RC, Jones PW. Mortality and morbidity in Cushing's disease over 50 years in Stoke-on-Trent, UK: audit and meta-analysis of literature. J Clin Endocrinol Metab. 2011;96(3):632–642. doi:10.1210/jc.2010-1942.

65. Hammer GD, Tyrrell JB, Lamborn KR, Applebury CB, Hannegan ET, et al. Transsphenoidal microsurgery for Cushing's disease: initial outcome and long-term results. J Clin Endocrinol Metab. 2004;89(12):6348–6357. doi:10.1210/jc.2003-032180.

66. Faggiano A, Pivonello R, Spiezia S, De Martino MC, Filippella M, et al. Cardiovascular risk factors and common carotid artery caliber and stiffness in patients with Cushing's disease during active disease and 1 year after disease remission. J Clin Endocrinol Metab. 2003;88(6):2527–2533. doi:10.1210/jc.2002-021558.

67. Davies JH, Storr HL, Davies K, Monson JP, Besser GM, et al. Final adult height and body mass index after cure of paediatric Cushing's disease. Clin Endocrinol. 2005;62(4):466–472. doi:10.1111/j.1365-2265.2005.02244.x.

3.12. Estradiol

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding serum estrogen level to the EDL as an IVD to assess reproductive function in women, to include menstrual irregularities and menopausal status, and to evaluate infertility and guide its treatment.

B. Applicant

World Health Organization

C. WHO technical department

Sexual and Reproductive Health and Research
Contraception and Fertility Care

D. Background (from application)

Disease condition and impact on patients

Estrogens are steroid hormones that are responsible for the development and regulation of the female reproductive system and secondary sex characteristics. In women, they are synthesized primarily in the ovaries (in response to follicle stimulating hormone (FSH) stimulation) and in the placenta during pregnancy. Lesser amounts of estrogens are produced in other tissues, such as liver, adipose, adrenal, skin and brain. In men, estrogens are synthesized in response to FSH in the testes but are produced in much lower amounts than in women.

There are three main naturally occurring estrogens: estrone (E1), estradiol (E2) and estriol (E3) (*1*). In terms of estrogenic activity, E2 is the predominant estrogen during reproductive years and thus is the most clinically relevant. After menopause when E2 levels fall, E1 becomes the predominant circulating estrogen. E3 is the main estrogen during pregnancy; it does not have a significant role in non-pregnant women or men (*1*). Given its predominant role and importance in clinical utility, this application focuses on the measurement of E2.

Measuring serum E2 levels has an integral role in the assessment of reproductive function in women, specifically infertility, menstrual irregularities and menopausal status. E2 levels are also commonly measured to monitor ovulation induction, and during preparation for in vitro fertilization (IVF).

Infertility. Infertility is recognized as an essential component of reproductive health by the UN Programme of Action of the International Conference on Population and Development (*2*). Paradoxically, nations with the highest overall fertility are also the ones with the greatest prevalence of infertility and these often include LMICs. Given the economic, resource, cultural and religious constraints in these countries, infertility services among them will vary significantly. Assays that measure E2 levels are, however, relatively non-invasive (requiring only a blood draw or finger stick), inexpensive and accessible from laboratories throughout the world. Measuring E2 levels can be used in infertility management to assess ovarian reserve (*3*) and predict IVF success (*4*) in women, and as an aid in determining the etiology of azoospermia in men (*5*).

Menstrual irregularities. E2 levels can also be useful in evaluating irregular menstrual cycles, particularly suspected anovulation and amenorrhoea, which generally indicate a defect at some point in the hypothalamic-pituitary-ovarian-uterine axis.

Anovulation. In addition to impacting fertility, ovulation disorders can also cause menstrual irregularities ranging from amenorrhoea to dysfunctional uterine bleeding (DUB).

Primary amenorrhoea is defined as no menses by age 14 in the absence of development of secondary sexual characteristics, or no menses by age 16 regardless of sexual development (*6*). Most cases are caused by gonadal dysgenesis, most commonly Turner syndrome or (HH) as well as Kallmann syndrome. Secondary amenorrhoea is defined in women who have menstruated before as no menses for 6 months or the equivalent of at least three cycles (*6*). It is most commonly caused by primary ovarian insufficiency and menopause. Hypothalamic amenorrhoea, another category of secondary amenorrhoea, is essentially a diagnosis of exclusion in the face of low estrogen levels and low or normal FSH. It is often associated with the excessive exercise and/or weight loss

that can be seen in elite athletes (*7*). But it can also be precipitated by weight loss due to malnutrition or extreme psychological stress. Although there are no known studies that have documented this, hypothalamic amenorrhoea is likely underdiagnosed in LMICs. Because these women are estrogen deficient, they are at risk for its natural clinical outcomes, and failure to identify their deficiency may also lead to conditions such as osteoporosis and genital atrophy.

Anovulation can also lead to DUB (*8*), which causes anaemia and can negatively impact psychological well-being, personal economics and social interactions, particularly in LMICs. Studies report that women in LMICs face increased physical and psychological obstacles coping with menstrual flow compared to women in HICs. Menstruating women in LMICs often face challenges in accessing menstrual materials, as well as religious or cultural shunning, lack of privacy, absences from school or work due to stained clothing and overall poor self-esteem (*9*).

General reproductive function. Primary ovarian insufficiency (POI), previously called "premature menopause" or "premature ovarian failure", is a heterogeneous disorder resulting in the depletion or dysfunction of ovarian follicles with menses stopping before 40 years of age. Follicle depletion can be caused by chromosomal abnormalities (e.g. fragile X syndrome), chemotherapy or radiation therapy, autoimmune disease or infiltrative or infective processes. POI affects approximately 1% of women.

Does the test meet a medical need?

See "How the test is used".

How the test is used

Infertility. Since E2 is secreted by developing ovarian follicles, it can be used to evaluate ovarian reserve. The number of ovarian follicles present in the ovaries (oocytes) is established at birth and declines with time throughout a woman's lifetime without regeneration. It can therefore become necessary to confirm the presence of viable oocytes and the potential to ovulate in infertile women by measuring E2 and FSH levels (*10*). Although testing for anti-Müllerian hormone (AMH) has gained favour in predicting IVF success (*10*), measurement of E2 (with FSH) levels has also been shown to be a significant prognostic marker and may be more accessible in low-resource settings (*11*).

Menstrual irregularities. An E2 measurement provides an accessible and relatively inexpensive objective assessment of ovarian function. Once a diagnosis of anovulation is confirmed, measuring E2 levels can also help clarify the etiology, according to the classification adopted by WHO (*12*):

- Group I (hypogonadotropic hypogonadal anovulation) accounts for around 5–10% of anovulatory women and is characterized by low E2 in the presence of low FSH.
- Group II (normogonadotropic normestrogenic anovulation) accounts for around 75–85% of anovulatory women and is characterized by normal E2 with normal FSH.
- Group III (hypergonadotropic anovulation) accounts for 10–20% of anovulatory women and is characterized by suppressed E2 in the presence of elevated FSH.
- Group IV (hyperprolactinaemic anovulation) accounts for 5–10% of anovulatory women and is characterized by low E2 levels with generally low FSH.

Amenorrhoea and hypogonadism. There are many causes of primary and secondary amenorrhoea, and multiple laboratory measurements are used in their evaluation and diagnosis. E2 levels are primarily used in conjunction with FSH and luteinizing hormone (LH). Decreased E2 levels indicate hypogonadism, and subsequent measurement of gonadotropins is essential to distinguish whether the abnormality is gonadal or of pituitary/hypothalamic origin (*13*). Physical manifestations of gonadal dysgenesis may be subtle, so measuring reproductive hormone levels can facilitate diagnosis, particularly in LMICs, where laboratory menus and diagnostic resources may be limited. Early diagnosis of gonadal dysgenesis is essential, as these syndromes often involve congenital heart defects and other organ anomalies. Timely diagnosis also allows for initiation of growth hormone treatment in early childhood to address short stature and estrogen in adolescence to promote pubertal development and prevent osteoporosis (*14*). E2 levels can be used to monitor this treatment.

In males, increased E2 levels may help confirm a diagnosis of primary hypogonadism.

General reproductive function. When POI is suspected on the basis of history and physical exam, measuring E2 and FSH can confirm the diagnosis. If E2 levels indicate hypogonadism and basal FSH levels are elevated into the menopausal range, a repeat measurement 1 month later is indicated for final confirmation (*15*).

In precocious puberty (PP), E2, LH and FSH tend to be above the prepubertal range (*16*). E2 measurement in children suspected of having PP is used to support the diagnosis and to determine the etiology. Increased E2 levels can be caused by exogenous estrogens or an ovarian cyst that has produced transient estrogens. Early diagnosis of central PP allows for treatment with gonadotropin-releasing hormone (GnRH) agonists that can delay the onset of puberty until a more appropriate age. The psychosocial and physical burdens of

undiagnosed and untreated delayed or precocious puberty are significant, and it is important that all tools available to aid in a timely diagnosis be available.

The combination of low E2 and high FSH levels may be used to confirm menopausal status, particularly in women who have had hysterectomies (with ovaries intact) and in whom the absence of menses cannot be used as a reliable indicator of ovarian function (*17*).

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Infertility. Estimating the prevalence of global infertility is challenging due to differences in definitions and methodology and the lack of population-based studies (*18*). One analysis based on health surveys estimates that more than 40 million couples are affected by infertility (*18*). And data on global prevalence of childlessness suggest that as many as 70 million couples would benefit from medical intervention to achieve pregnancy (*19*). The socioeconomic impact of infertility is significant, particularly for women, who often suffer from social isolation, discrimination, disinheritance, depression, abuse, divorce and possible abandonment in old age. Infertility can also have broader negative impacts on families, particularly in LMICs, where children contribute to family incomes and older parents depend on their children for support.

Menstrual irregularities. The prevalence of menstrual disorders is estimated to range from 5% to ~36% (*20*); occurrence depends on age, nutritional status and country of residence.

Anovulation. Estimates of the incidence of anovulation are imprecise due to variables such as age, general health, nutritional status and country of residence. Up to 25% of infertile women are estimated to be anovulatory (*12*). Turner syndrome has a reported incidence of 1 per 2500 live births (*21*); Kallmann syndrome has a reported incidence of 1 in 120 000 women (*22*).

F. WHO or other clinical guidelines relevant to the test

A 2016 Endocrine Society clinical practice guideline on hormonal replacement in hypopituitarism in adults (*13*) recommends measuring E2, FSH and LH in females with oligomenorrhoea or amenorrhoea.

In 2014, the American College of Obstetricians and Gynecologists (ACOG) (*15*) listed measuring FSH and E2 (two random tests at least 1 month apart) when there has been menstrual irregularity for at least three consecutive months as a way of diagnosing and initially evaluating POI.

The American Society for Reproductive Medicine (ASRM) Practice Committee (*11*) states that ovarian reserve tests should include both biochemical tests and ultrasound imaging of the ovaries, and that biochemical tests should

include both basal measurements (of FSH, E2, inhibin B and AMH) and provocative tests such as the clomiphene citrate challenge test. The committee go on to say that basal E2 alone should not be used to screen for DOR, explaining that the test is only of value as an aid to correctly interpreting a "normal" basal FSH value. An early rise in serum E2 is a classic characteristic of reproductive ageing and can lower an otherwise elevated basal FSH level into the normal range, thereby causing a misinterpretation of the test. When the basal FSH concentration is "normal" but the E2 level is elevated (> 60–80 pg/mL) in the early follicular phase, there is limited evidence for an association with poor response, increased cancellation rates or lower pregnancy rates.

G. Basic test characteristics (from application)

Test formats available	Immunoassays of varying formats, for example electrochemiluminescence and chemiluminescence.
Specimen types	Serum, plasma
Equipment required	Automated analyser
Regulatory status	CE-marked and FDA-approved; also approved by other regulatory agencies, including Health Canada
Availability	Global
Price per test range	US$ 4.50–43 depending on the country
Instrument price range	US$ 35 000–85 000 depending on the instrument size

H. Evidence for diagnostic accuracy (from application)

Because E2 testing is a quantitative measurement that is not used in screening for diseases, and it is generally done in conjunction with other testing to support a diagnosis of disease or to monitor a treatment, sensitivity and specificity have not been determined for many indications.

I. Evidence for clinical usefulness and impact (from application)

A 2019 review by Mikhael et al. (*23*) suggests that, in general, basic evaluation of hypothalamic-pituitary failure includes measurement of E2 (with FSH and LH). Overall management of hypothalamic-pituitary failure will depend on patient needs. For example, in young women of reproductive age where E2 levels (with FSH) are used to diagnose ovarian insufficiency, management can include reassurance and education about alternate reproductive options. This may also apply to patients with mosaic Turner syndrome who become aware of impending follicular atresia and who wish to preserve fertility through oocyte or embryo cryopreservation.

In cases where E2 is used to confirm ovarian failure, either premature or due to natural menopause, the diagnosis allows for monitoring the effects of estrogen deficiency and consideration of hormonal replacement to prevent or delay the associated morbidity.

Infertility. A 2003 WHO study (*24*) looked at high singleton live birth rate following classical ovulation induction in normogonadotrophic anovulatory infertility. All patients started with clomiphene for ovulation induction: those who did not ovulate within three treatment cycles of incremental daily doses up to 150 mg for five consecutive days or ovulatory clomiphene citrate (CC) patients who did not conceive within six cycles underwent gonadotrophin induction of ovulation applying a step-down dose regimen. Of 240 consecutive women, there were 134 pregnancies ending in a singleton live birth (56% of women). The cumulative pregnancy rate after 12 and 24 months of follow-up was 50% and 71%, respectively. The authors concluded that classical ovulation induction produces very good results in normogonadotrophic anovulatory infertility.

A 2014 study in India by Prasad et al. (*25*) looked at the E2 level on downregulated day 2 and on the day of human chorionic gonadotropin (hCG) trigger as a predictor of IVF success. The authors found that for women with E2 levels > 31.2 pg/mL on downregulated day 2, 61.8% will achieve pregnancy; 59% of women with E2 levels greater than 1400 pg/mL on the day of hCG trigger will achieve pregnancy.

Menstrual irregularities. The clinical community generally accepts that hormonal therapy is highly successful in treating DUB. A 2004 study by Hurskainen et al. (*26*) followed more than 200 women with DUB who were treated with a triphasic preparation of norgestimate and ethinyl E2. More than 80% were documented to have improvement in bleeding, and this was significantly increased over the placebo group. The authors concluded that the triphasic combination of norgestimate and ethinyl E2 is an effective treatment for metrorrhagic, menometrorrhagic, oligomenorrhoeic and polymenorrheic dysfunctional uterine bleeding.

General reproductive function. In 2012, Lee et al. (27) studied 76 girls with PP who were treated with leuprolide acetate, a GnRH agonist, every 3 months. At 6 months, 98% exhibited LH suppression and 100% E2 suppression. The authors concluded that treatment with leuprolide acetate 3-month depot formulations (11.25 and 30 mg) effectively suppressed the GnRH axis.

J. Evidence for economic impact and/or cost–effectiveness (from application)

The cost (in US dollars) of an E2 test in the USA is around US$ 43, compared with US$ 4.50–7 in India and US$ 34 in Australia. By contrast, an AMH test is around US$ 100 in the USA, US$ 17–28 in India and US$ 64 in Australia.

It is several times more expensive to predict IVF success through ultrasound determination of follicle number and size. An E2 measurement that prompts patients to abandon a cycle of assisted reproductive technology could also result in significant savings (thousands of dollars in the USA).

It is also relatively inexpensive to determine whether anovulation has occurred using an E2 measurement, particularly in cases of amenorrhoea. Given the significant health repercussions of anovulation in terms of disorders of sexual development (DSDs) and also the sequelae of estrogen deficiency, there should be no question that identifying these risks and reducing morbidity would be cost–effective.

K. Ethics, equity and human rights issues (from application)

No information provided.

L. Summary of evidence evaluation

The estradiol test is aimed at aiding the diagnosis of disorders related to reproductive dysfunction in women, more specifically to screen for decreased ovarian reserve. No evidence is presented on the diagnostic accuracy of the test for these conditions. Practice guidelines and committee opinions recommend using estradiol for a range of conditions including WHO I–III anovulation, amenorrhoea, menstrual irregularities and hypogonadism, although they warn against using basal estradiol alone to screen for decreased ovarian reserve.

The estradiol test has value as an aid to correct interpretation of a "normal" basal serum FSH value.

M. Summary of SAGE IVD deliberations

The consequences of infertility from a personal and social point of view can be dramatic. Fertility is considered a major public health issue by WHO. Estradiol is used in clinical practice as part of a battery of tests for diagnosing and managing a range of fertility issues. It forms part of several international guidelines for diagnosing WHO I–III anovulation. It is also acknowledged in guidelines to be of some (more limited) use in diagnosing and monitoring ovarian reserve, amenorrhoea, PP and hypogonadism. One of the medicines that is routinely used in fertility – clomiphene – is also included in the complementary list of the EML.

There is, however, potential for overuse of the test. SAGE IVD emphasized that LH testing goes hand in hand with FSH testing and that result interpretation requires appropriate laboratory infrastructure and the availability of fertility or endocrinology specialists. The group also emphasized that estradiol test results can only be interpreted correctly in combination with other tests, for example FSH.

SAGE IVD noted that the specimen types listed in the estradiol submission to the EDL did not include urine, even though urine testing is often used in practice.

N. SAGE IVD recommendations

SAGE IVD recommended including the estradiol test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, within a new subsection for endocrine disorders);
- using an immunoassay format;
- to aid in the diagnosis of primary and secondary amenorrhoea, anovulation, gonadal dysfunction and precocious puberty; and
- to aid in the evaluation and management of infertility.

The group requested the addition of a note to the test category entry in the EDL stating that it is only recommended for use in specialized health care settings.

The group further requested that the original submitter of the estradiol test category be asked to provide information on the applications of estradiol in urine to support an edit of the EDL entry next year.

O. References

1. Bulun S, Adashi Y. The physiology and pathology of the female reproductive axis. In: Melmed S, Polonsky K, Reed Larsen PR, editors. Williams textbook of endocrinology, 11th edition. Philadelphia: Saunders/Elsevier; 2008:541–614.
2. Nachtigall RD. International disparities in access to infertility services. Fertil Steril. 2006;85(4):871–875. doi:10.1016/j.fertnstert.2005.08.066.
3. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:1110, 982, 1045.
4. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:982.
5. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:1045.
6. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:343.
7. Hamilton-Fairley D, Taylor A. Anovulation. BMJ. 2003;327(7414):546–549. doi:10.1136/bmj.327.7414.546.
8. Committee on Practice Bulletins – Gynecology. Practice bulletin no. 136: management of abnormal uterine bleeding associated with ovulatory dysfunction. Obstet Gynecol. 2013;122(1):176–185. doi:10.1097/01.AOG.0000431815.52679.bb.
9. Hennegan J, Shannon AK, Rubil J, Schwab KJ, Melendez-Torres GJ. Women's and girls' experiences of menstruation in low- and middle-income countries: a systematic review and qualitative metasynthesis. PLoS Med. 2019;16(5):e1002803. doi:10.1371/journal.pmed.1002803.

10. Infertility workup for women's health specialist: ACOG Committee opinion, number 781. Obstet Gynecol. 2019;133(6):e377–e384. doi:10.1097/AOG.0000000000000327.
11. Practice Committee of the American Society for Reproductive Medicine. Testing and interpreting measures of ovarian reserve: a committee opinion. Fertil Steril. 2015;103:e9–e17. doi:10.1016/j.fertnstert.2014.12.093.
12. National Collaborating Centre for Women's and Children's Health. Fertility: assessment and treatment for people with fertility problems. NICE clinical guidelines, no. 156. London: Royal College of Obstetricians and Gynaecologists; 2013.
13. Fleseriu M, Hashim IA, Karavitaki N, Melmed S, Hassan Murad M, et al. Hormonal replacement in hypopituitarism in adults: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2016;101(11):3888–3921. doi:10.1210/jc.2016-2118.
14. Morgan T. Turner syndrome: diagnosis and management. Am Fam Physician. 2007;76(3):405–410.
15. Committee opinion number 605: primary ovarian insufficiency in adolescents and young women. Obstet Gynecol. 2014;124(1):193–197. doi:10.1097/01.AOG.0000451757.51964.98.
16. Ankarber-Lindgren C, Kristrom B, Norjavaara E. Physiological estrogen replacement therapy for puberty induction in girls: a clinical observational study. Horm Res Paediatr. 2014;81(4):239–244. doi:10.1159/000356922.
17. Polesel DN, Nozoe KT, Sanchez MZ, Pardo MCO, Bittencourt LR, et al. The follicle-stimulating hormone as best classifier for diagnosis of natural menopause. J Gynecol Surg. 2017;33(6):236–242. doi:10.1089/gyn.2017.0040.
18. Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. PLoS Med. 2012;9(12):e1001356. doi:10.1371/journal.pmed.1001356.
19. Ombelet W. Is global access to infertility care realistic? The Walking Egg Project. Reprod Biomed Online. 2014;28(3):267–272. doi:10.1016/j.rbmo.2013.11.013.
20. Kwak Y, Kim Y, Baek KA. Prevalence of irregular menstruation according to socioeconomic status: a population-based nationwide cross-sectional study. PLoS One. 2019;14(3):e0214071. doi:10.1371/journal.pone.0214071.
21. Viswanathan V, Eugaster EA. Etiology and treatment of hypogonadism in adolescents. Pediatr Clin North Am. 2011;58(5):1181. doi:10.1016/j.pcl.2011.07.009.
22. Kallmann syndrome. In: NIH/Genetics home reference [website]. Bethesda: US National Library of Medicine; 2020 (https://ghr.nlm.nih.gov/condition/kallmann-syndrome#genes, accessed 28 March 2020).
23. Mikhael S, Punjala-Patel A, Gavrilova-Jordan L. Hypothalamic-pituitary-ovarian axis disorders impacting female fertility. Biomedicines. 2019;7(1):5. doi:10.3390/biomedicines7010005.
24. Eijkimans MJ, Imani B, Mulders AG, Habbema JD, Fauser BC. High singleton live birth rate following classical ovulation induction in normogonadotrophic anovulatory infertility (WHO 2). Hum Reprod. 2003;18(11):2357–2362. doi:10.1093/humrep/deg459.
25. Prasad P, Kumar Y, Singhal M, Sharma S. Estradiol level on day 2 and day of trigger: a potential predictor of the IVF-ET success. J Obstet Gynaecol India. 2014;64(3):202–207. doi:10.1007/s13224-014-0515-6.
26. Hurskainen R, Teperi J, Rissanen P, Aalto AM, Grenman S, et al. Clinical outcomes and costs with the levonorgestrel-releasing intrauterine system or hysterectomy for treatment of menorrhagia: randomized trial 5-year follow-up. JAMA. 2004;291:1456–1463. doi:10.1001/jama.291.12.1456.

27. Lee PA, Klein K, Mauras N, Neely EK, Bloch CA, et al. Efficacy and safety of leuprolide acetate 3-month depot 11.25 milligrams or 30 milligrams for the treatment of central precocious puberty. J Clin Endocrinol Metab. 2012;97(5):1572–1580. doi:10.1210/jc.2011-2704.

3.13. Follicle-stimulating hormone

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding blood follicle-stimulating hormone (FSH) levels to the EDL as an IVD to aid in the diagnosis of diseases caused by malfunction of the hypothalamic-pituitary-gonadal (HPG) axis, or to aid in the clinical evaluation of female and male infertility, menstrual irregularities, pituitary disorders, precocious or delayed puberty, and ovarian or testicular dysfunction.

B. Applicant

World Health Organization

C. WHO technical department

Sexual and Reproductive Health and Research
Contraception and Fertility Care

D. Background (from application)

Disease condition and impact on patients

FSH is a polypeptide hormone produced by the anterior pituitary gland and is responsible for stimulating the development of ovarian follicles in women and for supporting testicular Sertoli cells critical for spermatogenesis in males. In both males and females, FSH secretion is regulated by a balance of positive and negative feedback mechanisms involving the hypothalamic-pituitary axis, the reproductive organs, and the pituitary and sex steroid hormones. Measuring FSH levels provides insight into the functioning of the HPG axis and is used in the clinical evaluation of infertility (both female and male), menstrual irregularities, pituitary disorders, precocious or delayed puberty, and ovarian/testicular dysfunction. Diseases caused by malfunction of the HPG axis are varied with wide-ranging impacts; they affect both HICs and LMICs.

Infertility. See "Socioeconomic impact".

Ovarian reserve. The number of ovarian follicles present in the ovaries (oocytes) is established at birth and declines with time throughout a woman's lifetime without regeneration.

Menstrual disorders. Irregular menses, particularly amenorrhoea and oligomenorrhoea, generally indicate a defect at some point in the hypothalamic-pituitary-ovarian-uterine axis.

Polycystic ovary syndrome (PCOS) causes irregular menstrual cycles, polycystic ovaries and hirsutism (*1*); it may also include infertility, insulin-resistance, impaired glucose tolerance (type 2 diabetes) and dyslipidaemia (*1*). It is associated with negative psychosocial impacts, and the premature pubertal growth spurt and accelerated bone maturation can also result in reduced adult height (*2*).

Gonadal dysfunction. Hypergonadotropic hypogonadism is most commonly caused by Turner syndrome in females and Klinefelter syndrome in males. Exogenous administration of GnRH is often used to differentiate primary hypergonadotropic and HH (*3*). Following administration of exogenous GnRH, FSH levels will be totally or partially absent in HH but will gradually appear in response to the exogenous GnRH in HH. Children with hypogonadism experience delayed puberty if left untreated. Timely diagnosis of hypogonadism allows treatment with prepubertal hormone replacement (*4*). For girls, administering estrogen allows secondary sex characteristics to emerge and facilitates adequate breast and uterine development. Early treatment with human growth hormone may also significantly increase height. For boys, testosterone administration is used to induce puberty (*4*).

Does the test meet a medical need?

Infertility. FSH levels can help diagnose potential causes of infertility. For example, they may help distinguish a gonadal from a pituitary cause of anovulation or azoospermia; or they may be used to give an indication of ovarian reserve and predict IVF success (*5*). Although AMH has gained favour for these purposes (*6*), FSH is a less expensive and more accessible test and so offers a more reasonable choice, especially for LMICs introducing infertility services (including IVF).

Ovarian reserve. Ovarian reserve testing can inform women of their reproductive lifespan. Confirming the presence of viable oocytes and the potential to ovulate is also useful for female cancer patients of reproductive age who are being treated with gonadotoxic therapy.

Menstrual disorders. FSH levels can be used in the evaluation of irregular menstrual bleeding, particularly oligomenorrhoea and amenorrhoea, by helping to distinguish ovulatory dysfunction from uterine abnormalities as the cause.

Confirmation that irregular menstrual bleeding is caused by hypothalamic-pituitary-ovarian (HPO) dysfunction allows corrective treatment with hormonal intervention and carries a high degree of success in regulating cycles. In addition, appropriate hormonal replacement will aid in preventing the sequelae of estrogen deficiency.

PCOS. An elevated ratio of LH to FSH (greater than 3) is often used with history and physical exam to support a diagnosis of PCOS (*8*). Appropriate diagnosis of PCOS allows hormonal treatments that can regulate menstrual cycles, induce ovulation when indicated, decrease androgenic effects and also reduce the risks of metabolic abnormalities such as diabetes and metabolic syndrome, which frequently occur with PCOS.

Gonadal dysfunction. Gonadal dysfunction can be evaluated by measuring FSH levels to distinguish between primary gonadal failure (hypergonadotropic) and deficient gonadal stimulation (hypogonadotropic hypogonadism) (*3*). High FSH levels also provide confirmation of hypogonadism in the evaluation of primary amenorrhoea. Early diagnosis of hypogonadal syndromes allows more complete evaluation of associated anomalies. Furthermore, administration of growth hormone to young girls in childhood and hormone replacement at puberty for both girls and boys can reduce morbidity and mortality, as well as improve quality of life (*4*).

FSH levels can also aid in the diagnosis of suspected PP (*2*). In gonadotropin-dependent PP, elevated FSH levels rule out adrenal hyperplasia or CNS lesions (*9*). In gonadotropin-independent PP, investigating the source of exogenous estrogens can lead to a more timely diagnosis of gonadal hormone-producing cysts or tumours. The psychosocial and physical burdens of undiagnosed and untreated delayed or precocious puberty are significant, and it is important that all tools to aid in a timely diagnosis be available.

How the test is used

FSH levels are generally used alongside physical exam and other laboratory tests or procedures to confirm a diagnosis. For example, in cases where the physical exam of a child suggests delayed puberty, FSH can help distinguish between a pituitary-hypothalamic disorder and a primary gonadal dysfunction. FSH can be done with other laboratory tests, for example with LH levels to determine the etiology of hypogonadism. The ratio of LH/FSH may also be used to support a diagnosis of PCOS.

Infertility. FSH levels are usually included in a basic infertility workup, generally as part of a panel of reproductive markers. Elevated FSH levels can indicate decreased ovarian function, and consequent anovulation and insufficient ovarian

reserve. FSH levels measured on menstrual cycle days 2, 3 and 4 can also be used in evaluating whether an IVF cycle will be successful, although the sensitivity for predicting poor response is highly variable and dependent on the cut-off point. The sensitivity for predicting failure to achieve pregnancy is poorer (*10*).

FSH levels may also be useful in predicting diminished ovarian reserve in response to a clomiphene challenge. Studies show that measuring FSH levels after a clomiphene challenge is less specific than basal FSH but results in increasing specificity (*11*). The knowledge that an IVF cycle is likely to be unsuccessful can save costs and may also help alleviate the disappointment and stress that couples experience when assisted reproduction fails.

Ovarian reserve. Other methods, including AMH and antral follicular count, already exist to evaluate ovarian reserve in subfertile women (*12*), but measuring FSH levels at predetermined times of a menstrual cycle (typically day 3) remains an important test (*10*). Single FSH measures are not predictive of the perimenopausal state or the timing of menopause, but elevated FSH levels can confirm menopausal status (*13*), particularly in women who have had hysterectomies (with ovaries intact) and in whom the absence of menses cannot be used as a reliable indicator of ovarian function.

Menstrual disorders. FSH levels are used in evaluating the etiology of irregular menstrual bleeding, particularly amenorrhoea and oligomenorrhoea. Elevated FSH levels indicate ovarian dysfunction or failure, while normal or low levels suggest pituitary or hypothalamic failure (*14*). Multiple diagnostic criteria have been adopted for PCOS, but the common denominator appears to be oligoovulation and androgen excess.

Gonadal dysfunction. Measuring FSH after GnRH stimulation can help diagnose suspected PP (*15*). An increase in FSH levels in response to a GnRH challenge indicates central PP (gonadotropic-dependent), while no increase supports a diagnosis of peripheral (gonadotropic-independent) PP. Early diagnosis of central PP allows for treatment with GnRH agonists that can delay the onset of puberty until a more appropriate age.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Infertility. Estimates suggest that up to 186 million women globally are infertile (*16*, *17*). Although differing methods and definitions have been used to derive infertility burden, a recent review estimated the global prevalence of infertility to be 9% (*18*). Data on worldwide prevalence of childlessness further estimate that as many as 70 million couples would benefit from medical intervention to achieve pregnancy (*19*). There are not enough data to assess global infertility prevalence

trends over the past 20 years. But these trends can be impacted by a rise in sexually transmitted infections (STIs) with subsequent impaired reproductive organ function, lifestyle changes and delayed childbearing (*16–18*).

The economic and social impact of infertility is significant, particularly for women, who will often suffer from social isolation, discrimination, disinheritance, depression, abuse, divorce and possible abandonment in old age. Infertility as a common cause of childlessness can also have a broader negative economic impact on families, particularly in LMICs, where children contribute to family incomes and older parents depend on their children for support.

Infertility is recognized as an essential component of reproductive health by The UN Programme of Action of the International Conference on Population and Development (*20*). Paradoxically, nations with the highest overall fertility are also the ones with the greatest prevalence of infertility and these often include LMICs. Given the economic, resource, cultural and religious constraints in these countries, infertility services among them vary significantly. Assays that measure serum FSH levels are, however, relatively non-invasive (requiring only a blood draw or finger stick), inexpensive and accessible from laboratories throughout the world.

Menstrual disorders. Estimates of the prevalence of menstrual disorders range from 5% to ~36% (*21*); occurrence depends on age, nutritional status and country of residence. PCOS, which is a subset of menstrual irregularity, is thought to be the most common endocrine disorder found in women of reproductive age (*22*) and impacts all races and ethnicities. In unspecified populations, PCOS has a reported incidence rate of 3–10%, although more precise incidence is unknown due to underdiagnosis.

Gonadal dysfunction. Turner syndrome has an incidence of 1 per 2500 live births (*4*). Klinefelter syndrome has an incidence of 1 per 1000 live births (*1*). HH is rarer. The incidence of PP is estimated to be 1 per 5000–10 000 children, and occurs in females at ten times the frequency in males (*2*). Gonadotropin-independent PP is about five times less common than gonadotropin-dependent PP (*2*).

Early diagnosis of DSDs can reduce deaths and illness and improve physical and psychological well-being. But there is considerable variation in how these disorders are managed in resource-rich and resource-poor countries. Closing the gap requires establishing specialized treatment centres, appropriate diagnostic algorithms, strong laboratory quality assurance and accessible treatments (*23*).

F. WHO or other clinical guidelines relevant to the test

A 2016 Endocrine Society clinical practice guideline on hormonal replacement in hypopituitarism in adults (*3*) recommends measuring serum estradiol (E2),

FSH and LH in females with oligomenorrhoea or amenorrhoea; and serum T, FSH and LH in males with suspected hypogonadism.

In 2014, ACOG (24) listed measuring FSH and estradiol (two random tests at least 1 month apart) when there has been menstrual irregularity for at least three consecutive months as a way of diagnosing and initially evaluating primary ovarian insufficiency. For example, it identifies measuring FSH and estradiol levels to rule out causes of ovarian insufficiency such as pregnancy, thyroid disease and hyperprolactinaemia. It goes on to state that if gonadotropins are elevated into the menopausal range (typically, FSH levels of greater than 30–40 mIU/mL), a repeat FSH measurement should be done a month later; if the FSH is still elevated, a diagnosis of primary ovarian insufficiency can be established.

The ASRM (*10*) Practice Committee acknowledged that high FSH values have been associated with, but do not necessarily predict, both poor ovarian stimulation and the failure to conceive. It stated that "a single FSH value has very limited reliability because of inter- and intra-cycle variability (particularly if it is not elevated). An elevated FSH value has good specificity but may represent a false positive especially when used in a low-risk population. Given the inter-assay variability of FSH, the cut-off point selected by an IVF program ideally should be based on its own data or on data from studies using the same FSH assay".

G. Basic test characteristics (from application)

Test formats available	Immunoassays of varying formats, for example electrochemiluminescence and chemiluminescence
Specimen types	Serum, plasma
Equipment required	Automated analyser
Regulatory status	CE-marked and FDA-approved; also approved by other regulatory agencies, including Health Canada
Availability	Global
Price per test range	US$ 4.50–43 depending on the country
Instrument price range	US$ 35 000–85 000 depending on the instrument size

H. Evidence for diagnostic accuracy (from application)

A large systematic review (37 studies) published by Broekmans et al. in 2006 (*25*) evaluated clinical accuracy to predict IVF outcomes. Sensitivity was seen to vary and was highly dependent on FSH cut-off points. Cut-off points less than 10 IU/L had a sensitivity to predicting a poor outcome of 65–100%, with specificity of 5–100%. Positive likelihood ratios for a positive result ranged from

1.1 to 4.4 IU/L. The sensitivity to predicting non-pregnancy at cut-off points less than 10 IU/L was 26–78% with specificity of 27–87%. Positive likelihood ratios for a positive result using these lower cut-off points ranged from 1.1 to 6.8 IU/L.

The authors conclude that basal FSH may only be adequate for predicting poor response and non-pregnancy at high cut-off points. Further, it may not be suitable as a diagnostic test to exclude patients from undergoing IVF, but more as a screening test for counselling patients and determining further diagnostic steps.

Despite the review's limited support for the use of basal FSH levels in assisted reproductive procedures, LMICs offering IVF may favour using basal FSH at higher cut-off points as it is accessible and less costly than other diagnostic tools for predicting outcome.

I. Evidence for clinical usefulness and impact (from application)

A 2019 review by Mikhael et al. (*26*) suggests that, in general, basic evaluation of hypothalamic-pituitary failure includes measurement of FSH (with LH and estradiol). Overall management of hypothalamic-pituitary failure depends on patient needs. For example, in young women of reproductive age where FSH levels (with estradiol) are used to diagnose ovarian insufficiency, management can include reassurance and education about other reproductive options. This may also apply to patients with mosaic Turner syndrome who become aware of impending follicular atresia and who wish to preserve fertility through oocyte or embryo cryopreservation.

In cases where ovarian failure is confirmed, either premature or due to natural menopause, the diagnosis allows for monitoring the effects of estrogen deficiency and consideration of hormonal replacement to prevent or delay the associated morbidity.

Infertility. A 2003 WHO study (*27*) looked at high singleton live birth rate following classical ovulation induction in normogonadotrophic anovulatory infertility. All patients started with clomiphene for ovulation induction; those who did not ovulate within three treatment cycles of incremental daily doses up to 150 mg for five consecutive days or ovulatory CC patients who did not conceive within six cycles underwent gonadotropin induction of ovulation applying a step-down dose regimen. Of 240 consecutive women, there were 134 pregnancies ending in a singleton live birth (56% of women). The cumulative pregnancy rate after 12 and 24 months of follow-up was 50% and 71%, respectively. The study concluded that classical ovulation induction produces very good results in normogonadotrophic anovulatory infertility.

Akande et al. (*28*), looked at using FSH to assess the risk of IVF treatment cancellation due to poor ovarian response in 536 women undergoing IVF.

Modelling showed that high FSH and age were independent predictors of poor response to hyperstimulation and were independently associated with the risk of treatment cancellation.

In 2000, Creus et al. (*29*) looked at day 3 serum inhibin B, FSH and age as predictors of assisted reproduction treatment outcome. They studied 120 women undergoing assisted reproductive technology, including 40 consecutive cycles cancelled due to poor follicular response. The association of basal FSH with cancellation rate was significant with a predictive value of 79% and was independent of and stronger than the effects of age and inhibin B ($P < 0.05$). Basal FSH concentration was a better predictor of cancellation rate than age, but age was a stronger predictor of pregnancy rate.

Menstrual irregularities. The clinical community generally accepts that hormonal therapy is highly successful at treating DUB. A 2004 study by Hurskainen et al. (*30*) followed more than 200 women with DUB who were treated with a triphasic preparation of norgestimate and ethinyl E2. More than 80% were documented to have improvement in bleeding, and this was significantly increased over the placebo group. The authors concluded that the triphasic combination of norgestimate and ethinyl E2 is an effective treatment for metrorrhagic, menometrorrhagic, oligomenorrhoeic, and polymenorrheic dysfunctional uterine bleeding.

General reproductive function. Linglart et al. (*31*) looked at the use of recombinant growth hormone on 64 young girls with Turner syndrome. After 4 years a gain in mean height was observed compared with a decrease in mean height in the control group. The authors concluded that early treatment with growth hormone helps to prevent natural evolution towards short stature in most girls with Turner syndrome. In 2012, Lee et al. (*32*) studied 76 girls with PP who were treated with leuprolide acetate, a GnRH agonist, every 3 months. At 6 months, 98% exhibited LH suppression and 100% E2 suppression. The authors concluded that treatment with leuprolide acetate 3-month depot formulations (11.25 and 30 mg) effectively suppressed the GnRH axis.

J. Evidence for economic impact and/or cost–effectiveness (from application)

The cost of an FSH test in the USA is around US$ 23, compared with US$ 4–10 in India, US$ 22 in Australia and US$ 34 in Uganda. Prices among manufacturers are competitive and are not likely to differ significantly. By contrast, an AMH test costs around US$ 100 in the USA, US$ 17–28 in India, US$ 64 in Australia and US$ 80.50 in Uganda. It is several times more expensive to predict IVF success through ultrasound determination of follicle number and size.

An FSH measurement that prompts patients to abandon a cycle of assisted reproductive technology could result in significant savings (thousands of dollars in the USA).

It is also relatively inexpensive to determine whether anovulation has occurred using an FSH measurement, particularly in cases of amenorrhoea. Given the significant syndromes and health issues that are associated with anovulation, such as disorders of sexual development or PCOS, there should be no question that confirming anovulation and reducing its morbidity would be cost-effective. Identifying anovulatory women also allows preventive measures to be taken, for example to address estrogen deficiency or the metabolic sequelae of untreated chronic anovulation. The International Osteoporosis Foundation estimates the cost of osteoporotic fractures at €37 billion every year in the European Union, and US$ 19 billion in the USA. Estimates of the cost of diabetes and metabolic syndrome are less definable, but identifying those at risk and taking preventative measures should offset the cost of measuring reproductive hormone levels when the diagnosis is suspected.

K. Ethics, equity and human rights issues (from application)

FSH measurement plays a significant role in the diagnosis and management of infertility, particularly as it relates to assisted reproductive technology. But infertility may not be assigned a high priority in LMIC health care systems. To address this issue, WHO recommended in 2001 that infertility be viewed as a world health problem and encouraged the development of lower-cost assisted reproductive technology. Progress towards this end needs to be accelerated before the full potential of FSH as an infertility diagnostic test can be appreciated.

FSH levels are also used to identify DSDs attributable to hypogonadism, and identifying these disorders is essential to prevent further illness and death. But it is just as important to ensure that the resources needed to replace deficient hormones or to treat the disease are available, even in LMICs, to decrease overall morbidity. Identifying PP should similarly carry with it an obligation to provide treatment that delays puberty.

Measurement of FSH levels should be accessible to most global populations. Its use will reduce inequalities, provided that health care systems are both equipped and prepared to follow through and make available the treatments that are indicated by the results.

L. Summary of evidence evaluation

This evidence portfolio makes clear the need for FSH testing in conjunction with other tests for the indication of infertility and ovarian reserves. It is primarily built on the predictive value of FSH for ovarian response and/or pregnancy after IVF (to tailor evaluation and management in women suspected of having HPO axis disorders impacting fertility). FSH appears to be an essential part in the test

battery (together with LH and estrogen) with which to make these predictions. There are many assays; all showed adequate sensitivity and acceptable intra-assay variability.

Standardized assays show high specificity (83–100%) for predicting poor response to stimulation (usually defined as < 2–3 follicles or %4 retrieved oocytes) using multiple cut-off points above 10 IU/L (10–20 IU/L). But because FSH is measured in conjunction with other laboratory tests, it is difficult to evaluate the value or impact of FSH testing separately. So the evaluation of this submission should be done in conjunction with the submissions for the other basic tests required (LH and estrogen).

FSH testing is a basic element (alongside other laboratory tests) in tailoring the evaluation and management of women with infertility issues. There is evidence that the assays have significant inter- and intra-cycle variability, which limits their reliability. The overall correlation among different FSH assays is excellent, but absolute values can differ from one another. Clinicians may find it difficult to generalize FSH cut-off points reported in the medical literature to their practices unless they are using the very same assay and reference preparation. The accuracy in clinical practice depends on the cut-off point used.

M. Summary of SAGE IVD deliberations

The consequences of infertility from a personal and social point of view can be dramatic. Fertility is considered a major public health issue by WHO. And testing FSH levels is part of several international guidelines and best practice for a large number of indications in fertility medicine, gonadal disorders, anovulation, amenorrhoea and menstrual irregularities. One of the medicines that is routinely used in fertility – clomiphene – is also included in the complementary list of the EML.

There is, however, potential for overuse of the test, and little evidence base available to review for the clinical impact of FSH testing alone. SAGE IVD emphasized that FSH testing and result interpretation requires appropriate laboratory infrastructure and the availability of fertility or endocrinology specialists. The group also highlighted the importance of timing in sample collection (although timing requirements may vary depending on the condition being monitored). A WHO representative informed SAGE IVD that WHO is developing guidelines for diagnosing and managing infertility that will cover the nuances of sample collection for different conditions and circumstances. These are intended to be available later in 2020.

SAGE IVD noted that while FSH testing is critical to evaluate and manage infertility, additional investigations may be needed for a definitive diagnosis.

N. SAGE IVD recommendations

SAGE IVD recommended including the Follicle-Stimulating Hormone (FSH) test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, within a new subsection for endocrine disorders);
- using an immunoassay format;
- to aid in the diagnosis of primary and secondary amenorrhoea, anovulation, gonadal dysfunction and precocious puberty; and
- to aid in the evaluation and management of infertility.
- The group requested that the FSH entry in the EDL include a link to the new WHO guidelines for diagnosing and managing infertility as and when these are published.

The group further requested the addition of a note to the test category entry in the EDL stating that it is only recommended for use in specialized health care settings.

O. References

1. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:411–417.
2. Sinha SK, Hoffman RP. Precocious pseudopuberty. In: Medscape [website]. New York: Medscape; 2020 (https://emedicine.medscape.com/article/923876-overview-a6, accessed 27 March 2020).
3. Fleseriu M, Hashim IA, Karavitaki N, Melmed S, Hassan Murad M, et al. Hormonal replacement in hypopituitarism in adults: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2016;101(11):3888–3921. doi:10.1210/jc.2016-2118.
4. Viswanathan V, Eugaster EA. Etiology and treatment of hypogonadism in adolescents. Pediatr Clin North Am. 2011;58(5):1181. doi:10.1016/j.pcl.2011.07.009.
5. Taylor A. ABC of sub-fertility: making a diagnosis. BMJ. 2003;327(7413):494–497. doi:10.1136/bmj.327.7413.494.
6. Tal R, Seifer DB. Ovarian reserve testing: a user's guide. Am J Obstet Gynecol. 2017;217(2):129–140. doi:10.1016/j.ajog.2017.02.027.
7. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:396.
8. Hsu M, Liou TH, Liang SJ. Inappropriate gonadatropic secretion in polycystic ovary syndrome. Fertil Steril. 2009;91(4):1168–1174. doi:10.1016/j.fertnstert.2008.01.036.
9. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:326.
10. Practice Committee of the American Society of Reproductive Medicine. Testing and interpreting measures of ovarian reserve: a committee opinion. Fertil Steril. 2015;103:e9–e17. doi:10.1016/j.fertnstert.2014.12.093.

11. Hendriks DJ, Mol BW, Bancsi LF, te Velde ER, Broekmans FJ. The clomiphene citrate challenge test for the prediction of poor ovarian response and nonpregnancy in patients undergoing in vitro fertilization: a systematic review. Fertil Steril. 2006;86(4):807–818. doi:10.1016/j.fertnstert.2006.03.033.
12. Infertility workup for women's health specialist: ACOG Committee opinion, number 781. Obstet Gynecol. 2019;133(6):e377–e384. doi:10.1097/AOG.0000000000003271.
13. Polesel DN, Nozoe KT, Sanchez MZ, Pardo MCO, Bittencourt LR, et al. The follicle-stimulating hormone as best classifier for diagnosis of natural menopause. J Gynecol Surg. 2017;33(6):236–242. doi:10.1089/gyn.2017.0040.
14. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility. 9th ed. Philadelphia: Wolters Kluwer; 2019:399.
15. Kaplowitz PB, Hoffman RP. What is the role of gonadotropin testing in the diagnosis of precocious puberty? In: Medscape [website]. New York: Medscape; 2020 (https://www.medscape.com/answers/924002-95756/what-is-the-role-of-gonadotropin-testing-in-the-diagnosis-of-precocious-puberty, accessed 27 March 2020).
16. Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. PLoS Med. 2012;9(12):e1001356. doi:10.1371/journal.pmed.1001356.
17. Rutstein SO, Shah IH. Infecundity, infertility, and childlessness in developing countries. Calverton: ORC Macro; 2004.
18. Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. Hum Reprod. 2007;22:1506–1512. doi:10.1093/humrep/dem046.
19. Ombelet W. Is global access to infertility care realistic? The Walking Egg Project. Reprod Biomed Online. 2014;28(3):267–272. doi:10.1016/j.rbmo.2013.11.013.
20. Nachtigall RD. International disparities in access to infertility services. Fertil Steril. 2006;85(4):871–875. doi:10.1016/j.fertnstert.2005.08.066.
21. Kwak Y, Kim Y, Baek KA. Prevalence of irregular menstruation according to socioeconomic status: a population based nationwide cross-sectional study. PLoS One. 2019;14(3):e0214071. doi:10.1371/journal.pone.0214071.
22. Wolf WM, Wattick RA, Kinkade ON, Olfert MD. Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity. Int J Environ Res Public Health. 2018;15(11):2589. doi:10.3390/ijerph15112589.
23. Warne GW, Raza J. Disorders of sex development (DSDs), their presentation and management in different cultures. Rev Endocr Metab Disord. 2008;9:227–236. doi:10.1007/s11154-008-9084-2.
24. Committee opinion number 605: primary ovarian insufficiency in adolescents and young women. Obstet Gynecol. 2014;124(1):193–197. doi:10.1097/01.AOG.0000451757.51964.98.
25. Broekmans FJ, Kwee J, Hendriks DJ, Mol BW, Lambalk CB. A systematic review of tests predicting ovarian reserve and IVF outcome. Hum Reprod Update. 2006;12(6):685–718. doi:10.1093/humupd/dml034.
26. Mikhael S, Punjala-Patel A, Gavrilova-Jordan L. Hypothalamic-pituitary-ovarian axis disorders impacting female fertility. Biomedicines. 2019;7(1):5. doi:10.3390/biomedicines7010005.
27. Eijkimans MJC, Imani B, Mulders AGMGJ, Habbema JDF, Fauser BCJM. High singleton live birth rate following classical ovulation induction in normogonadotrophic anovulatory infertility (WHO 2). Hum Reprod. 2003;18(11):2357–2362. doi:10.1093/humrep/deg459.

28. Akande VA, Keay SD, Hunt LP, Mathur RS, Jenkins JM, et al. The practical implications of a raised serum FSH and age on the risk of IVF treatment cancellation due to a poor ovarian response. J Assist Reprod Genet. 2004;21(7):257–262. doi:10.1023/b:jarg.0000042011.67696.85.

29. Creus M, Peñarrubia J, Fábregues F, Vidal E, Carmona F, et al. Day 3 serum inhibin B and FSH and age as predictors of assisted reproduction treatment outcome. Hum Reprod. 2000;15(11):2341–2346. doi:10.1093/humrep/15.11.2341.

30. Hurskainen R, Teperi J, Rissanen P, Aalto AM, Grenman S, et al. Clinical outcomes and costs with the levonorgestrel-releasing intrauterine system or hysterectomy for treatment of menorrhagia: randomized trial 5-year follow-up. JAMA 2004;291:1456–1463. doi:10.1001/jama.291.12.1456.

31. Linglart A, Cabrol S, Berlier P, Stuckens C, Wagner K, et al. Growth hormone treatment before the age of 4 years prevents short stature in young girls with Turner syndrome. Eur J Endocrinol. 2011;164(6):891–897. doi:10.1530/EJE-10-1048.

32. Lee PA, Klein K, Mauras N, Neely EK, Bloch CA, et al. Efficacy and safety of leuprolide acetate 3-month depot 11.25 milligrams or 30 milligrams for the treatment of central precocious puberty. J Clin Endocrinol Metab. 2012;97(5):1572–1580. doi:10.1210/jc.2011-2704.

3.14. Luteinizing hormone

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding serum LH levels to the EDL as an IVD to aid in the diagnosis of diseases caused by malfunction of the HPG axis; and to aid the clinical evaluation of infertility (both female and male), menstrual irregularities, pituitary disorders, precocious or delayed puberty, and ovarian/testicular dysfunction.

B. Applicant

World Health Organization

C. WHO technical department

Sexual and Reproductive Health and Research
Contraception and Fertility Care

D. Background (from application)

Disease condition and impact on patients

LH is a polypeptide hormone produced by the anterior pituitary gland. It induces ovulation and the secretion of progesterone in women and stimulates the production of testosterone in men. LH works synergistically with FSH. The

measurement of LH levels, often alongside FSH, is used to aid in the diagnosis of DSDs – for example Turner and Klinefelter syndromes, and PCOS – and to evaluate the causes of infertility, amenorrhoea and irregular menstrual bleeding, and suspected Leydig cell insufficiency (*1*).

DSDs include hypergonadotropic hypogonadism, which is most commonly caused by Turner syndrome in females and Klinefelter syndrome in males. Children with hypogonadism experience delayed puberty if left untreated. Timely diagnosis of hypogonadism allows treatment with prepubertal hormone (sex steroid) replacement (*2*). For girls, administering estrogen allows secondary sex characteristics to emerge as well as adequate breast and uterine development. Early treatment with human growth hormone may also significantly increase height. For boys, testosterone administration is used to induce puberty (*2*). Measuring LH after stimulation with GnRH can also aid in the diagnosis of suspected PP (*3*). PP is defined as the onset of puberty before 8 years of age in girls and 9 years of age in boys (*4*). The psychosocial and physical burdens of undiagnosed and untreated delayed or precocious puberty are significant, and it is important that all tools to aid in a timely diagnosis be available.

PCOS is a complex disorder characterized by chronic anovulation. Common symptoms include irregular menstrual cycle, polycystic ovaries and hirsutism; less common symptoms include infertility, insulin resistance, impaired glucose tolerance (type 2 diabetes) and dyslipidaemia (5). Multiple diagnostic criteria have been adopted for PCOS, but the common denominator appears to be oligoovulation and androgen excess.

Infertility diagnostics and management use LH levels to document ovulation. Recent progress in urinary LH testing, which allows qualitative measurement on specimens applied to paper and various materials, has supplanted serum LH as a useful tool (*6*). Quantitative serum LH levels may, however, still be used to guide management decisions during IVF (*7*). In males, LH levels are used to evaluate azoospermia (*8*).

Menstrual disorder diagnosis can use LH levels to evaluate irregular menstrual bleeding, particularly oligomenorrhoea and amenorrhoea, by helping to distinguish ovulatory dysfunction from uterine abnormalities as the cause. The use of qualitative urinary LH measurement to evaluate ovulation often now replaces serum levels.

Does the test meet a medical need?

DSDs. LH levels, often with FSH, are used to distinguish between primary gonadal failure (hypergonadotropic hypogonadism) and deficient gonadal stimulation (hypogonadotropic hypogonadism) (*9*). Early diagnosis of hypogonadal syndromes

allows timely evaluation of the associated anomalies. Administration of growth hormone to young girls in childhood and hormone replacement at puberty for both girls and boys can reduce morbidity and mortality, as well as improve the quality of life (*2*).

LH levels are also used in diagnosing PP. In gonadotropin-dependent PP, elevated LH levels rule out adrenal hyperplasia or CNS lesions (10). In gonadotropin-independent PP, investigating the source of exogenous estrogens can lead to a more timely diagnosis of gonadal hormone-producing cysts or tumours.

Early diagnosis of DSDs can reduce morbidity and mortality and improve physical and psychological well-being. Failure to diagnose and treat these conditions, particularly in LMICs, can have a negative economic impact on families where children who are essential to their parents' survival must be economically independent, meaning that they are employable and able to marry (*11*). Despite the importance of addressing DSDs, there is considerable variation in how these disorders are managed in HICs and LMICs (*11*).

PCOS. LH results facilitate the diagnosis and treatment of PCOS with the short-term benefit of restoring ovulation and HPO balance (*12*). Appropriate diagnosis of PCOS allows hormonal treatments that can regulate menstrual cycles, induce ovulation when indicated, decrease androgenic effects and also reduce the risks of metabolic abnormalities that frequently occur with PCOS (*5*).

Equally important are the preventative long-term benefits of reducing the risk of developing diabetes and the metabolic syndrome (*5*). With the alarmingly rapid increase in obesity and diabetes in LMICs (*13*), rigorous attempts to diagnose and treat PCOS will have a significant positive impact on public health.

Menstrual disorders. Anovulatory or dysfunctional uterine bleeding can impact women's health and well-being by causing unscheduled and often excessive bleeding. Measuring LH levels is part of the laboratory investigation needed to arrive at the diagnosis. Once the diagnosis is established, bleeding can be controlled by administering appropriate hormonal therapy and carries a high degree of success. Morbidity is decreased by reducing the risk of anaemia due to blood loss and endometrial hyperplasia from estrogen overstimulation (*5*). Appropriate hormonal replacement also helps prevent the sequelae of estrogen deficiency. Economic impact is also lessened by reducing absences from work or school; social isolation, often directed at menstruating women in LMICs, is avoided too.

Infertility. LH levels are important in evaluating male infertility due to azoospermia. The public health impact of addressing this cause of infertility is essential for preserving the nuclear family, but may be particularly important in those countries with a patriarchal culture.

How the test is used

DSDs. When a diagnosis of hypogonadism is suspected based on history, physical exam and basal gonadotropin levels, exogenous administration of GnRH is often used to differentiate primary gonadal failure (hypergonadotropic) from deficient gonadal stimulation (hypogonadotropic hypogonadism) (*9*). LH levels will be totally or partially absent in hypogonadotropic cases but will gradually appear in hypergonadotropic ones. Measuring LH after GnRH stimulation can also help diagnose suspected PP (*3*). An increase in LH levels in response to a GnRH challenge is indicative of central PP (gonadotropic dependent), while no increase supports a diagnosis of peripheral (gonadotropic independent) PP. Early diagnosis of central PP allows for treatment with GnRH agonists that can delay the onset of puberty until a more appropriate age.

PCOS. An elevated ratio of LH to FSH (greater than 3) is often used with history and physical exam to support a diagnosis of PCOS (*12*).

Infertility. Elevated LH levels can indicate decreased ovarian function and consequent anovulation. If detected during an infertility workup, elevated LH levels could further suggest a diagnosis of PCOS and prompt confirmation of the presence of polycystic ovaries by ultrasound exam (*14*). Measuring LH levels may play a more significant role in evaluating male infertility by determining the etiology of azoospermia; low levels indicate a central dysfunction, while elevated levels are indicative of gonadal failure (*14*).

Menstrual disorders. Irregular menses, particularly amenorrhoea and oligomenorrhoea, are generally indicative of a defect at some point in the HPO axis, and LH levels can be used to evaluate the cause. Elevated LH levels indicate ovarian dysfunction or failure, while normal or low levels suggest pituitary or hypothalamic failure (*15*).

E. Public health relevance (from application)

Prevalence and socioeconomic impact

DSDs. Turner syndrome has an incidence of 1 per 2500 live births (*2*). Klinefelter syndrome has an incidence of 1 per 1000 live births (*16*). HH is rarer. The incidence of PP is estimated at 1 per 5000–10 000 population. Gonadotropin-independent PP is about five times less common than gonadotropin-dependent PP (*17*).

PCOS. PCOS is thought to be the most common endocrine disorder found in women of reproductive age and impacts all races and ethnicities. In unspecified populations, PCOS has a reported incidence rate of 3–10% (*18*), although more precise incidence is unknown due to underdiagnosis.

Infertility. Estimates suggest that between up to 186 million women globally are infertile (*8*, *16*). Although differing methods and definitions have been used to derive infertility burden, a recent review estimated global prevalence of infertility to be 9% (*19*). Data on worldwide prevalence of childlessness further estimate that as many as 70 million couples would benefit from medical intervention to achieve pregnancy (*20*). There are not enough data to assess global infertility prevalence trends over the past 20 years. But these trends can be impacted by a rise in STIs with subsequent impaired reproductive organ function, lifestyle changes and delayed childbearing (*8*, *16*, *19*).

The economic and social impact of infertility is significant, particularly for women, who often suffer from social isolation, discrimination, disinheritance, depression, abuse, divorce and possible abandonment in old age. Infertility as a common cause of childlessness can also have a broader negative economic impact on families, particularly in LMICs, where children contribute to family incomes and older parents depend on their children for support.

Infertility is recognized as an essential component of reproductive health by the UN Programme of Action of the International Conference on Population and Development (*21*). Paradoxically, nations with the highest overall fertility are also the ones with the greatest prevalence of infertility; these often include LMICs. Given the economic, resource, cultural and religious constraints, infertility services among countries vary significantly. Assays that measure serum LH levels are, however, relatively non-invasive (requiring only a blood draw or finger stick), inexpensive and accessible from laboratories throughout the world.

Menstrual disorders. Estimates of the prevalence of menstrual disorders range from 5% to ~36% (*22*); occurrence depends on age, nutritional status and country of residence.

F. WHO or other clinical guidelines relevant to the test

A 2016 Endocrine Society clinical practice guideline on hormonal replacement in hypopituitarism in adults (*9*) recommends measuring serum estradiol, FSH and LH in females with oligomenorrhoea or amenorrhoea; and serum T, FSH and LH in males with suspected hypogonadism.

In 2003, the Rotterdam European Society of Human Reproduction and Embryology (ESHRE)/ASRM-Sponsored PCOS Consensus Workshop Group published a revised consensus on diagnostic criteria and long-term health risks related to PCOS (*23*). It states that "serum LH levels should not be considered necessary for the clinical diagnosis of PCOS. LH levels could be useful as a secondary parameter (especially in lean women with amenorrhea or in research)".

In 2018, ACOG (*24*) identified gonadotropin measures to determine the cause of amenorrhoea as an optional test to consider.

A 2013 Endocrine Society clinical practice guideline on the diagnosis and treatment of POCs (25) lists serum LH and FSH in a table of diagnoses to consider for excluding hypothalamic amenorrhoea in select women, depending on presentation.

G. Basic test characteristics (from application)

Test format	Immunoassays of varying formats, for example electrochemiluminescence or chemiluminescence.
Specimen types	Serum, plasma
Equipment required	Automated analyser
Regulatory status	CE-marked and FDA-approved; also approved by other regulatory agencies, including Health Canada
Availability	Global
Price per test range	US$ 4.50–43, depending on the country
Instrument price range	US$ 35 000–85 000, depending on the instrument size

H. Evidence for diagnostic accuracy (from application)

A 2012 review by Harrington and Palmert (*26*) identified a total of 19 studies published in English over the past 30 years whose primary objective was to assess the sensitivity of a diagnostic test in differentiating HH from constitutional delay of growth and puberty (CDGP) in adolescents. Although some studies reviewed confirmed the utility of basal LH in discriminating between complete HH and CDPG, consistent results were not observed due perhaps to assay variability. In administering stimulation testing, LH levels following GnRH agonists provided better discrimination. A single inhibin B level was also investigated and showed better predictive value in some studies. In conclusion, basal and stimulated LH levels have limited utility in distinguishing HH from CDPG. Inhibin B may provide better results, but further investigation is needed.

A second review by Kalia et al. (*27*) describes and compares various assay methods used in serum LH measurement in varied clinical conditions. The authors review different methodologies with benefits and limitations, emphasizing reproducibility and sensitivity.

Houk et al. (*28*) evaluated a single basal LH measurement using a commercially available assay and reported a sensitivity of 93% and specificity of 100% for diagnosing central PP.

Resende et al. (*29*) measured basal and GnRH-stimulated LH levels of 315 prepubertal and pubertal children with an immunochemiluminometric assay. Basal LH levels were able to differentiate pubertal from prepubertal stage

in boys; in girls, GnRH-stimulated levels were needed to diagnose maturity of the HPG axis.

I. Evidence for clinical usefulness and impact (from application)

Dunkel et al. (*30*) describe various etiologies, and methods for timely diagnosis, of delayed puberty. When the diagnosis is confirmed, the management to induce puberty is also reviewed in detail.

Dwyer et al. (*31*) discuss the psychological impact of hypogonadism, the treatment and the transition to adulthood.

Chen et al. (*32*) review diagnosis and management of precocious puberty, with treatment outcomes.

DSDs. A study by Linglart et al. (*33*) looked at 64 young girls with Turner syndrome who were treated with recombinant growth hormone. After 4 years a gain in mean height was observed compared with a decrease in mean height in the control group. The authors concluded that early treatment with growth hormone helps to prevent natural evolution towards short stature in most girls with Turner syndrome. This emphasizes the importance of early diagnosis.

Lee et al. (*34*) looked at 76 girls with PP who were treated with leuprolide acetate, a GnRH agonist, every 3 months. At 6 months, 98% exhibited LH suppression and 100% had estradiol suppression. The authors concluded that treatment with leuprolide acetate 3-month depot formulations (11.25 and 30 mg) effectively suppressed the GnRH axis, again demonstrating the clinical importance of a clear diagnosis of PP.

PCOS. A study by Nestler et al. (*35*) looked at the effects of metformin on spontaneous and clomiphene-induced ovulation in PCOS. To restore ovulatory function, women with PCOS were given clomiphene plus metformin vs clomiphene alone. Overall, 89% of women who received metformin had spontaneous ovulation compared with 12% with clomiphene alone. The authors concluded that the ovulatory response to clomiphene can be increased in women with PCOS by decreasing insulin secretion with metformin.

Menstrual disorders. The clinical community generally accepts that hormonal therapy is highly successful in treating dysfunctional uterine bleeding. Hurskainen et al. (*36*) followed more than 200 women with dysfunctional uterine bleeding who were treated with a triphasic preparation of norgestimate and ethinyl E2. More than 80% had documented improvement in bleeding, and this was significantly increased over the placebo group. The conclusion was that the triphasic combination of norgestimate and ethinyl E2 is an effective treatment for metrorrhagic, menometrorrhagic, oligomenorrhagic, and polymenorrheic dysfunctional uterine bleeding.

J. Evidence for economic impact and/or cost–effectiveness (from application)

The cost (in US dollars) of an LH test in the USA is around US$ 23, compared with US$ 4–10 in India, US$ 22 in Australia and US$ 34 in Uganda. Prices among manufacturers are competitive and are not likely to differ significantly.

DSDs. LH measures are only somewhat successful in discriminating between HH and CDPG, while inhibin B had better predictive value. But inhibin B may not be available in LMICs and where available, the cost will be higher than an LH test. Measuring LH levels for suspected DSDs, in addition to physical exam and medical history, is a cost-effective step towards diagnosis and management.

Menstrual disorders. An LH measurement is a relatively inexpensive test that can determine whether anovulation has occurred, particularly in cases of amenorrhoea. Given the significant syndromes and health issues that are associated with anovulation, such as DSDs or PCOS, there should be no question that confirming anovulation and reducing its morbidity would be cost-effective. Identifying anovulatory women also allows preventative measures to be taken, for example to address estrogen deficiency or the metabolic sequelae of untreated chronic anovulation. The International Osteoporosis Foundation estimates the cost of osteoporotic fractures, which are increased in the hypoestrogen state, at €37 billion every year in the European Union, and US$ 19 billion each year in the USA. Estimates of the cost of diabetes and metabolic syndrome are less definable, but identifying those at risk and taking preventative measures should offset the cost of measuring reproductive hormone levels when the diagnosis is suspected.

K. Ethics, equity and human rights issues (from application)

LH measurement plays a significant part in the diagnosis of DSDs attributable to hypogonadism, and identifying these disorders is essential to prevent further illness, death and economic impact. But it is just as important to ensure that the resources needed to replace deficient hormones or to treat the disease are available, even in LMICs, to decrease overall morbidity. Identifying PP should similarly carry with it an obligation to provide treatment that delays puberty.

LH levels are also used to diagnose and manage infertility, particularly as it relates to assisted reproductive technology. But infertility may not be assigned a high priority in LMIC health care systems. To address this issue, WHO recommended in 2001 that infertility be viewed as a world health problem and encouraged the development of lower-cost assisted reproductive technology. Progress towards this end needs to be accelerated before the full potential of LH as an infertility diagnostic test can be appreciated.

Measurement of LH levels should be accessible to most global populations. Its use will reduce inequalities in health care, provided that systems delivering

health care are both equipped and prepared to follow through and make available the treatments that are indicated by the results.

L. Summary of evidence evaluation

The LH test is proposed to aid in the diagnoses of various diseases, including DSDs and PCOS, to evaluate the causes of infertility, amenorrhoea and/or irregular menstrual bleeding and suspected Leydig cell insufficiency. Although guidelines and expert opinions state that LH tests are necessary to diagnose these conditions, there is very limited evidence on their diagnostic accuracy and no evidence on their clinical utility on their own.

M. Summary of SAGE IVD deliberations

The consequences of infertility from a personal and social point of view can be dramatic. Fertility is considered a major public health issue by WHO. And testing LH levels is part of several international guidelines and best practice for a large number of indications in fertility medicine, gonadal disorders, anovulation, amenorrhoea and menstrual irregularities. One of the medicines that is routinely used in fertility – clomiphene – is also included in the complementary list of the EML.

There is, however, potential for overuse of the test. And there is little evidence available to review for the clinical impact of LH testing. SAGE IVD emphasized that LH testing goes hand in hand with FSH testing; and that result interpretation requires appropriate laboratory infrastructure and the availability of fertility or endocrinology specialists. The group also highlighted the importance of timing in sample collection (although timing requirements may vary depending on the condition being monitored).

SAGE IVD noted that the specimen types listed in the LH submission to the EDL did not include urine, even though urine LH is often used in practice, appears in at least one international guideline and is the subject of a published Cochrane review.

The group further noted that while LH testing is critical to evaluate and manage infertility, additional investigations may be needed for a definitive diagnosis.

N. SAGE IVD recommendations

SAGE IVD recommended including the Luteinizing Hormone (LH) test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, within a new subsection for endocrine disorders);
- using an immunoassay format;

- to aid in the diagnosis of primary and secondary amenorrhoea, anovulation, gonadal dysfunction and precocious puberty; and
- to aid in the evaluation and management of infertility.
- The group requested the addition of a note to the test category entry in the EDL stating that it is only recommended for use in specialized health care settings.

The group further requested that the original submitter of the LH test category be asked to provide information on the applications of LH in urine to support an edit of the EDL entry next year.

O. References

1. Reproductive hormones: the right test, at the right time, for the right patient. Best Practice Advocacy Centre New Zealand (bpacnz); 2013.
2. Viswanathan V, Eugaster EA. Etiology and treatment of hypogonadism in adolescents. Pediatr Clin North Am. 2011;58(5):1181–1200. doi:10.1016/j.pcl.2011.07.009.
3. Kaplowitz PB, Hoffman RP. What is the role of gonadotropin testing in the diagnosis of precocious puberty? In: Medscape [website]. New York: Medscape; 2020 (https://www.medscape.com/answers/924002-95756/what-is-the-role-of-gonadotropin-testing-in-the-diagnosis-of-precocious-puberty, accessed 27 March 2020).
4. Precocious puberty. In: Mayo Clinic/Patient care and health information [website]. Minnesota: Mayo Clinic; 2020 (https://www.mayoclinic.org/diseases-conditions/precocious-puberty/symptoms-causes/syc-20351811, accessed 27 March 2020).
5. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:411–417.
6. Su HW, Yi YC, Wei TY, Chang TC, Cheng CM. Detection of ovulation, a review of currently available methods. Bioeng Transl Med. 2017;2(3):238–246. doi:10.1002/btm2.10058.
7. Kolibianakis EM, Collins J, Tarlatzis B, Papanikolaou E, Devroey P. Are endogenous LH levels during ovarian stimulation for IVF using GnRH analogues associated with the probability of ongoing pregnancy? A systematic review. Hum Reprod Update. 2006;12(1):3–12. doi:10.1093/humupd/dmi030.
8. Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. PLoS Med. 2012;9(12):e1001356. doi:10.1371/journal.pmed.1001356.
9. Fleseriu M, Hashim IA, Karavitaki N, Melmed S, Hassan Murad M, et al. Hormonal replacement in hypopituitarism in adults: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2016;101(11):3888–3921. doi:10.1210/jc.2016-2118.
10. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:326.
11. Warne GW, Raza J. Disorders of sex development (DSDs), their presentation and management in different cultures. Rev Endocr Metab Disord. 2008;9:227–236. doi:10.1007/s11154-008-9084-2.
12. Hsu M, Liou TH, Liang SJ. Inappropriate gonadatropic secretion in polycystic ovary syndrome. Fertil Steril. 2009;91(4):1168–1174. doi:10.1210/jcem.82.11.4377.

13. Shen J, Kondal D, Rubinstein A, Irazola V, Gutierrez L, et al. A multiethnic study of pre-diabetes and diabetes in LMIC. Glob Heart. 2016 Mar;11(1):61–70. doi:10.1016/j.gheart.2015.12.015.

14. Taylor A. Making a diagnosis. BMJ. 2003;327(7413):494–497. doi:10.1136/bmj.327.7413.494.

15. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:399.

16. Rutstein SO, Shah IH. Infecundity, infertility, and childlessness in developing countries. Calverton: ORC Macro; 2004.

17. Sinha SK, Hoffman RP. Precocious pseudopuberty. In: Medscape [website]. New York: Medscape; 2020 (https://emedicine.medscape.com/article/923876-overview-a6, accessed 27 March 2020).

18. Wolf WM, Wattick RA, Kinkade ON, Olfert MD. Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity. Int J Environ Res Public Health. 2018;15(11):2589. doi:10.3390/ijerph15112589.

19. Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. Hum Reprod. 2007;22:1506–1512. doi:10.1093/humrep/dem046.

20. Ombelet W. Is global access to infertility care realistic? The Walking Egg Project. Reprod Biomed Online. 2014;28(3):267–272. doi:10.1016/j.rbmo.2013.11.013.

21. Nachtigall RD. International disparities in access to infertility services. Fertil Steril. 2006;85(4):871–875. doi:10.1016/j.fertnstert.2005.08.066.

22. Kwak Y, Kim Y, Baek KA. Prevalence of irregular menstruation according to socioeconomic status: a population-based nationwide cross-sectional study. PLoS One. 2019;14(3):e0214071. doi:10.1371/journal.pone.0214071.

23. Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil Steril. 2004;81(1):19–25. doi:10.1016/j.fertnstert.2003.10.004.

24. ACOG Practice Bulletin No. 194: polycystic ovary syndrome. Obstet Gynecol. 2018;131(6):e157–e171. doi:10.1097/AOG.0000000000002656.

25. Legro RS, Arslanian SA, Ehrmann DA, Hoeger KM, Murad MH, et al. Diagnosis and treatment of polycystic ovary syndrome: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2013;98(12):4565–4592. doi:10.1210/jc.2013-2350.

26. Harrington J, Palmert MR. Clinical review: distinguishing constitutional delay of growth and puberty from isolated hypogonadotropic hypogonadism: critical appraisal of available diagnostic tests. J Clin Endocrin Metab. 2012;97(9):3056–3067. doi:10.1210/jc.2012-1598.

27. Kalia V, Jadhav AN, Bhutani KK. Luteinizing hormone estimation. Endocr Res. 2009;30(1):1–17. doi:10.1081/erc-120029888.

28. Houk CP, Kunselman AR, Lee PA. Adequacy of a single unstimulated luteinizing hormone level to diagnose central precocious puberty in girls. Pediatrics. 2009;123(6):e1059–e1063. doi10.1542/peds.2008-1180.

29. Resende EAMR, Lara BHJ, Reis JD, Ferreira BP, Pereira GA, et al. Assessment of basal and gonadotropin-releasing hormone-stimulated gonadotropins by immunochemiluminometric and immunofluorometric assays in normal children. J Clin Endocrinol Metab. 2007;92(4):1424–1429. doi:10.1210/jc.2006-1569.

30. Dunkel L, Quinton R. Transition in endocrinology: induction of puberty. Eur J Endocrinol. 2014;170(6):R229–239. doi:10.1530/EJE-13-0894.

31. Dwyer AA, Phan-Hug F, Hauschild M, Elowe-Gruau E, Pitteloud N. Transition in endocrinology: hypogonadism in adolescence. Eur J Endocrinol. 2015;173:15–24. doi:10.1530/EJE-14-0947.

32. Chen M, Eugster EA. Central precocious puberty: update on diagnosis and treatment. Paediatr Drugs. 2015;17(4):273–281. doi:10.1007/s40272-015-0130-8.

33. Linglart A, Cabrol S, Berlier P, Stuckens C, Wagner K, et al. Growth hormone treatment before the age of 4 years prevents short stature in young girls with Turner syndrome. Eur J Endocrinol. 2011;164(6):891–817. doi:10.1530/EJE-10-1048.

34. Lee PA, Klein K, Mauras N, Neely EK, Bloch CA, et al. Efficacy and safety of leuprolide acetate 3-month depot 11.25 milligrams or 30 milligrams for the treatment of central precocious puberty. J Clin Endocrinol Metab. 2012;97(5):1572–1580. doi:10.1210/jc.2011.2704.

35. Nestler JE, Jakubowicz DJ, Evans WS, Pasquali R. Effects of metformin on spontaneous and clomiphene-induced ovulation in the polycystic ovary syndrome. N Engl J Med. 1998;338:1876–1880. doi:10.1056/NEJM199806253382603.

36. Hurskainen R, Teperi J, Rissanen P, Aalto AM, Grenman S, et al. Clinical outcomes and costs with the levonorgestrel-releasing intrauterine system or hysterectomy for treatment of menorrhagia: randomized trial 5-year follow-up. JAMA. 2004;291:1456–1463. doi:10.1001/jama.291.12.1456.

3.15. Progesterone

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding serum progesterone levels to the EDL as an IVD to confirm evidence of ovulation and assess luteal phase function during fertility investigations.

B. Applicant

World Health Organization

C. WHO technical department

Sexual and Reproductive Health and Research
Contraception and Fertility Care

D. Background (from application)

Disease condition and impact on patients

Progesterone is a steroid hormone that is mainly formed in the cells of the corpus luteum of the ovary and during pregnancy in the placenta. Progesterone concentrations fluctuate during the menstrual cycle with levels barely detectable in the follicular phase and then rising with the increased synthesis that occurs

after ovulation in the luteal phase. A surge in progesterone occurs approximately 1 day before ovulation.

Progesterone increases the glandular component of the lining of the uterus (endometrium) in preparation to implant a fertilized ovum; during pregnancy, progesterone inhibits uterine contractions. Progesterone, in synergy with estrogen, also affects breast tissue by promoting the proliferation of the alveoli of the mammary gland.

Infertility. Anovulation may be related to subfertility, PCOS, genetic disorders, obesity and other etiologies (*1*). PCOS is a complex disorder characterized by chronic anovulation and accounts for 70% of anovulatory infertility (*2*).

Luteal phase deficiency (LPD). LPD corresponds to a duration of less than 11 days from ovulation until the onset of menses. Most experts believe it is a defect of corpus luteum progesterone output, both in amount and duration, that results in inadequate stimulation of the endometrium for implanting the fertilized ovum. It has also been implicated in early pregnancy loss.

Does the test meet a medical need?

Infertility. Although studies vary in support of serum progesterone levels to confirm ovulation, a single progesterone determination in the mid-luteal phase (approximately day 21 of the menstrual cycle or 7 days before anticipated menses) is frequently used (*2*). It has a reported sensitivity and specificity equal to or greater than 90% (*3*). A wide range of disorders can be considered once anovulation is confirmed, including hypogonadotropic and hypothalamic hypogonadism, premature or natural ovarian insufficiency, hyperprolactinaemia, thyroid dysfunction and late-onset congenital adrenal hyperplasia (*4*). All have a significant public health impact; but their diagnoses require more definitive testing than measuring progesterone levels.

LPD. Because LPD as a cause of infertility has yet to be proven (*5*), interpreting results and applying them to patient management is challenging. But, given that progesterone is produced by the follicle, the follicle's abnormal development could logically result in LPD. LPD may also impact fertility through an abnormal endometrial response to normal progesterone production. Despite the clear association of progesterone and the development of the corpus luteum, using progesterone to diagnose LPD is complicated by the pulsatile nature of progesterone secretion. Attempts to overcome this limitation by evaluating pooled luteal phase progesterone samples to improve sensitivity and specificity have been only somewhat successful (*6*). Nonetheless, clinicians have used empiric treatment with the understanding that evidence-based treatment regimens have not been developed.

Further supporting the concept of LPD and its association with infertility, studies indicate that treatment with ovulation-inducing agents such as clomiphene or gonadotropins that optimize follicular development and number has been used with some success (*7*, *8*).

Progesterone levels have also been used to predict the outcome of early pregnancy when complications occur. Early diagnoses of a non-viable pregnancy allow consideration of management options that may help avoid unanticipated haemorrhage or pain (*9*, *10*). Complications from spontaneous (and unsafe induced) abortions are recognized globally as a major public health concern.

PCOS. Determining progesterone levels in women suspected of having PCOS but who are not amenorrhoeic can confirm anovulation and support the diagnosis. Appropriately diagnosing PCOS allows hormonal treatments that can regulate menstrual cycles, induce ovulation when indicated, decrease androgenic effects and also reduce the risks of metabolic abnormalities that often occur with PCOS (*11*).

How the test is used

Progesterone levels are primarily used in fertility investigations to confirm ovulation, evaluate PCOS and assess the proper functioning of the luteal phase. Apart from fertility evaluation, measuring progesterone may help establish lack of ovulation; but it cannot diagnose the wide range of conditions associated with anovulation, including hypothalamic and pituitary dysfunction as well as genetic disorders. Because the diagnosis of these disorders will depend on more definitive tests, this submission of progesterone as a diagnostic device focuses on disorders related to reproductive dysfunction.

Infertility. Although there are many other tests that can signal ovulation (basal body temperature, cervical mucous, urinary LH testing), evaluating luteal phase progesterone level is often part of the basic workup of subfertile women to determine whether ovulation has occurred (*1*).

LPD. LPD is diagnosed through a combination of assessments that may include determining progesterone levels, assessing basal body temperature, biopsy of the endometrium or sonographic imaging. Although progesterone measurement is clinically indicated in evaluating LPD, experts disagree on the cut-off point for abnormal levels, as well as the frequency and timing of testing. The accuracy of diagnosing LPD may be improved by evaluating mid-luteal progesterone levels on pooled samples. Progesterone can also be used to predict the outcome of early pregnancy in women experiencing complications such as pain or bleeding (*12*).

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Infertility. Estimates suggest that up to 186 million women globally are infertile (*13*, *14*). Although differing methods and definitions have been used to derive infertility burden, a recent review estimated global prevalence of infertility to be 9% (*15*). Data on worldwide prevalence of childlessness further estimate that as many as 70 million couples would benefit from medical intervention to achieve pregnancy (*16*). There are not enough data to assess global infertility prevalence trends over the past 20 years. But these trends can be impacted by a rise in STIs with subsequent impaired reproductive organ function, lifestyle changes and delayed childbearing (*13*–*15*).

The economic and social impact of infertility is significant, particularly for women, who will often suffer from social isolation, discrimination, disinheritance, depression, abuse, divorce and possible abandonment in old age. Infertility as a common cause of childlessness can also have a broader negative economic impact on families, particularly in LMICs, where children contribute to family incomes and older parents depend on their children for support.

PCOS is thought to be the most common endocrine disorder found in women of reproductive age and impacts all races and ethnicities. In unspecified populations PCOS has a reported incidence rate of 3–10% (*17*), although more precise incidence is unknown due to underdiagnosis. Common symptoms include irregular menstrual cycle, polycystic ovaries and hirsutism (*18*), but may also include infertility, insulin resistance, impaired glucose tolerance (type 2 diabetes) and dyslipidaemia (*18*). Multiple diagnostic criteria have been adopted for PCOS, but the common denominator appears to be oligoovulation and androgen excess.

LPD. Estimates of LPD prevalence among infertile women vary considerably from around 4% to 30%, with variability attributed to lack of consensus on both its definition and diagnostic criteria. A shortened luteal phase is, however, reported to occur in 5% of ovulatory cycles and accounts for between 25% and 40% of recurrent pregnancy losses (*19*, *20*).

F. WHO or other clinical guidelines relevant to the test

There are no relevant guidelines for LPD.

The 2006 guidelines by the Royal College of Obstetricians and Gynaecologists (*12*) considers whether serum progesterone assay has a role in predicting pregnancy outcome. They state that such an assay can be a useful adjunct when ultrasound suggests pregnancy of unknown location. In these cases, serum progesterone levels below 25 nmol/L are associated with pregnancies subsequently

confirmed to be non-viable. But the guidelines warn against uterine evacuation based on a low initial progesterone, as viable pregnancies have been reported with initial levels less than 15.9 nmol/L.

In the presence of pregnancy of unknown location, a serum progesterone less than 20 nmol/L predicts spontaneous pregnancy resolution with a sensitivity of 93% and specificity of 94%. One advantage is that the need for formal uterine evacuation can be reduced with a policy of expectant management. Levels above 25 nmol/L are "likely to indicate" and above 60 nmol/L are "strongly associated with" pregnancies subsequently shown to be normal. Overall, it is not possible to define a specific discriminatory value for a single serum progesterone result that will allow absolute clinical confirmation of viability or non-viability.

G. Basic test characteristics (from application)

Test formats available	Immunoassays of varying formats, for example electrochemiluminescence and chemiluminescence
Specimen types	Serum, plasma
Equipment required	Automated analyser
Regulatory status	CE-marked and FDA-approved; also approved by other regulatory agencies, including Health Canada
Availability	Global
Price per test range	US$ 5.50–43 depending on the country
Instrument price range	US$ 35 000–85 000 depending on the instrument size

H. Evidence for diagnostic accuracy (from application)

A 2012 review by Verhaegen et al. (*21*) involved 19 cohort studies, including more than 7000 women, and evaluated the diagnostic accuracy of a single serum progesterone measurement to predict pregnancy outcomes in women experiencing pain or bleeding with inconclusive ultrasound exam or with pain or bleeding alone. Sensitivity and specificity were highly dependent on the progesterone threshold selected.

In women with pain or bleeding and an inconclusive ultrasound exam, a single progesterone measurement predicted a non-viable pregnancy with pooled sensitivity of 74.6% (95% CI: 50.6–89.4), specificity of 98.4% (95% CI: 90.9–99.7), positive likelihood ratio of 45 (7.1–289) and negative likelihood ratio of 0.26 (0.12–0.57).

In women with pain or bleeding alone, progesterone predicted a non-viable pregnancy:

- at a threshold of 10 ng/mL, with a pooled sensitivity of 66.5% (95% CI: 53.6–77.4), specificity of 96.3% (95% CI: 91.1–98.5), positive likelihood ratio of 18 (7.2–45), and negative likelihood ratio of 0.35 (0.24–0.50);
- at a threshold of 15 ng/mL, with a pooled sensitivity of 83.3% (95% CI: 66.6–92.6), specificity of 87.5% (95% CI: 78.5–93.1), positive likelihood ratio of 6.7 (3.8–12) and negative likelihood ratio of 0.35 (0.09–0.5); and
- at a threshold of 20 ng/mL, with pooled sensitivity of 85.7% (95% CI: 72.3–93.2), specificity of 66.6% (95% CI: 47–91.8), positive likelihood ratio of 2.6 (1.5–4.5) and negative likelihood ratio of 0.22 (0.1–0.47).

These findings show that a single progesterone measurement for women in early pregnancy presenting with bleeding or pain and inconclusive ultrasound assessments can rule out a viable pregnancy.

In 1996, McCord et al. (*22*) looked at the accuracy of screening serum progesterone to diagnose ectopic pregnancy (EP) and to identify a cut-off value that provides the best compromise between test sensitivity and specificity. The authors collected single progesterone measurements from 3674 pregnancies; they defined outcomes as EP, viable intrauterine pregnancy (IUP) and spontaneous abortion (SAB), and analysed the diagnostic accuracy of the test. They found that the diagnostic accuracy for:

- EP vs IUP was 88.7%, standard error of the mean (SEM) = 0.1%;
- SAB vs IUP was 93.8%, SEM = 0.4%; and
- SAB plus EP vs IUP was 92.8%, SEM = 0.4%.

They concluded that for progesterone > to 17.5 ng/mL, patients thought to be at risk for EP may be followed reasonably without ultrasound or further invasive diagnostic studies.

I. Evidence for clinical usefulness and impact (from application)

A 2011 review by Van der Linden et al. (*23*) describes various treatments during IVF for presumed LPD and concludes that progesterone supplementation seems to be an important aspect of any assisted reproductive treatment (ART).

Vaisbuch et al. (*24*) surveyed clinicians providing ART in 2014 and reported that all 408 centres across 82 countries used some form of progesterone for luteal phase support. Since evidence-based data are lacking, a review by Mesen et al. in 2015 (*25*) discusses the controversy over whether LPD is a diagnosable disorder that is proven to cause infertility. The authors note that

there are dual theories of LPD etiology (low levels of progesterone with a poorly functioning corpus luteum or a deficient endometrial response to normal hormonal levels), which may explain the lack of specific data to define LPD or to prove its association with infertility. The pulsatile nature of progesterone secretion also makes it hard to determine normal from subnormal levels. Although the review did not report a sensitivity or specificity for detecting LPD by progesterone level, a large study in normally menstruating women was cited where low progesterone levels were found to be significant for LPD by univariate analysis. The authors concluded that, despite the lack of evidence-based data, progesterone is critical for reproduction and LPD is a plausible cause of infertility and early pregnancy loss.

A prospective study by Schliep et al. in 2014 (*26*) followed 259 women 18–44 years of age for up to two menstrual cycles and concluded that identifying ovulation in combination with a well-timed luteal progesterone measurement may serve as a cost-effective and specific tool for clinicians and researchers to assess LPD.

J. Evidence for economic impact and/or cost–effectiveness (from application)

Progesterone measurement is an inexpensive test that can confirm ovulation as part of a basic infertility workup. Progesterone levels, despite the lack of evidence-based data, are used to support an LPD diagnosis, after which supplemental progesterone can be administered as a cost-effective way of facilitating conception. Supplementing ART cycles with progesterone can also decrease the cost of ART by increasing the success rate.

Using a single progesterone measurement to predict a non-viable pregnancy in women with bleeding or pain may be particularly useful and cost-effective in LMICs where ultrasound equipment is not readily available. Timely diagnosis of a non-viable pregnancy allows planning expectant or medical management as needed and will decrease the number of complications and associated costs.

K. Ethics, equity and human rights issues (from application)

Progesterone levels are used to diagnose and manage infertility, which may not be a high priority in LMIC health care systems. In 2001, WHO recommended viewing infertility as a world health problem and encouraged the development of lower-cost ARTs. Progesterone measurement can be used towards this goal and can ultimately reduce inequalities in health care as it relates to infertility and childlessness.

Using progesterone to identify non-viable pregnancies gives women the option of medical management even where ultrasound examination and other testing are not available.

L. Summary of evidence evaluation

The serum progesterone test is aimed at aiding the diagnosis of disorders related to reproductive dysfunction, more specifically infertility and anovulation, LPD and PCOS. No evidence is presented on the diagnostic accuracy or clinical utility of the serum progesterone test for these three conditions. Evidence provided shows that a single low-progesterone measurement for women in early pregnancy presenting with bleeding or pain and inconclusive ultrasound results can rule out a viable pregnancy.

M. Summary of SAGE IVD deliberations

The consequences of infertility from a personal and social point of view can be dramatic. Fertility is considered a major public health issue by WHO. Progesterone forms part of international guidelines and best practice to determine whether ovulation has taken place during infertility evaluation and treatment. Importantly, guidelines do not commonly support its use to assess LPD, or as the primary assessment tool to diagnose amenorrhoea or menstrual irregularities due to hypogonadism, hyperprolactinaemia, thyroid dysfunction or late-onset congenital adrenal hyperplasia.

Because progesterone testing has such limited indications, it is particularly prone to overuse. SAGE IVD emphasized that LH testing and result interpretation require appropriate laboratory infrastructure and the availability of fertility or endocrinology specialists.

N. SAGE IVD recommendations

SAGE IVD recommended including the progesterone test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, within a new subsection for endocrine disorders);
- using an immunoassay format;
- to confirm ovulation during infertility evaluation and treatment.
- The group requested the addition of a note to the test category entry in the EDL stating that it is only recommended for use in specialized health care settings.

The group further requested the submission of testosterone as a potential addition to the EDL, for full review and consideration next year.

O. References

1. Infertility workup for the women's health specialist. 2019. ACOG Committee opinion, number 781. Obstet Gynecol. 2019;133,(6):e377–e384. doi:10.1097/AOG.000000000000327.
2. Hamilton-Fairley D, Taylor A. Anovulation. BMJ. 2003 Sep 6;327(7414):546–549. doi:10.1136/bmj.327.7414.546.
3. Leiva R, Bouchard T, Boehringer H, Abulla S, Ecochard R. Random serum progesterone threshold to confirm ovulation. Steroids. 2015;101:125–129. doi:10.1016/j.steroids.2015.06.013.
4. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:399.
5. Practice Committee of the American Society for Reproductive Medicine. The clinical relevance of luteal phase deficiency: a committee opinion. Fertil Steril. 2012;98(5):1112–1117. doi:10.1016/j.fertnstert.2012.06.050.
6. Jordan J, Craig K, Clifton DK, Soules MR. Luteal phase defect: the sensitivity and specificity of diagnostic methods in common clinical use. Fertil Steril. 1994;62(1):54–62. doi:10.1016/s0015-0282(16)56815-0.
7. Guzick DS, Zeleznik A. Efficacy of clomiphene citrate in the treatment of luteal phase deficiency: quantity versus quality of preovulatory follicles. Fertil Steril. 1990;54(2):206–210. doi:10.1016/s0015-0282(16)53690-5.
8. Li TC, Dinga SH, Anstie B, Tuckerman E, Wood K, et al. Use of human menopausal gonadotropins in the treatment of endometrial defects associated with recurrent miscarriage: preliminary report. Fertil Steril. 2001;75(2):434–437. doi:10.1016/s0015-0282(00)01708-8.
9. American College of Obstetricians and Gynecologists. ACOG Committee opinion no. 427: misoprostol for postabortion care. Obstet Gynecol. 2009;113:465–468. doi:10.1097/AOG.0b013e31819930f9.
10. Department of Reproductive Health and Research. Safe abortion: technical and policy guidance for health systems. Geneva: World Health Organization; 2012.
11. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:411–417.
12. Hinshaw K, Fayyad A, Munjuluri P. The management of early pregnancy loss. Royal College of Obstetricians and Gynaecologists; 2006.
13. Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. PLoS Med. 2012;9(12):e1001356. doi:10.1371/journal.pmed.1001356.
14. Rutstein SO, Shah IH. Infecundity, infertility, and childlessness in developing countries. Calverton: ORC Macro; 2004.
15. Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment seeking: potential need and demand for infertility medical care. Hum Reprod. 2007;22:1506–1512. doi:10.1093/humrep/dem046.
16. Ombelet W. Is global access to infertility care realistic? The Walking Egg Project. Reprod Biomed Online. 2014;28(3):267–272. doi:10.1016/j.rbmo.2013.11.013.
17. Wolf WM, Wattick RA, Kinkade ON, Olfert MD. Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity. Int J Environ Res Public Health. 2018;15(11):2589. doi:10.3390/ijerph15112589.
18. Taylor HS, Pai L, Seli E, editors. Speroff's clinical gynecologic endocrinology and infertility, 9th edition. Philadelphia: Wolters Kluwer; 2019:411–417.

19. Filicori M, Butler JP, Crowley WF Jr. Neuroendocrine regulation of the corpus luteum in the human. Evidence for pulsatile progesterone secretion. J Clin Invest. 1984;73:1638–1647. doi:10.1172/JCI111370.

20. Wuttke W, Pitzel L, Seidlova-Wuttke D, Hinney B. LH pulses and the corpus luteum: the luteal phase deficiency (LPD). Vitam Horm. 2001;63:131–158. doi:10.1016/s0083-6729(01)63005-x.

21. Verhaegen J, Gallo ID, van Mello NM, Abdel-Aziz M, Takwoingi Y, et al. Accuracy of single progesterone test to predict early pregnancy outcome in women with pain or bleeding: meta-analysis of cohort studies. BMJ. 2012;345:e6077. doi:10.1136/bmj.e6077.

22. McCord ML, Muram D, Buster JE, Arheart KL, Stovall TG. Single serum progesterone as a screen for ectopic pregnancy: exchanging specificity and sensitivity to obtain optimal test performance. Fertil Steril. 1996;66(4):513–516. doi:10.1016/s0015-0282(16)58560-4.

23. Van der Linden M, Buckingham K, Farquhar C, Kremer JAM, Metwally M. Luteal phase support for assisted reproduction cycles. Cochrane Database Syst Rev. 2015;7;CD009154. doi:10.1002/14651858.CD009154.pub3.

24. Vaisbuch E, de Ziegler D, Leong M, Weissman A, Shoham Z. Luteal-phase support in assisted reproduction treatment: real-life practices reported worldwide by an updated website survey. Reprod Biomed Online. 2014;28(3):330–335. doi:10.1016/j.rbmo.2013.10.022.

25. Mesen TB, Young SL. Progesterone and the luteal phase: a requisite to reproduction. Obstet Gynecol Clin North Am. 2015;42(1):135–151. doi:10.1016/j.ogc.2014.10.003.

26. Schliep KC, Mumford SL, Hammoud AO, Stanford JB, Kissell KA, et al. Luteal phase deficiency in regularly menstruating women: prevalence and overlap in identification based on clinical and biochemical diagnostic criteria. J Clin Endocrinol Metab. 2014;99(6):E1007–E1014. doi:10.1210/jc.2013-353".

3.16. Prolactin

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding prolactin (PRL) levels (serum, plasma) to the EDL as an IVD for diagnosing prolactinoma and following up during prolactinoma management.

B. Applicant

British Columbia Children's Hospital
Global Pediatric Endocrinology and Diabetes

C. WHO technical department

Sexual and Reproductive Health and Research
Contraception and Fertility Care

D. Background (from application)

Disease condition and impact on patients

The main function of PRL lies in the development of mammary glands, milk synthesis and maintenance of milk secretion during pregnancy and lactation. Serum prolactin levels rise rapidly during pregnancy with increase in the size and number of lactotrophs. During lactation, suckling induces rapid secretion of PRL via a neuroendocrine reflex pathway. In the absence of pregnancy, hyperprolactinaemia (elevated levels of PRL) may present with symptoms of HH, including menstrual disturbance and infertility or visual symptoms from a pituitary mass effect by a prolactinoma, the most common pituitary tumour.

Prolactinoma is a benign tumour of the pituitary gland. It is divided into microadenoma (< 10 mm) or macroadenoma (> 10 mm). Symptoms include infertility, galactorrhoea (milk leaking from nipples), amenorrhoea or oligomenorrhoea, loss of libido, vaginal dryness, acne, hirsutism as well as visual or neurological symptoms caused by the mass of the tumour itself. Treatment is primarily medical (bromocriptine or cabergoline as first-line agents); but a surgical approach may be needed in large tumours that compress other brain structures and do not respond to the pharmacological approach.

There are many other causes of functional hyperprolactinaemia, including stress, medicines, pregnancy and compression of the pituitary stalk by other tumours. These (usually) mild elevations of PRL may be symptomatic or asymptomatic and do not usually require pharmacological treatment.

Does the test meet a medical need?

Determining serum PRL is the key test for diagnosing hyperprolactinaemia secondary to prolactinoma. When history and PRL levels are suggestive, magnetic resonance imaging (MRI) is needed to define the presence of a lesion compatible with a pituitary tumour (*1*, *2*). A prolactinoma diagnosis cannot be made without measuring PRL levels, which makes the test necessary for specific pharmacological treatment.

Assay-specific normal values are higher in women than men and are generally lower than 25 μg/L. In patients with pituitary adenomas, including prolactinomas, lengthy active disease is associated with a greater risk of comorbidities and lower quality of life. Timely diagnosis and treatment is therefore recommended to prevent or at least limit deleterious effects of hormone excess. Compared with no treatment, appropriate treatment (pharmacological with dopamine agonists or surgical resection of the adenoma) can lead to disease remission, improved quality of life, decreased incidence and severity of comorbidities, and fewer deaths.

Regardless of cause, hyperprolactinaemia is frequently associated with hypogonadism and its consequences. Timely diagnosis and treatment have been

shown to reverse these symptoms and others. Normalizing PRL levels restores fertility in more than 80–90% of patients with prolactinomas, at a significantly lower cost compared with fertility induction treatments. Quality of life is impaired in patients with hyperprolactinaemia; indeed, studies have shown that it is inversely associated with PRL levels among these patients, reinforcing the importance of providing adequate disease control.

Treating prolactinomas aims to reverse clinical signs, decrease tumour size, restore gonadal function and other pituitary hormone deficiencies, and prevent tumour recurrence and progression. All these are usually achieved progressively, as the concentrations of serum PRL normalize. Measuring PRL can help during medical treatment adjustment. The Endocrine Society guidelines recommend periodic PRL measurement starting 1 month after therapy, to guide treatment intensification so as to achieve normal PRL levels and reverse hypogonadism.

How the test is used

PRL is measured in the serum in symptomatic patients through a single blood test. Dynamic testing has no demonstrated advantages. The first step in diagnosing and evaluating hyperprolactinaemia is a detailed history of clinical symptoms and possible causes, including medications, comorbidities and lifestyle factors. The physical examination evaluates signs of hypothyroidism, hypogonadism, renal failure and visual field defects. Blood samples should be collected under resting basal conditions without excessive venepuncture stress. A serum PRL result above the upper limit of normal after ruling out macroprolactin or other interferences (medicines, stress) is required to diagnose hyperprolactinaemia. Sometimes a repeat PRL test is needed if the initial prolactin result is borderline high, especially in children, or if the normal reference interval provided by the laboratories is not appropriate.

The use of PRL measurement to diagnose hyperprolactinaemia, however, is not simple (*3*, *4*).

First, elevated PRL levels can be caused by other conditions, which need to be ruled out before a diagnosis of hyperprolactinaemia can be considered. Although PRL levels above 200–250 µg/L suggest prolactinomas, they can also occasionally be found in people with macroprolactinaemia, drug-induced hyperprolactinaemia, chronic renal failure, primary hypothyroidism and pregnancy.

Second, macroprolactinaemia is a common finding that must be identified as it usually requires no treatment. Most macroprolactinaemic patients are asymptomatic, but many can present coincidental galactorrhoea or menstrual disorders, as well as neuroradiological abnormalities, due to the presence of other diseases. Macroprolactinaemia is the third most frequent cause of non-physiological hyperprolactinaemia after drugs and prolactinomas.

Third, the hook effect, which is an assay artifact caused by an extremely high PRL level, can result in falsely low PRL results. The hook effect should be considered in all cases of large (≥ 3 cm) pituitary adenomas associated with normal or mildly elevated PRL levels (≤ 250 μg/L). An overlooked hook effect may lead to incorrect diagnosis and unnecessary surgical intervention in patients with prolactinomas, and may be identified by repeating PRL measurement after a 1 : 100 serum sample dilution.

Fourth, patients with non-functioning pituitary adenomas typically have PRL levels of 100 μg/L, but there are frequent exceptions. Up to 25% of patients harbouring a microprolactinoma or a cystic macroprolactinoma may have PRL levels below 100 μg/L.

Serum PRL levels are key to follow-up in patients under medical treatment. They are used to titrate the dose of the dopamine agonist (*1*, *2*, *5*). There are no evidence-based recommendations for how often to measure PRL during treatment, but common practice is initially every 1–3 months, followed by every 3–6 months once PRL levels normalize, and yearly after treatment finishes (sooner if there are symptoms).

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Hyperprolactinaemia is one of the most frequently diagnosed clinical disorders in routine endocrine practice. It is prevalent in around 0.4% of the general adult population, but its prevalence increases substantially among people with reproductive diseases (*6*). Hyperprolactinaemia is noted in 15–20% of women with secondary amenorrhoea or oligomenorrhoea, in approximately 30% of those with galactorrhoea or infertility, and in 75% of those with both amenorrhoea and galactorrhoea. (*7*, *8*). Its prevalence was reported to be 5% in a family planning clinic and 17% among women with PCOS.

PRL-secreting tumours (prolactinomas) are the most frequently occurring pituitary tumours and account for up to 40% of pituitary adenomas in the clinical setting. They are divided into microadenomas (≤ 10 mm diameter) and macroprolactinomas (≥ 10 mm). They occur with an incidence of 6–10 cases per million population per year, which translates into a prevalence of approximately 60–100 cases per million (*1*, *8*). Recent research indicates, however, that the prevalence of all pituitary tumours, including prolactinomas, may be three to five times higher than once thought. A prevalence of 55 per 71 000 (775 per million) inhabitants was found in Belgium (*9*), with similar findings in other populations (*9*–*11*). The prevalence of pituitary adenomas has been estimated to be higher than 10% in imaging-based screening or autopsy studies, although most of these patients have small tumours and are asymptomatic (*12*).

In young adults, prolactinomas occur much more frequently in women than in men, with a female-to-male ratio of approximately 10 : 1 between the ages of 20 and 50 years (*8*). The highest incidence rate is found in women between 25 and 34 years of age: almost 24 per 100 000 person-years (*13*). It is possible that the diagnosis is more frequent in young adult women because of the sensitivity of menses to disruption by hyperprolactinaemia. The adenomas that occur in men are usually larger, in part due to the lack of symptoms or delay in seeking medical attention for symptoms such as erectile dysfunction; but adenomas in men may also have an inherently greater rate of growth. Over the age of 50 there are no differences across gender. Prolactinomas are rare in children and adolescents, but they still account for approximately half of all pituitary adenomas in that population (*8*, *13–15*).

Pregnancy, breastfeeding, stress, exercise and sleep can also elevate PRL levels in generally healthy people (*16*, *17*). Medication-related hyperprolactinaemia has been reported in up to 25–30% of patients treated with neuroleptics, neuroleptic-like drugs and antidepressants, and in 5% of patients taking H2-receptor antagonists. In such drug-induced cases, PRL elevation is usually mild but can be highly variable. Risperidone and metoclopramide can lead to PRL levels above 200 μg/L (*18*, *19*).

Hyperprolactinaemia inhibits the secretion of hypothalamic GnRH. Consequently, the most common symptoms of hyperprolactinaemia in premenopausal women are amenorrhoea and galactorrhoea, while symptoms in men include impotence and decreased libido. Hyperprolactinaemia is an important cause of infertility among both genders. Bone loss and vertebral fractures are common comorbidities of hyperprolactinaemia-mediated sex steroid attenuation (*20*). In particular, spinal bone density is decreased, although overt osteoporosis is rare. Metabolic consequences of untreated hyperprolactinaemia can be expected (*21*).

Even though there are few available data on mortality associated with hyperprolactinaemia, studies have suggested that high serum PRL levels might be associated with an increased risk of death (*22*). In a recent study, no increased mortality was observed in patients with pituitary microadenomas, but other subgroups – including pituitary macroadenomas, drug-induced and idiopathic hyperprolactinaemia – had an increased risk of death, with adjusted hazard ratios ranging between 2.8 and 3.7 (*23*).

Quality of life is significantly affected in patients with hyperprolactinaemia. Hyperprolactinaemia causes infertility by interfering with ovulation. Infertility has particularly severe consequences for men and women in LMICs, including stigma, detrimental psychological effects and loss of economic security (*24*). Patients with hyperprolactinaemia have lower scores in physical functioning, general health, social functioning, emotional aspect and mental health. Treatment

of prolactinoma, either with surgery or with bromocriptine and cabergoline, has been linked to a QALYs gain of approximately 20 years (*25*).

F. WHO or other clinical guidelines relevant to the test

In all guidelines, PRL levels are considered to be normal if less than 25 μg/L, with an association between tumour size and PRL levels. PRL levels above 250 μg/L usually indicate the presence of a prolactinoma, although some drugs, including risperidone and metoclopramide, may cause PRL elevations above 200 μg/L. If there is hyperprolactinaemia discordant with the clinical picture, macroprolactin measurement is recommended.

The Endocrine Society clinical guidelines (*26*) recommend dopamine agonist therapy to lower PRL levels, decrease tumour size and restore gonadal function for patients harbouring symptomatic PRL-secreting microadenomas or macroadenomas.

G. Basic test characteristics (from application)

Test formats available	Immunoassays of varying formats, for example electrochemiluminescence and chemiluminescence
Specimen types	Serum, plasma
Equipment required	Automated analyser
Regulatory status	CE-marked and FDA-approved; also approved by other regulatory agencies, including Health Canada
Availability	Global
Price per test range	Reimbursement of US$ 10–21 (in North America)
Instrument price range	US$ 40 000–200 000 depending on instrument size (in North America).

H. Evidence for diagnostic accuracy (from application)

No systematic review of the accuracy of the PRL test in clinical studies could be found.

Saleem et al. (*26*) look at the issues relating to the laboratory measurement of PRL and find considerable variability in routinely available PRL immunoassays as a result of differing reactivity towards monomeric PRL and macroprolactin and lack of commutability of the WHO third International Standard between routine methods. Macroprolactinaemia is a relatively common cause of interference in the PRL assay that may lead to incorrect diagnosis and unnecessary investigations. Measuring PRL after precipitation by polyethylene glycol when PRL levels

are above the reference interval is routinely used to identify macroprolactin, although harmonization of the precipitation process and reporting may improve clinical care.

As the differential diagnosis for hyperprolactinaemia is very broad, other causes must be ruled out before considering a diagnosis of prolactinoma (*27*). Kawaguchi et al. (*28*) did a retrospective chart review to differentiate prolactinomas from non-functioning pituitary tumours and found the area under the ROC to be 0.96, with 99% sensitivity and 81% specificity at a cut-off PRL value of 38.6 μg/L. But this PRL level also included 18.9% of patients with non-functioning pituitary adenomas. The study has some limitations, including the fact that the authors did not identify macroprolactin. This showed that PRL alone is not definitive for diagnosing prolactinoma. Neuroimaging may help reduce false-positive diagnoses. Some studies even propose using PRL/adenoma maximum diameter and PRL/adenoma volume to distinguish prolactinoma from other types of pituitary adenomas (*29*).

I. Evidence for clinical usefulness and impact (from application)

Measurement of serum PRL is required to diagnose and manage hyperprolactinaemia.

In 2012, Wang et al. (*30*) did a systematic review and meta-analysis of outcomes of hyperprolactinaemic patients, including micro- and macroprolactinomas, to provide evidence-based recommendations for practitioners. The authors aimed to compare efficacy and adverse effects of medications, surgery and radiotherapy in the treatment of hyperprolactinaemia. The review included eight randomized and 178 non-randomized studies (more than 3000 patients). Compared with no treatment, dopamine agonists significantly reduced PRL levels (weighted mean difference: –45; 95% CI: –77 to –11) and the likelihood of persistent hyperprolactinaemia. Cabergoline was more effective than bromocriptine in reducing persistent hyperprolactinaemia, amenorrhoea/oligomenorrhoea and galactorrhoea. A large body of non-comparative literature showed dopamine agonists improved other patient-important outcomes. Low-to-moderate quality evidence was found supporting improved outcomes with surgery and radiotherapy compared with no treatment in patients who were resistant to or intolerant of dopamine agonists. Their results provide evidence to support the use of dopamine agonists in reducing PRL levels and persistent hyperprolactinaemia, with cabergoline proving to be more efficacious than bromocriptine.

Another systematic review and meta-analysis of randomized controlled trials by dos Santos Nunes et al. (*31*) compared cabergoline with bromocriptine for treating patients with idiopathic hyperprolactinaemia and prolactinomas and found similar results. The meta-analysis of normalization of serum PRL levels

and menstruation with return of ovulatory cycle showed a significant difference in favour of cabergoline, with a risk ratio (RR) of 0.67 (95% CI: 0.57–0.80) and 0.74 (95% CI: 0.67–0.83), respectively. The number of adverse effects was significantly higher in the bromocriptine group (RR 1.43 [95% CI: 1.03–1.98]). This showed new evidence favouring the use of cabergoline compared with bromocriptine for treating prolactinomas and idiopathic hyperprolactinaemia. Further support for the use of cabergoline is provided by a systematic review on the treatment of giant prolactinomas by Huang et al. (*32*).

As for the impact of treatment on other comorbidities, a recent meta-analysis by D'Sylva et al. (*33*) showed that fracture prevalence was increased in patients with untreated hyperprolactinaemia compared with those on treatment, independent of gonadal function; this preliminary evidence would support treatment in postmenopausal women with no other symptoms.

There are several studies showing the clinical impact of treating hyperprolactinaemia; 10 have been chosen based on clinical relevance and publication date.

Various studies show that dopamine agonists normalize PRL levels and reduce the size of prolactinomas in a significant proportion of patients. For example, Berinder et al. (*34*) found that among women with hyperprolactinaemia, treatment with dopamine agonists (including bromocriptine, cabergoline and quinagolide) normalized PRL levels in 71% of patients, with 80% exhibiting total or partial tumour shrinkage. Similarly, Colao et al. (*35*) did a prospective study of patients with macroprolactinomas and found normal PRL levels were achieved within 6 months in 81% of patients receiving cabergoline, with 92% exhibiting significant tumour shrinkage. This effect has also been shown in patients with idiopathic hyperprolactinaemia: in Verhelst et al.'s (*36*) retrospective study, cabergoline normalized PRL levels in 92% of patients with idiopathic hyperprolactinaemia or a microprolactinoma and in 77% of patients with macroadenomas. Pinzone et al. (*37*) found that 80% of men with prolactinomas had normalized PRL after treatment with other dopamine agonists. And Ono et al. (*38*) showed that individualized high-dose cabergoline treatment can normalize hyperprolactinaemia and hypogonadism, irrespective of tumour size or preceding treatments, even for large tumours previously thought to have poor response (*38*).

Ezzat et al. (*12*) found that hyperprolactinaemia treatment induces restoration of menses. Ono et al. (*39*) also found that high-dose cabergoline in infertile women with prolactinoma showed high rates of pregnancy with uneventful outcomes. And de Rosa et al. (*40*, *41*) did studies in men which showed that treatment with cabergoline restored erectile function and sperm count and motility.

As for other comorbidities related to hyperprolactinaemia, recent studies by Auriemma et al. (*42*) and Berinder et al. (*43*) associate the treatment of prolactinoma with reduced prevalence of metabolic syndrome and improved metabolic profile.

J. Evidence for economic impact and/or cost–effectiveness (from application)

It is important to ensure proper use of the PRL test. For instance, health professionals should know that using antipsychotic medications causes hyperprolactinaemia. Inappropriate detection of hyperprolactinaemia in these patients has been associated with substantial health care costs. A retrospective study of 1000 patients receiving antipsychotic medications in the USA found that compared with the hyperprolactinaemia-free cohort, health care costs increased by nearly US$ 6000 in the hyperprolactinaemia group (*44*). There are no studies on health care costs of this condition in low-resource settings, but these data emphasize the importance of properly educating health care professionals and the risk of overusing this test.

K. Ethics, equity and human rights issues (from application)

The availability of the PRL test will enable the diagnosis of hyperprolactinaemia due to prolactinoma and other causes. Pharmacologic treatment of hyperprolactinaemia has proven to be clearly beneficial in terms of several patient outcomes. Their use controls hormonal secretion and tumour growth in approximately 80% of cases and also leads to resolution of symptoms, including infertility, amenorrhoea, sexual dysfunction and galactorrhoea.

L. Summary of evidence evaluation

The main goal of the test for prolactin is to diagnose hyperprolactinaemia (including caused by a prolactinoma) and to follow up patients treated for hyperprolactinaemia (including prolactinoma).

Guidelines suggest that prolactin measurement using immunoassays for detecting hyperprolactinaemia is appropriate. Both guidelines and reviews repeatedly emphasize that prolactin measurement alone is insufficient to further subclassify the etiologic cause of hyperprolactinaemia (e.g. pituitary tumour, drugs, idiopathic). For that, other tests such as anamnesis, biomarker measurement (thyroid-stimulating hormone (TSH), free T4 and creatinine) and MRI scans are needed.

The claims made on the diagnostic accuracy of PRL testing for hyperprolactinaemia are not made in terms of traditional diagnostic test accuracy measures, although they do suggest that it is adequate. The submission only lists one study with diagnostic test accuracy of PRL for etiologic causes of hyperprolactinaemia.

Claims made of impact on health outcomes or costs assume that a root cause of hyperprolactinaemia is identified, and are circumstantial. No evidence-based conclusion can be drawn for the effect of PRL testing on health outcomes or costs.

M. Summary of SAGE IVD deliberations

Prolactinoma is a pituitary tumour that responds to medical treatment and, if necessary, to surgical treatment. PRL testing is a simple test that is commonly used in clinical practice to diagnose and follow up on prolactinoma and conditions associated with hyperprolactinaemia, such as infertility and menstrual irregularities. While the EML does not include any medicines for managing hyperprolactinaemia, PRL testing for it is included in the guidelines of many professional societies. There is limited evidence regarding the test's impact on health outcomes or cost.

SAGE IVD raised concerns about the potential to overuse PRL testing as well as the potential for abnormal results that are not associated with hyperprolactinaemia. Like cortisol testing, PRL testing requires appropriate laboratory infrastructure and quality assurance. And, depending on the type of assay used, PRL testing can also be affected by biotin ingested by the patient.

Importantly, SAGE IVD emphasized the need to interpret PRL results with care, not least because macroprolactin, which is present in 3.7% of the general population, causes high levels of PRL but is inactive. The group noted that laboratory procedures such as precipitation by polyethylene glycol are required to estimate the amount of biological active monomeric PRL.

N. SAGE IVD recommendations

SAGE IVD recommended including the prolactin test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, within a new subsection for endocrine disorders);
- using an immunoassay format;
- to diagnose and monitor hyperprolactinaemia (including prolactinoma).

The group further requested the addition of a note to the test category entry in the EDL stating that it is only recommended for use in specialized health care settings.

O. References

1. Melmed S, Casanueva FF, Hoffman AR, Kleinberg DL, Montori VM, et al. Diagnosis and treatment of hyperprolactinemia: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2011;96(2):273–288. doi:10.1210/jc.2010-1692.

2. Casanueva FF, Molitch ME, Schlechte JA, Abs R, Bonert V, et al. Guidelines of the Pituitary Society for the diagnosis and management of prolactinomas. Clin Endocrinol. 2006;65(2):265–273. doi:10.1111/j.1365-2265.2006.02562.x.
3. Vilar L, Vilar CF, Lyra R, Freitas MDC. Pitfalls in the diagnostic evaluation of hyperprolactinemia. Neuroendocrinology. 2019;109(1):7–19. doi:10.1159/000499694.
4. Vilar L, Fleseriu M, Bronstein MD. Challenges and pitfalls in the diagnosis of hyperprolactinemia. Arq Bras Endocrinol Metabol. 2014;58(1):9–22. doi:10.1590/0004-2730000003002.
5. Snyder PJ. Management of hyperprolactinemia. In: UpToDate [website]. Waltham: UpToDate Inc; 2019 (https://www.uptodate.com/contents/management-of-hyperprolactinemia?topicRef=6638&source=see_link, accessed 30 March 2020).
6. Biller BM, Luciano A, Crosignani PG, Molitch M, Olive D, et al. Guidelines for the diagnosis and treatment of hyperprolactinemia. J Reprod Med. 1999;44(12):1075–1084.
7. Mancini T, Casanueva FF, Giustina A. Hyperprolactinemia and prolactinomas. Endocrinol Metab Clin North Am. 2008;37:67–99. doi:10.1016/j.ecl.2007.10.013.
8. Ciccarelli A, Daly AF, Beckers A. The epidemiology of prolactinomas. Pituitary. 2005;8(1):3–6. doi:10.1007/s11102-005-5079-0.
9. Beckers A, Adam C, Ciccarelli A, Rixhon M, Daly A, et al. Approaching the true prevalence of pituitary tumors. In: ENEA Congress 2004; Abstract O1.1:23 [cited 2019 Sep 23].
10. Fernandez A, Karavitaki N, Wass JA. Prevalence of pituitary adenomas: a community-based, cross-sectional study in Banbury (Oxfordshire, UK). Clin Endocrinol. 2010;72(3):377. doi:10.1111/j.1365-2265.2009.03667.x.
11. Daly AF, Rixhon M, Adam C, Dempegioti A, Tichomirowa MA, et al. High prevalence of pituitary adenomas: a cross-sectional study in the province of Liege, Belgium. J Clin Endocrinol Metab. 2006;91(12):4769–4775. doi:10.1210/jc.2006-1668.
12. Ezzat S, Asa SL, Couldwell WT, Barr CE, Dodge WE, et al. The prevalence of pituitary adenomas: a systematic review. Cancer. 2004;101(3):613–619. doi:10.1002/cncr.20412.
13. Davis JR, Farrell WE, Clayton RN. Pituitary tumours. Reproduction. 2001;121(3):363–371. doi:10.1530/rep.0.1210363.
14. Mindermann T, Wilson CB. Age-related and gender-related occurrence of pituitary adenomas. Clin Endocrinol. 1994;41(3):359–364. doi:10.1111/j.1365-2265.1994.tb02557.x.
15. Fideleff HL, Boquete HR, Suárez MG, Azaretzky M. Prolactinoma in children and adolescents. Horm Res. 2009;72(4):197–205. doi:10.1159/000236081.
16. Vance ML, Thorner MO. Prolactin: hyperprolactinemic syndromes and management. In: DeGroot LJ, editor. Textbook of Endocrinology, 2nd edition. Philadelphia: WB Saunders Company; 1989:408–418.
17. Arafah BM, Nasrallah MP. Pituitary tumors: pathophysiology, clinical manifestations and management. Endocr Relat Cancer. 2001;8(4):287–305. doi:10.1677/erc.0.0080287.
18. Petit A, Piednoir D, Germain ML, Trenque T. Drug-induced hyperprolactinemia: a case–non-case study from the national pharmacovigilance database. Therapie. 2003;58(2):159–163. doi:10.2515/therapie:2003023.
19. Molitch ME. Drugs and prolactin. Pituitary. 2008;11(2):209–218. doi:10.1007/s11102-008-0106-6.
20. Mazziotti G, Mancini T, Mormando M, De Menis E, Bianchi A, et al. High prevalence of radiological vertebral fractures in women with prolactin-secreting pituitary adenomas. Pituitary. 2011;14(4):299–306. doi:10.1007/s11102-011-0293-4.

21. Bernabeu I, Casanueva FF. Metabolic syndrome associated with hyperprolactinemia: a new indication for dopamine agonist treatment? Endocrine. 2013;44(2):273–274. doi:10.1007/s12020-013-9914-1.
22. Haring R, Friedrich N, Volzke H, Vasan RS, Felix SB, et al. Positive association of serum prolactin concentrations with all-cause and cardiovascular mortality. Eur Heart J. 2014;35(18):1215–1221. doi:10.1093/eurheartj/ehs233.
23. Soto-Pedre E, Newey PJ, Bevan JS, Leese GP. Morbidity and mortality in patients with hyperprolactinemia: the PROLEARS study. Endocr Connect. 2017;6(8):580–588. doi:10.1530/EC-17-0171.
24. Rouchou B. Consequences of infertility in developing countries. Perspect Public Health. 2013;133(3):174–179. doi:10.1177/1757913912472415.
25. Zygourakis CC, Imber BS, Chen R, Han SJ, Blevins L, et al. Cost-effectiveness analysis of surgical versus medical treatment of prolactinomas. J Neurol Surg B Skull Base. 2017;78(2):125–131. doi:10.1055/s-0036-1592193.
26. Saleem M, Martin H, Coates P. Prolactin biology and laboratory measurement: an update on physiology and current analytical issues. Clin Biochem Rev. 2018;39(1):3–16.
27. Vilar L, Freitas MC, Naves LA, Casulari LA, Azevedo M, et al. Diagnosis and management of hyperprolactinemia: results of a Brazilian multicenter study with 1234 patients. J Endocrinol Invest. 2008;31:436–444. doi:10.1007/bf03346388.
28. Kawaguchi T, Ogawa Y, Tominaga T. Diagnostic pitfalls of hyperprolactinemia: the importance of sequential pituitary imaging. BMC Res Notes. 2014;7:555. doi:10.1186/1756-0500-7-555.
29. Huang Y, Ding C, Zhang F, Xiao D, Zhao L, et al. Role of prolactin/adenoma maximum diameter and prolactin/adenoma volume in the differential diagnosis of prolactinomas and other types of pituitary adenomas. Oncol Lett. 2018;15(2):2010–2016. doi:10.3892/ol.2017.7462.
30. Wang AT, Mullan RJ, Lane MA, Hazem A, Prasad C, et al. Treatment of hyperprolactinemia: a systematic review and meta-analysis. Syst Rev. 2012;1:33. doi:10.1186/2046-4053-1-33.
31. dos Santos Nunes V, El Dib R, Boguszewski CL, Nogueira CR. Cabergoline versus bromocriptine in the treatment of hyperprolactinemia: a systematic review of randomized controlled trials and meta-analysis. Pituitary. 2011;14:259–265. doi:10.1007/s11102-010-0290-z.
32. Huang HY, Lin SJ, Zhao WG, Wu ZB. Cabergoline versus bromocriptine for the treatment of giant prolactinomas: a quantitative and systematic review. Metab Brain Dis. 2018;33(3):969–976. doi:10.1007/s11011-018-0217-3.
33. D'Sylva C, Khan T, Van Uum S, Fraser LA. Osteoporotic fractures in patients with untreated hyperprolactinemia vs. those taking dopamine agonists: a systematic review and meta-analysis. Neuro Endocrinol Lett. 2015;36(8):745–749.
34. Berinder K, Stackenäs I, Akre O, Lindén Hirschberg A, Hulting AL. Hyperprolactinaemia in 271 women: up to three decades of clinical follow-up. Clin Endocrinol. 2005;63(4):450–455. doi:10.1111/j.1365-2265.2005.02364.x.
35. Colao A, Di Sarno A, Landi ML, Scavuzzo F, Cappabianca P, et al. Macroprolactinoma shrinkage during cabergoline treatment is greater in naive patients than in patients pretreated with other dopamine agonists: a prospective study in 110 patients. J Clin Endocrinol Metab. 2000;85(6):2247–2252. doi:10.1210/jcem.85.6.6657.
36. Verhelst J, Abs R, Maiter D, van den Bruel A, Vandeweghe M, et al. Cabergoline in the treatment of hyperprolactinemia: a study in 455 patients. J Clin Endocrinol Metab. 1999;84(7):2518–2522. doi:10.1210/jcem.84.7.5810.

37. Pinzone JJ, Katznelson L, Danila DC, Pauler DK, Miller CS, et al. Primary medical therapy of micro- and macroprolactinomas in men. J Clin Endocrinol Metab. 2000;85(9):3053–3057. doi:10.1210/jcem.85.9.6798.

38. Ono M, Miki N, Kawamata T, Makino R, Amano K, et al. Prospective study of high-dose cabergoline treatment of prolactinomas in 150 patients. J Clin Endocrinol Metab. 2008;93(12):4721–4727. doi:10.1210/jc.2007-2758.

39. Ono M, Miki N, Amano K, Kawamata T, Seki T, et al. Individualized high-dose cabergoline therapy for hyperprolactinemic infertility in women with micro- and macroprolactinomas. J Clin Endocrinol Metab. 2010;95(6):2672–2679. doi:10.1210/jc.2009-2605.

40. de Rosa M, Zarrilli S, Vitale G, Di Somma C, Orio F, et al. Six months of treatment with cabergoline restores sexual potency in hyperprolactinemic males: an open longitudinal study monitoring nocturnal penile tumescence. J Clin Endocrinol Metab. 2004;89(2):621–625. doi:10.1210/jc.2003-030852.

41. de Rosa M, Ciccarelli A, Zarrilli S, Guerra E, Gaccione M, et al. The treatment with cabergoline for 24 months normalizes the quality of seminal fluid in hyperprolactinaemic males. Clin Endocrinol. 2006;64(3):307–313. doi:10.1111/j.1365-2265.2006.02461.x.

42. Auriemma RS, Granieri L, Galdiero M, Simeoli C, Perone Y, et al. Effect of cabergoline on metabolism in prolactinomas. Neuroendocrinology. 2013;98(4):299–310. doi:10.1159/000357810.

43. Berinder K, Nyström T, Höybye C, Hall K, Hulting AL. Insulin sensitivity and lipid profile in prolactinoma patients before and after normalization of prolactin by dopamine agonist therapy. Pituitary. 2011;14(3):199–207. doi:10.1007/s11102-010-0277-9.

44. Cloutier M, Greene M, Touya M, Gagnon-Sanschagrin P, Guerin A. A real-world analysis of healthcare costs and relative risk of hyperprolactinemia associated with antipsychotic treatments in the United States. J Med Econ. 2018;21(12):1183–1190. doi:10.1080/13693998.2018.1521415.

NEGLECTED TROPICAL DISEASES

3.17. *Trypanosoma* immunoglobulin G antibody

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a *T. cruzi* IgG antibody detection test to the EDL as an IVD for diagnosing chronic Chagas disease, screening pregnant women for Chagas disease, screening blood and organ donors, surveilling Chagas disease and to aid in the diagnosis of chagasic cardiomyopathy.

B. Applicant

Roche Diagnostics

C. WHO technical department

Control of Neglected Tropical Diseases

D. Background (from application)

Disease condition and impact on patients

Chagas disease, or American trypanosomiasis, is a parasitic, vector-borne disease caused by *T. cruzi*. In addition to vector transmission it can also be transmitted through food, congenital infection, blood transfusion, organ transplants and in the laboratory. More than 70 million people are estimated to be at risk of contracting Chagas disease, with around 38 000 new cases and 12 000 deaths each year. The disease affects around 8–10 million people in the Americas (*1*), where it is endemic. Vector control programmes have successfully reduced vector transmission of *T. cruzi* in these regions (*2*); but the disease has spread to other countries through migration and international travel and is now also present in Europe and the western Pacific region (*3*).

An estimated 15% of patients with chronic Chagas disease develop severe clinical manifestations and related complications, such as vascular accidents, intestinal complications and death (mainly due to arrhythmias and heart failure) (*4*).

WHO estimates that 1 125 000 women of fertile age are infected with *T. cruzi* in Latin America, with 8668 congenital infections each year. A further 40 000 infected women of childbearing age are estimated to live in the USA, where 60–315 congenital infections are expected to occur each year. In Latin America, the mean maternal–fetal transmission rate in chronic Chagas disease is estimated to be 4.7%, with the number of babies that would require testing each year an estimated 158 000–214 000 (mostly in Argentina, Brazil, Bolivia and Mexico).

Does the test meet a medical need?

WHO has recognized Chagas disease as a neglected tropical disease (NTD) since 2015. Early diagnosis and treatment are known to prevent congenital transmission and reduce the likelihood of disease progression. But Chagas disease remains underdiagnosed (*5*).

How the test is used

Diagnosis is usually made by serology (two positive serological tests), peripheral blood microscopy or PCR. A positive serology is indicative of either active *T. cruzi* infection or past exposure. Serologically positive asymptomatic patients can transmit the parasite to the vector insect and directly to other individuals via blood components, organ donation or through congenital infection.

The Chagas serological assays to be used for screening and diagnosis must be both highly sensitive and specific. This can be achieved by using a combination of antigens to develop tests, which maintains the high sensitivity without compromising the specificity. An adequate performance of such tests can

be reached by using recombinant antigens that cover all stages of the parasite life cycle, for example, flagellar calcium-binding protein, flagellar repetitive antigen and cruzipain for the determination of antibodies to *T. cruzi* (*6*).

E. Public health relevance (from application)

Prevalence and socioeconomic impact

The global annual burden of Chagas disease is calculated at more than US$ 7 billion per year, considering health care costs and DALYs from infected individuals. The acute phase of Chagas infection lasts 2 months; without curative treatment, all patients enter a chronic phase (either as carriers or as symptomatic individuals).

F. WHO or other clinical guidelines relevant to the test

Pan American Health Organization (PAHO) guidelines for diagnosing and treating Chagas disease (*2*) provide recommendations on the use of various technologies for different test purposes. They make a strong recommendation for using ELISA and immunochromatographic test (ICT) for population studies on prevalence of Chagas disease. They also make a conditional recommendation on using ELISA, haemagglutination inhibition assay (HIA) or indirect immunofluorescence (IIF) test for diagnosing patients with suspected Chagas disease, in an algorithm with two initial tests with differing antigens, followed by a third test in case of conflicting results.

In 2010, WHO published a report evaluating 19 anti-*T. cruzi* assays, including 11 EIAs, seven agglutination assays and one rapid test (*7*). The evaluations focused on operational characteristics such as sensitivity, specificity on a WHO serum/plasma evaluation panel, ease of use and suitability to small laboratories with limited facilities. The information is intended to support programme managers and users to decide which tests are best suited to their particular situation.

G. Basic test characteristics (from application)

Test formats available	**ELISA, HIA, IIF, chemiluminescent microparticle immunoassay (CMIA), RDT**
Specimen types	Serum, plasma
Equipment required	ELISA reader
Regulatory status	Some CE-marked
Availability	Global

Table *continued*

Price per test range	Not provided
Instrument price range	Not provided

H. Evidence for diagnostic accuracy (from application)

Various formats of IgG immunoassays have been evaluated for the diagnosis of chronic Chagas. Sensitivity and specificity vary greatly based on the antigens used to develop IgG assays.

Traditional HIAs, IIFs and ELISAs prepared with whole parasite lysates offer quantitative or semi-quantitative results. These, however, may cross-react with *Leishmania* spp. and *Trypanosoma rangeli*. IHA tests have sensitivities ranging from 96% to 98%, while IIF has a greater sensitivity of 99%, but is operationally more demanding. ELISAs have good sensitivity and specificity (*8*).

Newer-generation ELISAs, CMIAs and RDTs (i.e. ICTs and ELISAs with POC applicability) offer qualitative results. They have improved specificity but may have lower sensitivity. In a study from Colombia, Díaz et al. (*9*) compared commercially available whole-parasite lysate and recombinant antigen ELISA assays and found 99–100% sensitivity and 98–100% specificity.

WHO guidelines (*2*) evaluated the accuracy of ELISA, CMIA and ICT tests, and noted no substantial differences in terms of sensitivity.

I. Evidence for clinical usefulness and impact (from application)

Given the lack of evidence linking diagnostics to patient outcomes, clinical utility has been extrapolated from diagnostic accuracy studies (undetected, therefore untreated patients vs effect of trypanocidal treatment) (*2*).

In different evaluations, the diagnostic sensitivities of the various methods did not vary substantially. When used as a single test without secondary confirmation, all methods showed some incorrect classification of patients (false negatives or false positives) (*2*).

This is why PAHO guidelines recommend using two serological assays with different technologies for initial diagnosis. In resource-limited settings where a single ELISA test is used for initial diagnosis, PAHO recommends confirming a positive result with an additional test (*2*).

J. Evidence for economic impact and/or cost–effectiveness (from application)

Bartsch et al. (*10*) evaluated the economic impact and outcomes of identifying and treating different proportions of patients in the acute and indeterminate disease states in a 2000-person village in Yucatán, Mexico. The authors concluded

that managing as few as 5% of Chagas cases annually (acute and indeterminate stages) reduces transmission and provides economic and health benefits.

The PAHO guidelines (*2*) conclude that the ELISA method is more resource effective than ICT and CMIA. But the diagnostic accuracy of ELISA commercial assays varies significantly.

K. Ethics, equity and human rights issues (from application)

None identified.

Testing reduces inequity by providing diagnosis. Easier-to-use methods such as ELISA or ICT are more likely to have a positive impact on equity.

L. Summary of evidence evaluation

Several recent systematic reviews have identified and pooled studies of the accuracy of immunologically based tests which show that they have high levels of performance, suitable for use across different settings. Estimates of sensitivity for ELISA, ICT and CMIA are 97% (CI: 96–98%), 94% (CI: 91–96%) and 99% (CI: 97–100%), respectively. Estimates of specificity for ELISA, ICT and CMIA are 98% (CI: 97–99%), 97% (CI: 96–98%) and 98% (CI: 91–99%), respectively.

PAHO guidelines have carefully considered the consequences of test errors in different circumstances and make clear recommendations as to which tests are preferable in different settings.

M. Summary of SAGE IVD deliberations

The *T. cruzi* IgG test has the potential to significantly increase access to diagnosis and, consequently, treatment in endemic countries. Automated tests have improved performance compared with other formats, but their complexity and costs could limit their use in countries where Chagas disease is endemic. In these cases, RDTs can be used. SAGE IVD members highlighted the debate surrounding the need for one or two positive tests to diagnose *T. cruzi* infection; the group also noted that while current WHO/PAHO guidelines do not recommend using RDTs as stand-alone tests, new evidence indicates their potential use as diagnostic tools in endemic countries. Further evidence is being generated to confirm this finding. SAGE IVD sought advice from WHO's Neglected Tropical Diseases team on the potential use of the *T. cruzi* test in primary care settings.

The submission did not include enough details about the different *T. cruzi* IgG detection assays available. Many different immunoassay formats exist (e.g. ELISA, ELISA-r, immunofluorescence assay [IFA], IHA, CMIA and ICT/RDT) which are used for different purposes, mainly chronic Chagas disease diagnosis and *T. cruzi* infection surveillance. The diagnostic algorithms and immunoassays used vary with the different purposes, populations targeted and regions.

N. SAGE IVD recommendations

SAGE IVD recommended including the *Trypanosoma cruzi* IgG antibody test category in the third EDL 3:

- as a disease-specific IVD for use in community settings and health facilities without laboratories, and as a disease-specific IVD for use in clinical laboratories (EDL 3, Sections I.b and II.b, Neglected tropical diseases);
- using an immunoassay format or an RDT (only in settings where laboratory-based methods are not available);
- for surveillance of *Trypanosoma cruzi* infection;
- to screen girls, women of childbearing age and pregnant women without previous treatment for *Trypanosoma cruzi* infection;
- to screen children and other at-risk populations; and
- to aid in the diagnosis (RDT) or to diagnose (immunoassay) chronic *Trypanosoma cruzi* infection (Chagas disease).

The group noted that an immunoassay to screen for *T. cruzi* in blood donations is already listed in the EDL 2, as a disease-specific IVD for blood screening laboratories (Section II.c); but highlighted the potential value of this test category for screening organ donors too. It recommended reviewing Section II.c of the EDL to include donated organs as well as blood.

O. References

1. Rassi A Jr, Rassi A, Marin-Neto JA. Chagas disease. Lancet. 2010;375(9723):1388–1402. doi:10.1016/S0140-6736(10)60061-X.
2. Guidelines for the diagnosis and treatment of Chagas disease. Washington (DC): Pan American Health Organization; 2019.
3. Gascon J, Bern C, Pinazo MJ. Chagas disease in Spain, the United States and other non-endemic countries. Acta Trop. 2010;115(1–2):22–27. doi:10.1016/j.actatropica.2009.07.019.
4. Lee BY, Bacon KM, Wateska AR, Bottazzi ME, Dumonteil E, Hotez PJ. Modeling the economic value of a Chagas' disease therapeutic vaccine. Hum Vaccin Immunother. 2012;8(9):1293–1301. doi:10.4161/hv.20966.
5. Basile L, Jansa JM, Carlier Y, Salamanca DD, Angheben A, et al. Chagas disease in European countries: the challenge of a surveillance system. Euro Surveill. 2011;16(37). doi:10.2807/ese.16.37.19968.en.
6. Balouz V, Agüero F, Buscaglia CA. Chagas disease diagnostic applications: present knowledge and future steps. Adv Parasitol. 2017;97:1–45. doi:10.1016/bs.apar.2016.10.001.
7. Anti-trypanosoma cruzi assays: operational characteristics. Geneva: World Health Organization; 2010.
8. Control of Chagas disease. In: WHO Expert Committee on the Control of Chagas Disease: second report. Geneva: World Health Organization; 2002 (WHO Technical Report Series, No. 905).

9. Díaz RA, Forsyth C, Bernal OA, Marchiol A, Duran MB, et al. Comparative evaluation of immunoassays to improve access to diagnosis for Chagas disease in Colombia. Int J Infect Dis. 2019;87:100–108. doi:10.1016/j.ijid.2019.07.022.

10. Bartsch SM, Avelis CM, Asti L, Hertenstein DL, Ndeffo-Mbah M, et al. The economic value of identifying and treating Chagas disease patients earlier and the impact on Trypanosoma cruzi transmission. PLoS Negl Trop Dis. 2018;12(11):e0006809. doi:10.1371/journal.pntd.0006809.

3.18. Visceral leishmaniasis direct agglutination test

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a visceral leishmaniasis (VL) direct agglutination test (DAT) to the EDL as an IVD for diagnosing VL.

B. Applicant

World Health Organization

C. WHO technical department

Control of Neglected Tropical Diseases (Leishmaniasis Control Programme)

D. Background (from application)

Note: This submission was made in response to a recommendation by SAGE IVD during its second annual meeting in March 2019.

Disease condition and impact on patients

The leishmaniases are a group of diseases caused by protozoan parasites from more than 20 Leishmania species that are transmitted by the bite of infected female phlebotomine sandflies (*1*). The disease affects some of the poorest people in the world, and is associated with malnutrition, population displacement, poor housing, a weak immune system and lack of financial resources.

There are three main forms of leishmaniases: visceral, also known as kala-azar and the most serious form of the disease; cutaneous (CL), the most common; and mucocutaneous, the most destructive. VL is a life-threatening systemic disease and is considered the second-largest parasitic killer disease after malaria. It is almost always fatal in more than 95% of cases without adequate and timely treatment. The disease occurs in poverty-stricken remote areas and affects the poorest people, and is associated with malnutrition, population

displacement, poor housing conditions and lack of financial resources. The typical *L. donovani* complex is the causative agent of VL.

The disease is characterized by prolonged fever, weight loss, enlargement of the spleen, liver and lymph nodes, anaemia and low blood cell count. In children, other symptoms include diarrhoea, cough, abdominal distension and growth retardation. If the disease remains untreated, it progresses to debilitation, bleeding and secondary infections, resulting in death.

The mode of transmission is anthroponotic and zoonotic. Human-to-human transmission is predominant in astern Africa and the Indian subcontinent; zoonotic transmission (through dogs) is found in the Mediterranean region and the Americas (*2*).

In endemic areas, a significant proportion of healthy people are exposed to infection but remain asymptomatic. However, the precise ratio between clinical cases and asymptomatic individuals varies from country to country: 1 : 5 in Kenya (*3*), 1 : 8 in Brazil (*4*), 1 : 5.6 in Ethiopia (*5*), 1 : 8.9 in India and Nepal (*6*), and 1 : 11 in Sudan (*7*).

Does the test meet a medical need?

The signs and symptoms of VL are non-specific, sharing clinical presentation with various other tropical diseases such as malaria, TB and enteric fever. This makes it mandatory to follow a diagnostic algorithm in a clinically suspected case of VL.

The gold standard for diagnosis is microscopic detection of the parasite in specimens. The most sensitive of these techniques is the splenic puncture followed by bone marrow and lymph node aspirates with decreasing order of sensitivity, respectively. These are invasive methods that require expertise and equipment and are only available at the referral level. Splenic aspirate carries the risk of fatal bleeding as a complication of the technique. Kosack et al. 2017 (*8*) suggest that the ideal diagnostic test for VL-endemic settings should be available at the level where they are needed most and be accurate, sensitive, specific, user-friendly, robust and equipment-free (i.e. it should meet the ASSURE criteria). The development of diagnostic tests to improve VL case management has been rated as one of the most pressing needs to tackle infectious diseases in LMICs.

It is important to diagnose VL correctly not only because untreated disease can be fatal but also because treatments are limited, expensive and have toxic side effects. Diagnostic tests should not give false-negative results that might result in death; nor should they give false-positive results that may cause people without VL to receive toxic treatments (*9*).

The invasive methods of diagnosis and associated complications have driven development of non-invasive serological tests like the DAT (*10–12*).

How the test is used

The DAT is a semi-quantitative test that uses microplates in which increasing dilutions of a patient's serum or blood are mixed with stained killed promastigotes of *L. donovani*. The DAT is a freeze-dried suspension of trypsin-treated fixed and stained culture of *L. donovani* promastigotes. If antibodies against the parasite are present in the sample, agglutination is visible with the naked eye.

Although the test requires minimal laboratory set-up, skilled laboratory personnel and a moderate level of training are needed to perform the test and interpret its results (*13*). During infection with VL, circulating antibodies are produced against the surface antigens of the invading parasites. The DAT detects antibodies to *L. donovani* in the blood or serum of those infected by means of direct agglutination. In the absence of antibodies to *Leishmania*, the DAT antigen accumulates at the bottom of the plate to form a dark blue spot. If antibodies to *Leishmania* are present, the antigen forms a pale blue film over the well, which constitutes a positive result (*13*).

The DAT test procedure requires an overnight incubation period, so results are not immediately available. That is why DATs are recommended at the district hospital level. Despite these limitations, DAT has excellent clinical accuracy and high precision in diagnosing VL cases. By enabling early and correct diagnosis and timely start of treatment, it can significantly impact patient outcomes, including curing the disease and thus helping to control VL in the endemic community.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Out of the 194 countries and six territories reporting to WHO, 75 (38%) are currently considered endemic for VL: most of these are in South-East Asia, eastern Africa and Brazil. In Asia and eastern Africa, the disease is caused by *L. donovani*; in Brazil (and elsewhere in Latin America, Europe and North Africa), it is caused by L. infantum (*14*).

There are an estimated 300 000 new cases of VL and 20 000–30 000 associated deaths each year, of which only 25–45% are reported to WHO. In 2017, more than 95% of cases were reported from just 10 countries: Bangladesh, Brazil, China, Ethiopia, India, Kenya, Nepal, Somalia, South Sudan and Sudan (*15*).

The disease has a huge economic impact on affected populations and communities. For example, in Bihar, India, 83% of households in communities with high attack rates of VL were also among the poorest 40%. Evidence is arguably most complete for VL, with studies from multiple countries showing that even when diagnosis and medicines are provided free of charge, between 25% and 75% of households of sufferers experience some type of financial catastrophe.

F. WHO or other clinical guidelines relevant to the test

The WHO Expert Committee on the Control of the Leishmaniases (*16*) issued international guidelines on controlling VL in 2010.

Several countries where VL is endemic also have national guidelines for diagnosing, treating and preventing the disease, including Ethiopia (*17*), Kenya (*18*), South Sudan (*20*), Sudan (*19*) and Uganda (*21*).

G. Basic test characteristics (from application)

Test formats available	Freeze-dried antigen DAT
Specimen types	Serum, DBS
Equipment required	Microplates, plate sealers, syringes, plastic stirring rods, dispensing and pipetting devices
Regulatory status	No information provided
Availability	Worldwide on request from Amsterdam University Medical Center
Price per test range	US$ 3–6
Instrument price range	Not applicable

H. Evidence for diagnostic accuracy (from application)

DAT is an easy-to-perform test that is widely applicable with high sensitivity (90–100%) and specificity (95–100%) and that has routinely been used in some regions for the past 2 decades *(22–25)*. The test can be carried out using plasma, serum or even urine samples, making it suitable for both field and laboratory application (*26–32*).

There have been several validation studies on the clinical accuracy of DAT. A meta-analysis by Chappuis et al. (*32*) found DAT to be almost 1% more sensitive and 2% more specific than the rK39 strip test.

A systematic review with meta-analysis of the rK39 strip test compared with DAT and IIF test and ELISA showed that sensitivity was 94.23% and specificity 89.97% and that the likelihood ratio of a positive test was 9.39 (*33*).

Another study in Sudan by Abdullah et al. (*30*) evaluated the DAT for serodiagnosis of VL based on freeze-dried *L. donovani* antigen and found it comparable with standard liquid antigen by testing serum and blood samples. The freeze-dried DAT was found to have a sensitivity of 96.8% and a specificity of 96.2%.

I. Evidence for clinical usefulness and impact (from application)

There are no systematic reviews of the test's clinical utility or impact on patient management and care. There are, however, several studies available to show that early diagnosis and complete treatment positively impact the control of VL in very high endemic settings.

In one study, about 6.2% of normal persons in VL endemic areas were found to be reactive to DAT, with 3.6% becoming seropositive in a year's time (*34*). In another study, the same author showed a strong association between serological status and probability of progression to clinical VL in prospective cohort studies in India and Nepal by using DAT titres, thereby demonstrating the utility of DAT not only as a diagnostic test but also for field surveillance (*35*).

J. Evidence for economic impact and/or cost–effectiveness (from application)

Boelaert et al. (*31*) show that the availability of DAT has proved convenient for use in field conditions. Since the available options are invasive techniques (splenic, bone marrow or lymph node aspirations) or serologic tests (DAT or rK39), the latter are clearly more cost-effective.

K. Ethics, equity and human rights issues (from application)

Availability of serological tests for VL have greatly improved access to diagnosis and treatment.

L. Summary of evidence evaluation

There is evidence that the accuracy of the DAT is comparable to the rK39 and suitable for use as a diagnostic test in endemic regions, as is recommended in WHO and many national guidelines. Estimates of accuracy are only available from reviews more than 10 years old that suggest sensitivity of 95% (CI: 93–96%) and specificity of 86% (CI: 72–93%).

M. Summary of SAGE IVD deliberations

As well as being supported by WHO guidelines for the diagnosis of VL, the DAT is a non-invasive, cost-effective test compared with alternative tests based on microscopy of bone marrow or lymph node aspirates or spleen puncture. It has a higher diagnostic accuracy compared with rK39 RDT in eastern Africa; and a higher sensitivity, negative predictive value (NPV) and PPV compared with rK39 for screening and diagnosis among HIV-infected patients. Its higher NPV among HIV-infected patients could help avoid unnecessary treatment with toxic medications in non-infected people.

Importantly, the DAT can be used in test-and-treat strategies targeting VL in LMICs, where the disease is often endemic, because it is easy to perform

and requires minimum laboratory skills; it can be done on DBSs that are collected in primary care settings and sent to higher care levels.

N. SAGE IVD recommendations

SAGE IVD recommended including the visceral leishmaniasis direct agglutination test in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, Neglected tropical diseases);
- in an agglutination assay format;
- to aid in the diagnosis of clinically suspected visceral leishmaniasis.

The group noted the need to ensure that the entry for this test category in EDL 3 is clearly linked to WHO guidelines stating that the VL DAT must be used in combination with a clinical definition.

O. References

1. Herwaldt BL. Leishmaniasis. Lancet. 1999;354(9185):1191–1199. doi:10.1016/S0140-6736(98)10178-2.
2. Chappuis F, Sundar S, Hailu A, Ghalib H, Rijal S, et al. Visceral leishmaniasis: what are the needs for diagnosis, treatment and control? Nat Rev Microbiol. 2007;5(11):873–882. doi:10.1038/nrmicro1748.
3. Schaefer KU, Kurtzhals JA, Gachihi GS, Muller AS, Kager PA. A prospective sero-epidemiological study of visceral leishmaniasis in Baringo District, Rift Valley Province, Kenya. Trans R Soc Trop Med Hyg. 1995;89(5):471–475. doi:10.1016/0035-9203(95)90070-5.
4. Evans TG, Teixeira MJ, McAuliffe IT, Vasconcelos I, Vasconcelos AW, et al. Epidemiology of visceral leishmaniasis in northeast Brazil. J Infect Dis. 1992;166(5):1124–1132. doi:10.1093/infdis/166.5.1124.
5. Ali A, Ashford RW. Visceral leishmaniasis in Ethiopia. IV. Prevalence, incidence and relation of infection to disease in an endemic area. Ann Trop Med Parasitol. 1994;88(3):289–293. doi:10.1080/00034983.1994.11812869.
6. Ostyn B, Gidwani K, Khanal B, Picado A, Chappuis F, et al. Incidence of symptomatic and asymptomatic Leishmania donovani infections in high-endemic foci in India and Nepal: a prospective study. PLoS Negl Trop Dis. 2011;5(10):e1284. doi:10.1371/journal.pntd.0001284.
7. Zijlstra EE, el-Hassan AM, Ismael A, Ghalib HW. Endemic kala-azar in eastern Sudan: a longitudinal study on the incidence of clinical and subclinical infection and post-kala-azar dermal leishmaniasis. Am J Trop Med Hyg. 1994;51(6):826–836. doi:10.4269/ajtmh.1994.51.826.
8. Kosack CS, Page AL, Klatser PR. A guide to aid the selection of diagnostic tests. Bull World Health Organ. 2017;95:639–645. doi:10.2471/BLT.16.187468.
9. Boelaert M, Verdonck K, Menten J, Sunyoto T, van Griensven J, et al. Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease. Cochrane Database Syst Rev. 2014;6:CD009135. doi:10.1002/14651858.CD009135.pub2.

10. El Harith A, Kolk AH, Kager PA, Leeuwenburg J, Muigai R, et al. A simple and economical direct agglutination test for serodiagnosis and sero-epidemiological studies of visceral leishmaniasis, Trans R Soc Trop Med Hyg. 1986;80(4):583–536. doi:10.1016/0035-9203(86)90149-5.

11. El Harith A, Kolk AH, Kager PA, Leeuwenburg J, Faber FJ, et al. Evaluation of a newly developed direct agglutination test (DAT) for serodiagnosis and sero-epidemiological studies of visceral leishmaniasis: comparison with IFAT and ELISA. Trans R Soc Trop Med Hyg. 1987;81(4):603–606. doi:10.1016/0035-9203(87)90423-8.

12. El Harith A, Kolk AH, Leeuwenburg J, Muigai R, Huigen E, et al. Improvement of a direct agglutination test for field studies of visceral leishmaniasis. J Clin Microbiol. 1988;26:1321–1325.

13. Adams ER, Jacquet D, Schoone G, Gidwani K, Boelaert M, et al. Leishmaniasis direct agglutination test: using pictorials as training materials to reduce inter-reader variability and improve accuracy. PLoS Negl Trop Dis. 2012;6(12):e1946. doi:10.1371/journal.pntd.0001946.

14. Boelaert M, Criel B, Leeuwenburg J, Van Damme W, Le Ray D, et al. Visceral leishmaniasis control: a public health perspective. Trans R Soc Trop Med Hyg. 2000;94(5):465–471. doi:10.1016/s0035-9203(00)90055-5.

15. Leishmaniasis. In: WHO/Newsroom/Fact sheets [website]. Geneva: World Health Organization; 2020 (https://www.who.int/news-room/fact-sheets/detail/leishmaniasis, accessed 25 November 2019).

16. Control of the leishmaniases: report of a meeting of the WHO Expert Committee on the Control of Leishmaniases. Geneva: World Health Organization; 2010 (WHO Technical Report Series, No 949; https://apps.who.int/iris/handle/10665/44412 accessed 25 March 2020).

17. Guideline for diagnosis, treatment and prevention of leishmaniasis in Ethiopia, 2nd edition. Addis Ababa: Ministry of Health Ethiopia; 2013.

18. Prevention, diagnosis and treatment of visceral leishmaniasis (kala-azar) in Kenya: national guidelines for health workers. Nairobi: Kenya Ministry of Health; 2017 (https://www.who.int/leishmaniasis/burden/Kala_Azar_Kenya_2017.pdf?ua=1, accessed 25 March 2020).

19. Manual for the diagnosis and treatment of leishmaniasis. Khartoum: Sudan Federal Ministry of Health; 2014 (https://www.who.int/leishmaniasis/burden/Manual_for_the_diagnosis_and_treatment_Leishmaniasis_Guideline_Sudan_2014.pdf?ua=1, accessed 25 March 2020).

20. Guidelines for diagnosis, treatment and prevention of visceral leishmaniasis in South Sudan. Juba: Ministry of Health (https://www.who.int/leishmaniasis/burden/Guidelines_for_diagnosis_treatment_and_prevention_of_VL_in_South_Sudan.pdf, accessed 25 March 2020).

21. Guidelines for the diagnosis, treatment and prevention of visceral leishmaniasis in Uganda. Kampala: Uganda Ministry of Health; 2019 (https://www.who.int/leishmaniasis/burden/MOH_Uganda_Guidelines_diagnosis_treatment_prevention_VL.pdf?ua=1, accessed 25 March 2020).

22. Sreenivas G, Ansari NA, Singh R, Subba Raju BV, Bhatheja R, et al. Diagnosis of visceral leishmaniasis: comparative potential of amastigote antigen, recombinant antigen and PCR. Br J Biomed Sci. 2002;59:218–222.

23. el Harith A, el Mutasim M, Mansour D, Fadil Mustafa E, Arvidson H. Use of glycerol as an alternative to freeze-drying for long-term preservation of antigen for the direct agglutination test. Trop Med Int Health. 2003;8:1025–1029. doi:10.1046/j.1360-2276.2003.01129.x.

24. el Mutasim, Mansour MD, Abass EM, Hassan WM, el Harith A. Evaluation of a glycerol-preserved antigen in the direct agglutination test for diagnosis of visceral leishmaniasis at rural level in eastern Sudan. J Med Microbiol. 2006;55:1343–1347. doi:10.1099/jmm.0.46753-0.

25. Jacquet D, Boelaert M, Seaman J, Rijal S, Sundar S, et al. Comparative evaluation of freeze-dried and liquid antigens in the direct agglutination test for serodiagnosis of visceral leishmaniasis (Itma-Dat/Vl). Trop Med Int Health. 2006;11:1777–1784. doi:10.1111/j.1365-3156.2006.01743.x.

26. Sundar S, Singh RK, Bimal SK, Gidwani K, Mishra A, et al. Comparative evaluation of parasitology and serological tests in the diagnosis of visceral leishmaniasis in India: a phase III diagnostic accuracy study. Trop Med Int Health. 2007;12:284–289. doi:10.1111/j.1365-3156.2006.01775.x.

27. Mandal J, Khurana S, Dubey ML, Bhatia P, Varma N, et al. Evaluation of direct agglutination test, rK39 test, and ELISA for the diagnosis of visceral leishmaniasis. Am J Trop Med Hyg. 2008;79:76–78.

28. Bhattarai NR, Van der Auwera G, Khanal B, De Doncker S, Rijal S, et al. PCR and direct agglutination as Leishmania infection markers among healthy Nepalese subjects living in areas endemic for kala-azar. Trop Med Int Health. 2009;14:404–411. doi:10.1111/j.1365-3156.2009.02242.x.

29. Oliveira E, Pedras MJ, de Assis IE, Rabello A. Improvement of direct agglutination test (DAT) for laboratory diagnosis of visceral leishmaniasis in Brazil. Trans R Soc Trop Med Hyg. 2009;103:1279–1281. doi:10.1016/j.trstmh.2009.04.007.

30. Abdallah KAA, Nour BYM, Schallig HDFH, Mergani A, Hamid Z, et al. Evaluation of the direct agglutination test based on freeze-dried Leishmania donovani promastigotes for the serodiagnosis of visceral leishmaniasis in Sudanese patients. Trop Med Int Health. 2004;9(10):1127–1131. doi:10.1111/j.1365-3156.2004.01308.x.

31. Boelaert M, el Safi S, Mousa H, Githure J, Mbati P, et al. Multi-centre evaluation of repeatability and reproducibility of the direct agglutination test for visceral leishmaniasis. Trop Med Int Health. 1999;4(1):31–37. doi:10.1046/j.1365-3156.1999.00348.x.

32. Chappuis F, Rijal S, Soto A, Menten J, Boelaert M. A meta-analysis of the diagnostic performance of the direct agglutination test and rK39 dipstick for visceral leishmaniasis. BMJ. 2006;333(7571):723. doi:10.1136/bmj.38917.503056.7C.

33. Maia Z, Lírio M, Mistro S, Mendes CMC, Mehta SR, et al. Comparative study of rK39 Leishmania antigen for serodiagnosis of visceral leishmaniasis: systematic review with meta-analysis. PLoS Negl Trop Dis. 2012;6(1):e1484. doi:10.1371/journal.pntd.0001484.

34. Hasker E, Kansal S, Malaviya P, Gidwani K, Picado A, et al. Latent infection with Leishmania donovani in highly endemic villages in Bihar, India. PLoS Negl Trop Dis. 2013;7(4). doi:10.1371/journal.pntd.0002053.

35. Hasker E, Malaviya P, Gidwani K, Picado A, Ostyn B, et al. Strong association between serological status and probability of progression to clinical visceral leishmaniasis in prospective cohort studies in India and Nepal. PLoS Negl Trop Dis. 2014;8(1):e2657. doi:10.1371/journal.pntd.0002657.

PNEUMOCYSTIS PNEUMONIA

3.19. *Pneumocystis jirovecii* nucleic acid test

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a nucleic acid amplification test for *P. jirovecii* DNA to the EDL as an IVD to diagnose and exclude *Pneumocystis* pneumonia (PCP).

B. Applicant

Global Action Fund for Fungal Infections

C. WHO technical department

None

D. Background (from application)

Disease condition and impact on patients

PCP is a life-threatening illness of largely immunosuppressed patients, such as those with HIV/AIDS (*1*). When diagnosed rapidly and treated, survival rates are high. The etiologic agent of PCP is *P. jirovecii*, a human-only, commensal fungus that does not grow on microbiological agar. Other mammals have their own *Pneumocystis* species. AIDS-related PCP has a variable incidence: 5.9–55% in adults and 5–49% in children (1). In the 1980s, during the HIV pandemic, PCP was one of the most prevalent AIDS-defining diseases in HICs; this remains an issue in cases of undiagnosed HIV, particularly in LMICs (*2*).

Untreated, the mortality is 100%. In HICs, the mortality is about 10% among AIDS patients, but it is about 30% in LMICs (*1*). Overall, the day 30 all-cause mortality is ~28%, ranging from 10% to 62%; it is higher in HIV-negative patients due to fulminant disease (*3*). In addition to high-dose cotrimoxazole, adjunctive corticosteroids reduce mortality in AIDS in moderate and severe cases, but are unhelpful in non-AIDS cases. Second-line therapy requires clindamycin and primaquine.

Does the test meet a medical need?

The historical difficulties in diagnosing PCP (lack of culture and poor sensitivity of microscopic investigation) have resulted in many cases being diagnosed on

clinical suspicion (*3*). Clinically, signs are non-specific. Typically, bilateral chest infection presents with respiratory signs of distress, which can be mild in the HIV positive but are typically more severe in HIV-negative. Radiological signs can be absent in the early stages; when present, although typical of PCP, they are generally non-specific (bilateral ground-glass opacification, leading to consolidation) and could also represent an underlying condition (*4*, *5*). With the advent of modern diagnostic techniques – real-time PCR and (1-3)-β-D-glucan (BDG) – reliable laboratory-based diagnosis can now be achieved.

Cotrimoxazole (trimethoprim-sulfamethoxazole) is the most effective agent for both preventing and treating PCP. In the absence of laboratory diagnosis, cotrimoxazole is instituted empirically and a 3-week course prescribed. If a precise laboratory diagnosis is made, empirical and possibly inappropriate use of cotrimoxazole could potentially be prevented.

How the test is used

The current gold standard for confirming a diagnosis remains histological and/or microscopic identification of cysts and trophic forms of *P. jirovecii* in clinical specimens, usually respiratory samples using conventional or immunofluorescent (IF) antibody stains. Standard microscopic investigation is highly subjective; this can influence specificity and sensitivity. While IF microscopy improves sensitivity, it is still suboptimal (67%).

PCP PCR is primarily a diagnostic test. It is used in conjunction with a lactate dehydrogenase (LDH) test, oxygen saturation and radiology. Detection of *Pneumocystis* DNA in blood is a poor prognostic feature.

E. Public health relevance (from application)

Prevalence

There are an estimated 400 000 PCP cases in AIDS patients annually and more than 100 000 in patients with other immunosuppressive diseases (*6*). In addition to patients diagnosed with HIV, an ever increasingly susceptible HIV-negative population is at risk of PCP (*3*, *7*–*11*).

Those at increased risk include patients with solid tumours or suffering from haematological malignancy, solid organ transplant recipients, patients with autoimmune and inflammatory conditions receiving immunomodulating therapies (e.g. high-dose corticosteroids or anti-TNF [tumour necrosis factor] therapy) and patients diagnosed with primary immune deficiencies (*11*, *12*).

Generally, the incidence of disease is low (*8*–*10*). But the epidemiology of PCP is changing, associated with more aggressive immunosuppressive and immunomodulatory approaches when managing auto-immune conditions (RA and vasculitis), pre-existing respiratory conditions (COPD), and other conditions (haematology and transplants, particularly renal transplants) (*3*, *7*).

Socioeconomic impact

In a recent analysis in the USA, mean inpatient costs for PCP patients on various insurance plans were US$ 23 342–63 388 per visit; for outpatients they were US$ 526–1061 per visit (*13*).

F. WHO or other clinical guidelines relevant to the test

The following organizations all have guidelines relevant to the test:

- British Society for Medical Mycology 2015 (*14*)
- Infectious Diseases Working Party of the German Society of Hematology and Medical Oncology (DGHO) 2015 (*15*)
- Fifth European Conference on Infections in Leukaemia (ECIL-5) 2016 (*16*)
- American Society of Transplantation Infectious Diseases Community of Practice 2019 (*17*)
- CDC, the National Institutes of Health and the HIV Medicine Association of IDSA 2019 (*18*).

ECIL-5 is a joint venture of the European Group for Blood Marrow Transplantation, the European Organisation for Research and Treatment of Cancer, the Immunocompromised Host Society and the European LeukemiaNet.

G. Basic test characteristics (from application)

Test formats available	Nucleic acid amplification
Specimen types	Respiratory (sputum, BAL fluid)
Equipment required	PCR thermocycler
Regulatory status	CE-marked
Availability	Global, but mostly in high- and middle-income countries
Price per test range	US$ 7–15, notwithstanding control tests
Instrument price range	US$ 6000–45 000

H. Evidence for diagnostic accuracy (from application)

PCP PCR is a semi-quantitative or quantitative test, validated by various manufacturers against sputum, BAL and other respiratory specimens. Measures of diagnostic accuracy considered therefore include accuracy of various thresholds to differentiate infection versus colonization.

Meta-analytical reviews of PCP PCR on respiratory samples generate excellent sensitivity of ≥ 97%; the subsequent NPV of ≥ 99% is sufficient to rule out PCP when PCR is negative (*19–21*). Despite the detection of possible *Pneumocystis* colonization, positivity in respiratory samples readily confirms disease as shown by positive likelihood ratios of ≥ 10.

PCR testing has been applied to a range of specimen types (BAL fluid, blood). PCR testing of BAL fluid is preferred, but positivity in upper airway samples (sputum, induced sputum, oral washes and nasopharyngeal aspirates), once thought to represent detection of transient colonization, likely reflects a significant burden lower in the respiratory tract and is specific for PCP (*16*).

Expectorated and induced sputum samples have been studied in Africa and have good sensitivity (*22*, *23*). In Namibia, 475 samples were analysed and 25 (5.3%) samples were positive for *P. jirovecii*; 17 (3.6%) using both real-time PCR and Grocott's methenamine silver (GMS) staining and eight (1.7%) using real-time PCR only (*22*). *P. jirovecii* was present in 8/150 (5.3%) HIV-positive and TB smear-negative patients, and in 12/227 (5.3%) TB smear-negative patients with an unknown HIV status. In South Africa, *P. jirovecii* was identified in 51% (156/305) and 67% (204/305) specimens using immunofluorescence and real-time PCR, respectively (*23*). The cut-off value for the real-time PCR that best predicted the reference IFA results was 78 copies/5 µL (area under ROC curve 0.92). The sensitivity and specificity of real-time PCR using this cut-off was 94.6% and 89.1%, respectively, compared with the IFA.

In children, PCP PCR is invaluable for diagnosis. Morrow et al. (*24*) studied 202 children (median age 3.2 months) in South Africa, including 124 (61.4%) who were infected with HIV. They identified PCP in 109 (54%) children using PCR, compared with 43 (21%) using IFA and GMS ($P < 0.0001$). Most PCP cases (88, 81%) occurred in HIV-infected children. All 21 cases (19%) occurring in HIV-negative children had another risk factor for PCP.

Das et al. (*25*) investigated 94 immunocompromised children with pneumonia for PCP in India. PCR detected *P. jirovecii* in 14 children. The occurrence of PCP in HIV-infected children was 43% (6/14), renal disease on immunosuppressants 45% (4/9), primary immune deficiency 19% (2/11) and malignancies on chemotherapy 4% (2/57).

Wang et al. (*26*) suggest that the presence of *Pneumocystis* DNA in blood samples is a poor prognostic marker. Detection of DNA in the plasma of HIV-positive patients was significantly higher in deceased patients (79%) compared with survivors (14%), as was the burden of disease (deceased: 54 610 copies/mL vs survivors: 935 copies/mL.

Choi et al. (*27*) use PCR to determine prognosis. In their study of 81 HIV-negative PCP patients with respiratory failure that were initially PCR positive, PCP PCR-negative conversion was associated with a good prognosis, generating a hazard ratio of 0.433 (95% CI: 0.203–0.928, $P = 0.031$).

Performance of PCR may vary between HIV-positive and HIV-negative patients. The lower burden encountered in HIV-negative patients may lower sensitivity, but PCR specificity remains high (*16*). Nevertheless, PCR negativity when testing BAL fluid can be used to exclude disease irrespective of the underlying condition (ECIL guidelines level of evidence: II; grade of recommendation: A).

I. Evidence for clinical usefulness and impact (from application)

Almost no real-life studies of the role of PCP PCR in LMICs have been done because the test is not generally available in these countries. Oladele et al. (*28*) reviewed the topic but did not do a systematic review, as this is not possible with the current state of knowledge.

J. Evidence for economic impact and/or cost–effectiveness (from application)

No reviews or primary studies of the economic impact or cost–effectiveness of the PCP PCR test are available.

The price per test (excluding control testing, human resource and processing costs) varies.

The development of the first freeze-dried assay using loop-mediated isothermal amplification (LAMP) technology is expected to greatly reduce shipping cost and enable transport to centres without a major airport. While the price per assay or the equipment cost cannot be determined, it is notable that no extraction system is required for this test format; the result is available in 25 minutes.

K. Ethics, equity and human rights issues (from application)

The test is for identifying and excluding PCP in vulnerable immunocompromised and HIV-infected individuals; it raises no concern for equity or breach of human rights.

L. Summary of evidence evaluation

The guidelines related to the test used rigorous methods, and the recommendations were based on evidence varying from low to moderate quality. The meta-analysis suffered from high heterogeneity, but the confidence intervals were rather small. The primary studies selected children hospitalized with suspected PCP in a prospective manner and using valid reference standards. All patients were included in the analysis. The challenge is that PCR seems to detect more PCP cases than the current reference standard.

In summary, the test was evaluated in various subgroups of patients on various specimens. No recommendations in guidelines were made specifically for children, and only two primary studies in children were added to the evidence

portfolio. With the sensitivity of the test at more than 97%, it is safe to rule out patients to avoid inappropriate use of cotrimoxazole. But the test cannot distinguish infection from carriage, which might hamper clinical usefulness.

M. Summary of SAGE IVD deliberations

PCP is a significant public health problem, whose incidence in LMICs is closely linked with the burden of HIV disease and, more specifically, the proportion of patients who still present with advanced disease. *P. jirovecii* cannot be cultured and has historically been diagnosed by stains or antibody-based microscopy, with limited sensitivity.

Molecular assays are valuable in confirming *P. jirovecii* as the causative agent of pneumonia, when combined with clinical signs and symptoms. In these cases, the assays are also reliable for excluding PCP and thus preventing inappropriate use of cotrimoxazole.

But current tests are not inexpensive; and they are only appropriate for centralized testing. SAGE IVD also noted that while the submission broadly covers NAT formats, the evidence base provided is predominantly for PCR tests. Group members noted that LAMP tests tend not to be as sensitive as PCR, although LAMP tests do have advantages, particularly in LMICs.

In considering the use of PCR tests, SAGE IVD raised some concerns about the potential for variability in results. Some procedures use conventional PCR, while others use real-time PCR. And the specimen type includes a broad range of respiratory tract specimens, from sputum to BAL to tracheal secretion. Diagnostic accuracies vary with both specimen and technology type. The group confirmed, however, that this potential for variability in results applies to many other conditions where different PCR types and specimens are used. SAGE-IVD also acknowledged that addressing quality assurance and standardization issues is beyond the scope of the EDL (although the group emphasized that this should be done as and when countries establish and implement their own national EDLs).

N. SAGE IVD recommendations

SAGE IVD recommended including the *P. jirovecii* test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b);
- using a nucleic acid test format;
- to aid in the diagnosis of Pneumocystis pneumonia.

The group requested the addition of a note to the test category entry in the EDL stating that it is particularly relevant in immunocompromised patients.

O. References

1. GAFFI priority fungal infections. In: GAFFI/Fact sheets [website]. Geneva: Global Action Fund for Fungal Infections; 2020 (https://www.gaffi.org/media/fact-sheets, accessed 12 March 2020).
2. Hay JW, Osmond DH, Jacobson MA. Projecting the medical costs of AIDS and ARC in the United States. J Acquir Immune Defic Syndr. 1988;1(5):466–485.
3. Maini R, Henderson KL, Sheridan EA, Lamagni T, Nichols G, et al. Increasing Pneumocystis pneumonia, England, UK, 2000–2010. Emerg Infect Dis. 2013;19(3):386–392. doi:10.3201/eid1903.121151.
4. Block B, Mehta T, Ortiz GM, Ferris SP, Vu TH, et al. Unusual radiographic presentation of Pneumocystis pneumonia in a patient with AIDS. Case Rep Infect Dis. 2017;3183525. doi:10.1155/2017/3183525.
5. Cereser L, Dallorto A, Candoni A, Volpetti S, Righi E, et al. Pneumocystis jirovecii pneumonia at chest high-resolution computed tomography (HRCT) in non-HIV immunocompromised patients: spectrum of findings and mimickers. Eur J Radiol. 2019;116:116–127. doi:10.1016/j.ejrad.2019.04.025.
6. Roadmap. In: GAFFI/Roadmap [website]. Geneva: Global Action Fund for Fungal Infections; 2020 (https://www.gaffi.org/roadmap, accessed 12 March 2020).
7. Patterson L, Coyle P, Curran T, Verlander NQ, Johnston J. Changing epidemiology of Pneumocystis pneumonia, Northern Ireland, UK and implications for prevention, 1 July 2011–31 July 2012. J Med Microbiol. 2017;66:1650–165. doi:10.1099/jmm.0.000617.
8. Pegorie M, Denning DW, Welfare W. Estimating the burden of invasive and serious fungal disease in the United Kingdom. J Infect. 2016;74(1):60–71. doi:10.1016/j.jinf.2016.10.005.
9. Buchacz K, Lau B, Jing Y, Bosch R, Abraham AG, et al. Incidence of AIDS-defining opportunistic infections in a multicohort analysis of HIV-infected persons in the United States and Canada, 2000–2010. J Infect Dis. 2016;214:862–872. doi:10.1093/infdis/jiw085.
10. Williams KM, Ahn KW, Chen M, Aljurf MD, Agwu AL, et al. The incidence, mortality and timing of Pneumocystis jirovecii pneumonia after hematopoietic cell transplantation: a CIBMTR analysis. Bone Marrow Transplant. 2016;51:573–580. doi:10.1038/bmt.2015.316.
11. Avino LJ, Naylor SM, Roecker AM. Pneumocystis jirovecii pneumonia in the non-HIV infected population. Ann Pharmacother. 2016;50(8):673–679. doi:10.1177/1060028016650107.
12. Szydłowicz M, Jakuszko K, Szymczak A, Piesiak P, Kowal A, et al. Prevalence and genotyping of Pneumocystis jirovecii in renal transplant recipients – preliminary report. Parasitol Res. 2018;118(1):181–189. doi:10.1007/s00436-018-6131-0.
13. Benedict K, Jackson BR, Chiller T, Beer KD. Estimation of direct healthcare costs of fungal diseases in the United States. Clin Infect Dis. 2019;68(11):1791–1797. doi:10.1093/cid/ciy776.
14. Schelenz S, Barnes RA, Barton RC, Cleverley JR, Lucas SB, et al. British Society for Medical Mycology best practice recommendations for the diagnosis of serious fungal diseases. Lancet Infect Dis. 2015;15:461–474. doi:10.1016/S1473-3099(15)70006-X.
15. Maschmeyer G, Carratalà J, Buchheidt D, Hamprecht A, Heussel CP, et al. Diagnosis and antimicrobial therapy of lung infiltrates in febrile neutropenic patients (allogeneic SCT excluded): updated guidelines of the Infectious Diseases Working Party (AGIHO) of the German Society of Hematology and Medical Oncology (DGHO). Ann Oncol. 2015;26(1):21–33. doi:10.1093/annonc/mdu192.

16. Alanio A, Hauser PM, Lagrou K, Melchers WJ, Helweg-Larsen J, et al. ECIL guidelines for the diagnosis of Pneumocystis jirovecii pneumonia in patients with haematological malignancies and stem cell transplant recipients. J Antimicrob Chemother. 2016:71(9):2386–2396. doi:10.1093/jac/dkw156.

17. Fishman JA, Gans H, AST Infectious Diseases Community of Practice. Pneumocystis jirovecii in solid organ transplantation: guidelines from the American Society of Transplantation Infectious Diseases Community of Practice. Clin Transplant. 2019;33(9):e13587. doi:10.1111/ctr.13587.

18. Kaplan JE, Benson C, Holmes KK, Brooks JT, Pau A, et al. Guidelines for the prevention and treatment of opportunistic infections in adults and adolescents with HIV: recommendations from CDC, the National Institutes of Health, and the HIV Medicine Association of the Infectious Diseases Society of America. MMWR Recomm Rep. 2009;58(RR-4):1–207.

19. Summah H, Zhu YG, Falagas ME, Vouloumanou EK, Qu JM. Use of real-time polymerase chain reaction for the diagnosis of Pneumocystis pneumonia in immunocompromised patients: a meta-analysis. Chin Med J. 2013;126:1965–1973.

20. Fan LC, Lu HW, Cheng KB, Li HP, Xu JF. Evaluation of PCR in bronchoalveolar lavage fluid for diagnosis of Pneumocystis jirovecii pneumonia: a bivariate meta-analysis and systematic review. PLoS One. 2013;8(9):e73099. doi:10.1371/journal.pone.0073099.

21. Lu Y, Ling G, Qiang C, Ming Q, Wu C, et al. PCR diagnosis of Pneumocystis pneumonia: a bivariate meta-analysis. J Clin Microbiol. 2011;49(12):4361–4363. doi:10.1128/JCM.06066-11.

22. Nowaseb V, Gaeb E, Fraczek MG, Richardson MD, Denning DW. Frequency of Pneumocystis jirovecii in sputum from HIV and TB patients in Namibia. J Infect Dev Ctries. 2014;8(3):349–357. doi:10.3855/jidc.3864.

23. Moodley B, Tempia S, Frean JA. Comparison of quantitative real-time PCR and direct immunofluorescence for the detection of Pneumocystis jirovecii. PLoS One. 2017;12(7):e0180589. doi:10.1371/journal.pone.0180589.

24. Morrow BM, Samuel CM, Zampoli M, Whitelaw A, Zar HJ. Pneumocystis pneumonia in South African children diagnosed by molecular methods. BMC Res Notes. 2014;7:26. doi:10.1186/1756-0500-7-26.

25. Das CK, Mirdha BR, Singh S, Seth R, Bagga A, et al. Use of induced sputum to determine the prevalence of Pneumocystis jirovecii in immunocompromised children with pneumonia. J Trop Pediatr. 2014;60(3):216–222. doi:10.1093/tropej/fmt112.

26. Wang M, Dai X, Huang Y, Sun J, Dong X, et al. The presence of Pneumocystis jirovecii DNA in plasma is associated with a higher mortality rate in patients with AIDS-associated Pneumocystis pneumonia. Med Mycol. 2018:Epub. doi:10.1093/mmy/myy110.

27. Choi JS, Lee S, Leem AY, Song JH, Kim SY, et al. Pneumocystis jirovecii pneumonia (PCP) PCR-negative conversion predicts prognosis of HIV-negative patients with PCP and acute respiratory failure. PLoS One. 2018;13(10):e0206231. doi:10.1371/journal.pone.0206231.

28. Oladele RO, Otu AA, Richardson MD, Denning DW. Diagnosis and management of Pneumocystis pneumonia in resource-poor settings. J Health Care Poor Underserved. 2018;29(1):107–158. doi:10.1353/hpu.2018"0010.

PRIMARY IMMUNODEFICIENCIES

3.20. Serum and urine protein electrophoresis for primary immunodeficiency

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding serum and urine PEP to the EDL as an IVD for diagnosing primary immunodeficiencies (PIDs), especially those caused by disorders of adaptive immunity.

B. Applicant

International Patient Organisation for Primary Immunodeficiencies

C. WHO technical department

Noncommunicable diseases

D. Background (from application)

Disease condition and impact on patients

People with PIDs have no protection against common pathogenic organisms, and as a result suffer lifelong life-threatening infections and increasing, permanent damage to various body organs, especially the lungs and intestines. This increasing damage with each infection renders the person more susceptible to more severe and frequent infections. Many of the conditions that are present in childhood are genetic, and parental consanguinity is a risk factor for developing PID. A positive family history is an important indication for screening family members for suspected PID.

In a study of 32 patients with PIDs, more than two-thirds experienced diagnostic delay, which led to severe illnesses requiring hospitalization, including pneumonia, meningitis, osteomyelitis and septicaemia (*1*). A research paper using data from the USA found that the proportion of persons with a hospital admission was higher each year from 2001 to 2005 among those with PID (13.9–18.6%) compared with those without a PID diagnosis (7.9–8.9%). People with a PID had significantly longer hospital stays. The most common comorbidities among hospitalized patients with PID included non-specific infection-associated comorbidities (fevers, splenomegaly, failure to thrive), respiratory infections, pathogen-specific infections and chronic lung disease (*2*).

The data from a new service to more than 1000 suspected PID patients in Asia showed that families often lost one or more children to an undiagnosed PID before the current child was diagnosed (*3*). A long delay in diagnosis and a large number of infectious episodes can further negatively impact health-related quality of life and lead to permanent functional impairments. A recent large-scale analysis of a cohort of 2212 patients with common variable immunodeficiencies (CVIDs) found that each year of diagnostic delay and each year of increased age at diagnosis were associated with a 1.7% and 4.5% increased risk of death, respectively (*1*).

In comparison with healthy children and adults, patients with PIDs experience measurably lower general health with higher hospitalization rates and increased limitations on physical, school and social activity (*1*).

Does the test meet a medical need?

Clinical data have clearly shown that appropriate treatment with polyvalent human immunoglobulins of patients affected by PIDs can be lifesaving as well as greatly improving quality of life. Some PIDs do not require immunoglobulin replacement therapy but can be cured by allogeneic stem cell transplantation (or even gene therapy in a few instances). PIDs are diverse, but they all increase susceptibility to infection and result in substantial illness and shortened lifespans. In most cases, prompt diagnosis and treatment can enable life-saving treatments (including antibiotics, replacement with stem cells and immunoglobulins) and improve quality and length of life (*4*). Yet failure to recognize PIDs is still a substantial challenge for clinicians around the world (*5*). One major problem is that many general practitioners, physicians and paediatricians are not familiar with PIDs and lack guidance on diagnostic tests. This is particularly true in LMICs. Use of the EML and WHO formulary is helping to tackle this problem and raise awareness of these treatable conditions.

There are many types of PIDs, depending on which part of the immune system is affected. Antibody failure is the most common, but more severe forms due to several immune defects in one individual are also easy to diagnose with the correct tools. Without diagnosis and replacement of stem cells, these more severe PIDs are fatal within 2 years of birth. The largest clinical data collection on the natural history of PIDs comes from a 1955–1966 UK Medical Research Council study, which aimed to capture nearly all the cases occurring in the United Kingdom. The study admitted 184 untreated patients (ascertainment has risen > 100-fold since that time) and found the 10-year survival rate to be 36%. Another study by Cunningham-Rundles and Bodian, on 248 treated patients diagnosed by the Immunodeficiency Clinic at Mount Sinai Medical Center from 1973 to 1998, found 10-year survival rates had risen to 78%, although they still lagged behind the USA's general figures of 98%. With more awareness and higher

doses of replacement immunoglobulins, survival rates have now improved further to 92%. Recent data from Iran confirms the hugely increased serious infection burden in patients without immunoglobulin treatment compared with those receiving polyvalent human immunoglobulins (*6*).

How the test is used

While it is true that PID patients rarely develop myeloma, primary antibody deficiency is a diagnosis of exclusion in younger and adult patients whose immune systems fail for unknown reasons (without haematological malignancy). The test is on the list for haematology, and should be recognized as essential for PIDs, especially in adults. Serum and urine PEP should be performed as part of the diagnostic investigation of suspected PIDs and secondary immunodeficiencies, as well as other diseases (plasma cell dyscrasias such as myeloma, lymphoma, chronic lymphatic leukaemia, etc.).

Serum and urine PEP is widely used in almost all countries, particularly poorer ones. It is important for adult PID patients (who represent more than 80% of all PIDs) and for less well resourced countries. These tests are very much used in Russia (adult hospital, 100%), Benin (serum PEP 50%), Kenya (less than 50%), Uruguay (100% of diagnosis), Bolivia (maybe 40%), Paraguay (70%) and Sudan (about 50–80%). Some of the other hospitals consulted did not use this technique because it is mainly used to diagnose adults and the hospitals mainly work with children.

One of the two experts consulted in Vietnam said: "We do not use these tests in our hospital since we can do immunoglobulin level. But I often suggest doctors in provincial hospitals do PEP when they have a PID suspected patient, as they cannot do other tests. The price is around US$ 15".

We have received similar comments from contacts in India, where the hospital consulted said it used to use PEP tests when immunoglobulin assay kits were not readily available or were expensive, adding that they are good for screening PID patients at risk for B cell lymphomas producing a monoclonal immunoglobulin.

E. Public health relevance (from application)

Prevalence

PIDs are a large and growing group of more than 350 different inherited disorders caused when some components of the immune system (mainly cells and proteins) do not work properly. Recent studies have shown that PIDs may be more common than previously estimated, and that as many as 1% of the population may be affected with a PID when all types and varieties are considered.

Social and economic impact

Diagnostic delays are not only damaging to the patient as the condition deteriorates and difficult for the family caring for their relative; they also negatively impact health care systems by causing inappropriate use of health resources through avoidable visits and admissions to hospital, involving a variety of different specialists for recurring infections. Modell et al. (*7*) have shown that early diagnosis reduces post-diagnosis costs compared with the year before diagnosis, even if regular immunoglobulin replacement therapy is needed. The authors calculate that the annual savings to the health care system for each diagnosed patient is US$ 85 882, and that even for patients who are diagnosed and treated with immunoglobulin, there are annual savings of US$ 55 882. In 2018, Condino-Neto and Espinosa-Rosales (*8*) similarly documented the financial impact associated with early diagnosis and management of PIDs in the USA and found that the average annual cost to the health care system for each undiagnosed patient with an underlying PID is US$ 102 552, while diagnosis and treatment create average annual savings of US$ 79 942 per patient.

F. WHO or other clinical guidelines relevant to the test

Several international guidelines recommend measuring immunoglobulin and immunofixation to diagnose myeloma and other conditions. For example, NICE recommend serum PEP to confirm the presence of a paraprotein. The same applies to diagnosing an immunodeficiency, for example, in guidelines from both the European Society of Immunodeficiencies (*9*) and the International Patient Organisation for Primary Immunodeficiencies (IPOPI) (*3*).

The IPOPI guidelines (*3*) on principles of care for PIDs identify the standardization of immunological diagnostic protocols, including protein analyses, immunophenotypes and in vivo and in vitro functional tests, as one of the six criteria for fast and reliable PID diagnoses.

A best practice article on paraprotein management, published in the *British Medical Journal* (*10*), illustrates the use of electrophoresis in diagnosing and monitoring plasma cell dyscrasias.

A 2009 article by Chapel and Cunningham-Rundles (*11*) provides an update in understanding and managing CVIDs and identifies the need for two relevant investigations:

- serum immunoglobulin levels and electrophoresis: full blood count, renal and liver function tests, including albumin for loss, serum IgG antibodies to exposure and immunisation antigens, and lymphocyte (T, B and NK cells) and B memory (isotype switched and naive); and

- urinary PEP: any investigations for specific diseases, for example fungal or mycobacterial antigens (if granuloma), TSH (if excessive fatigue), faecal fat (if malabsorption), sweat test (to exclude cystic fibrosis) and baseline respiratory tests, including CT.

Spickett et al. (*12*) write about the diagnosis and management of primary antibody deficiency by consultant immunologists in the United Kingdom.

Consensus recommendations by the Asia Pacific Immunoglobulins in Immunology Expert Group (*13*) on the use of immunoglobulin replacement therapy in PIDs list serum, urinary and faecal proteins as a necessary investigation for other causes of low immunoglobulin levels, including multiple myeloma, protein-losing enteropathy and nephrotic syndrome. Serum and urinary PEP should be performed in adults with reduced levels of IgG or IgM to exclude the presence of a paraprotein.

The International Union of Immune Societies PID Expert Group have guidelines (*14*) for clinicians to classify and diagnose PID at the bedside.

G. Basic test characteristics (from application)

Test formats available	Gel electrophoresis
Specimen types	Serum, urine
Equipment required	Electrophoresis instrument
Regulatory status	None
Availability	Worldwide
Price per test range	Around £5 in the United Kingdom; US$ 1.66 in India
Instrument price range	Not provided

H. Evidence for diagnostic accuracy (from application)

Serum paraproteins can be accurately and precisely measured by means of agarose gel electrophoresis and densitometry at 520 nm after staining with Coomassie Brilliant Blue G. The results obtained for this method agree closely with the quantity of paraprotein recoverable from preparative agarose gel electrophoresis (*15–17*).

CVIDs are a group of diseases in which failure to produce gammaglobulins (encompassing immunoglobulin) and protective antibodies results in symptoms, usually but not always, of recurrent bacterial infections. The common feature is lack of antibodies to pathogens. Differential diagnosis of patients with recurrent bacterial infections includes lymphoid malignancies and loss of immunoglobulin

due to renal/gut disease. Investigations in patients with suspected CVIDs include (among other tests) serum and urinary PEP (*18*–*20*).

Serum and urinary PEP exhibits good characteristics of workability, accuracy, precision and reliability. It can be recommended as a qualitative screening procedure for detecting abnormalities of the major proteins and as an important adjunct to specific quantitation of serum proteins (*21*).

I. Evidence for clinical usefulness and impact (from application)

The clinical immunology laboratory is a powerful adjunct to the clinician in the initial evaluation of immunodeficiency disorders. Clinicians order specific laboratory tests and interpret laboratory data that are useful in establishing a diagnosis of inherited (primary) or acquired (secondary) immunodeficiencies. The clinical immunologic evaluation of patients for immunodeficiency proceeds in an orderly fashion, from screening tests to sophisticated tests (*9*, *22*).

If serum immunoglobulin levels are abnormally high in a patient with repeated infection, particularly an adult, it may be important to perform serum PEP to determine whether a paraprotein is present in the form of a monoclonal spike as evidence of a lymphocyte malignancy (*23*).

In certain HICs, clinicians still have to deliberately request IgG, IgM and IgA levels and electrophoresis as separate tests to meaningfully assess serum immunoglobulin, whereas in other countries it is automatically recognized that these tests are inseparable and absolutely required to effectively interpret results.

In the 1980s, reports emerged about the unacceptable delay in diagnosis between the onset of recurrent infections and the recognition of immunodeficiency and the start of treatment. In 1995, consensus guidelines for general practitioners, physicians and paediatricians were jointly published by the Royal College of Physicians, the Royal College of Pathologists and the Primary Immunodeficiency Association in an attempt to facilitate early recognition of these disorders. In 2002 the diagnostic delay for primary antibody deficiency was reassessed, and the authors were pleased to report an improvement in the mean diagnostic delay compared to the 1989 study, but disappointment that the average diagnostic delay was 4·4 years (median 2 years). An audit of the prevalence of PID has shown that regions with no clinically led immunology service fail to recognize significant immunodeficiencies, with subsequent patient illness and death.

A series of basic screening tests are used which allow rapid decision-making regarding the need for further analysis. This is an important step in improving the outcome for PID patients by initiating suitable screening investigations at an early stage of presentation, as previous approaches were often overly complex for mainstream physicians and non-immunologists.

J. Evidence for economic impact and/or cost–effectiveness (from application)

Little comparative data on cost–effectiveness is available for these rare diseases.

K. Ethics, equity and human rights issues (from application)

None quoted.

L. Summary of evidence evaluation

In general, based on this portfolio, it can be recommended that serum PEP can be used for diagnosis of primary immunodeficiencies in situations where other diagnostic facilities are lacking. Serum PEP will not replace immunoglobulin level detection but is a good alternative for the diagnostic pathway in a PID-suspected patient when other tests cannot be done. No evidence was provided on the diagnostic accuracy measures of serum PEP. But guidelines recommend their use in cases where other testing facilities are lacking.

Early diagnosis shortens diagnostic delay that is distressing to the family, damaging to the patient and wasteful of health care resources. In general, when PIDs are left undiagnosed or are misdiagnosed, chronic illness and disability take a heavy toll on health care resources.

M. Summary of SAGE IVD deliberations

This is a resubmission of an application that was rejected last year for insufficient evidence.

After careful review and discussion, SAGE IVD members agreed that key evidence is still lacking. In particular, no evidence was provided on the diagnostic accuracy of serum PEP for PIDs. And SAGE IVD members noted that while serum PEP can probably detect deficiencies in IgG fairly reliably, it is much less accurate at detecting IgA or IgM deficiencies.

The group acknowledged that, despite the lack of primary studies on diagnostic accuracy, there does appear to be some agreement among professional societies that even though serum PEP cannot replace immunoglobulin detection, it may be a good alternative to use in the diagnostic pathway in PID-suspected patients if there are no other tests available. To that end, guidelines recommend using serum PEP to diagnose PIDs only in situations where other diagnostic facilities are lacking.

N. SAGE IVD recommendations

SAGE IVD recommended excluding the serum and urine PEP test category for diagnosing PIDs from the third EDL.

O. References

1. Jiang F, Torgerson TR, Ayars AG. Health-related quality of life in patients with primary immunodeficiency disease. Allergy Asthma Clin Immunol. 2015;11:27. doi:10.1186/s13223-015-0092-y.
2. Kobrynski L, Powell RW, Bowen S. Prevalence and morbidity of primary immunodeficiency diseases, United States 2001–2007. J Clin Immunol. 2014;34(8):954–961. doi:10.1007/s10875-014-0102-8.
3. Chapel H, Prevot J, Gaspar HB, Español T, Bonilla FA, et al. Primary immune deficiencies – principles of care. Front Immunol. 2014;15;5:627. doi:10.3389/fimmu.2014.00627.
4. Lindegren ML, Kobrynski L, Rasmussen SA, Moore CA, Grosse SD, et al. Applying public health strategies to primary immunodeficiency diseases: a potential approach to genetic disorders. CDC MMWR. 2004;53(RR01);1–29.
5. Sewell WAC, Khan S, Doré PC. Early indicators of immunodeficiency in adults and children: protocols for screening for primary immunological defects. Clin Exp Immunol. 2006;145(2):201–203. doi:10.1111/j.1365-2249.2006.03126.x.
6. Aghamohammadi A, Moin M, Farhoudi A, Rezaei N, Pourpak Z, et al. Efficacy of intravenous immunoglobulin on the prevalence of pneumonia in patients with agammaglobulinaemia. FEMS Immunol Med Microbiol. 2004;40(2):113–118. doi:10.1016/S0928-8244(03)00304-3.
7. Modell V, Orange JS, Quinn J, Modell F. Global report on primary immunodeficiencies: 2018 update from the Jeffrey Modell Centers Network on disease classification, regional trends, treatment modalities and physician reported outcomes. Immunol Res. 2018;66:367–380. doi:10.1007/s12026-018-8996-5.
8. Condino-Neto A, Espinosa-Rosales FJ. Changing the lives of people with primary immunodeficiencies (PI) with early testing and diagnosis. Front Immunol. 2018;9:1439. doi:10.3389/fimmu.2018.01439.
9. de Vries, E. Clinical Working Party of the European Society for Immunodeficiencies (ESID). Patient-centred screening for primary immunodeficiency: a multi-stage diagnostic protocol designed for non-immunologists. Clin Exp Immunol. 2006;145:204–214. doi:10.1111/j.1365-2249.2006.03138.x.
10. Smellie WS, Spickett GP. Paraprotein management. BMJ. 2006;333(7560):185–187. doi:10.1136/bmj.333.7560.185.
11. Chapel H, Cunningham-Rundles C. Update in understanding common variable immunodeficiency disorders (CVIDs) and the management of patients with these conditions. Br J Haematol. 2009;145(6):709–727. doi:10.1111/j.1365-2141.2009.07669.x.
12. Spickett GP, Askew T, Chapel H. Management of primary antibody deficiency by consultant immunologists in the United Kingdom: a paradigm for other rare diseases. Qual Health Care. 1995;4(4):263–268. doi:10.1136/qshc.4.4.263.
13. Consensus recommendations for the use of immunoglobulin replacement therapy in immune deficiency, 2nd edition. Parkville: Asia Pacific Immunoglobulins in Immunology Expert Group; 2009 (http://www.apiieg.org/files/1/APIIEG%20Consensus%20Recommendations%20Edition%201%20June%202008.pdf, accessed 2 April 2020).
14. Bousifha AA, Jeddane L, Ailal F, Al Herz W, Conley ME, et al. A phenotypic approach for IUIS PID classification and diagnosis: guidelines for clinicians at the bedside. J Clin Immunol. 2013;33(6):1078–1087. doi:10.1007/s10875-013-9901-6.
15. Guinan JE, Kenny DF, Gatenby PA. Detection and typing of paraproteins: a comparison between different methods in a routine diagnostic laboratory. Pathology. 1989;21(1):35–41. doi:10.3109/00313028909059528.

16. Keren DF. High-resolution electrophoresis and immunofixation, 2nd edition. Woburn: Butterworh-Heinmann; 1994.
17. Dennis PM, Biegler B, Papas R. Improved measurement of monoclonal paraproteins in serum using agarose gel electrophoresis. Ann Clin Biochem. 1987;24:73–76. doi:10.1177/000456328702400111.
18. Rikhi R, Bhattad S, Jindal A, Saikia B, Garg R, et al. Monoclonal gammopathy of unclear significance in a child with Wiskott-Aldrich syndrome: a rare occurrence. J Clin Immunol. 2019;39(1):7–10. doi:10.1007/s10875-018-0585-9.
19. Sanchez LA, Maggadottir SM, Pantell MS, Lugar P, Cunningham Rundles C, et al. Two sides of the same coin: pediatric-onset and adult-onset common variable immune deficiency. J Clin Immunol. 2017;37(6):592–602. doi:10.1007/s10875-017-0415-5.
20. Westh L, Mogensen TH, Dalgaard LS, Bernth Jensen JM, Katzenstein T, et al. Identification and characterization of a nationwide Danish adult common variable immunodeficiency cohort. Scand J Immunol. 2017;85(6):450–461. doi:10.1111/sji.12551.
21. Jeppson JO, Laurell CB, Franzén B. Agarose gel electrophoresis. Clin Chem. 1979;25:629–38.
22. Folds JD, Schmitz JL. Clinical and laboratory assessment of immunity. J Allergy Clin Immunol. 2003;111(Suppl. 2):S702–11. doi:10.1067/mai.2003.122.
23. Bonilla FA, Bernstein IL, Khan DA, Ballas ZK, Chinen J, et al. Practice parameter for the diagnosis and management of primary immunodeficiency. Ann Allergy Asthma Immunol. 2005;94 (5 Suppl 1):S1–63. doi:10.1016/s1081-1206(10)61142-8.

STREPTOCOCCAL PHARYNGITIS

3.21. Group A *Streptococcus* nucleic acid test

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a GAS NAT to the EDL as an IVD to diagnose streptococcal pharyngitis.

B. Applicant

World Health Organization

C. WHO technical department

Antimicrobial Resistance

D. Background (from application)

Disease condition and impact on patients

Streptococcal pharyngitis, also known as strep throat, is an infection of the back of the throat including the tonsils caused by GAS (*Streptococcus pyogenes*).

Common symptoms include fever, sore throat, red tonsils, and enlarged lymph nodes in the neck. A headache, abdominal pain, and nausea or vomiting may also occur. Symptoms typically begin 1–3 days after exposure and last 7–10 days. Complications that may arise if strep throat is not correctly diagnosed and managed include rheumatic fever with potential long-term heart damage or death, kidney inflammation and peritonsillar abscess, sepsis and necrotizing fasciitis.

Does the test meet a medical need?

Inappropriate antibiotic prescription is a major issue in many countries. The public health impact of the GAS NAT lies in assisting health care providers to use antibiotics appropriately for treating pharyngitis. The gold standard test for diagnosing strep throat, alongside clinical signs and symptoms, is a throat swab culture. But this takes 24–48 hours to report back, while a NAT gives a result in 5–20 minutes.

How the test is used

The test is used to diagnose streptococcal pharyngitis in the primary care setting or hospital POC setting, for example the emergency room. Although the latest guidelines do not reflect the use of a molecular test as a stand-alone test for detecting streptococcal pharyngitis in children and adults, the commercially available tests are approved for this application without needing to confirm with culture.

The typical sensitivity of a NAT is more than 95%; the newest POC tests are more than 98% sensitive. This means that confirmation of negative results by culture is no longer required and all patients who receive the test can be diagnosed within 5–20 minutes. In this way, all patients who need antibiotics will receive them in a more timely way compared with RADTs. That said, a small false-positive rate has been demonstrated which could lead to some patients getting antibiotics who do not need them. This, however, is a small group when considered against the large number who would have received unnecessary antibiotic treatment without the use of a rapid test (either RADT or nucleic acid) and is also small compared to the group that would have to wait for bacterial culture confirmation when RADTs are used.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

Strep throat is a common bacterial infection in children aged 5–15 years and is rare in children under 3 years of age. It is the cause of 15–40% of sore throats among children and 5–15% among adults. Cases are more common in late winter and early spring.

A lack of reliable data makes accurately estimating global disease burden difficult. A 2005 review estimated that at least 18.1 million people suffered from invasive GAS diseases, with another 1.78 million incident cases occurring each year (*1*). These estimates did not include the more than 111 million cases of streptococcal pyoderma and 616 million cases of GAS pharyngitis each year.

The health, economic and social burden of GAS pharyngitis (and skin infections) can be significant despite the fact that these diseases are relatively benign. Direct health costs come from antibiotic use, as well as missed schooldays for children and workdays for their parents; but their causal association with invasive infection also has clinical and public health implications. A small 2008 study in Boston, USA, estimated that the total cost of GAS pharyngitis among children in the USA ranges from US$ 224 to US$ 539 million per year, half of which is non-medical costs related to school- and workdays missed (*2*).

By far the heaviest burden of GAS infections falls on LMICs, because children living in crowded or unsanitary conditions are at much higher risk not only of contracting the benign forms of infections but also of these not being diagnosed correctly and progressing to invasive disease (*3*).

F. WHO or other clinical guidelines relevant to the test

No guidelines currently recommend use of the GAS NAT test. In its 2012 guidelines, IDSA does not mention using NATs to diagnose GAS pharyngitis. Few commercial tests were available when the guidelines were developed.

G. Basic test characteristics (from application)

Test format	Nucleic acid amplification test (NAT)
Specimen types	Throat swab
Equipment required	POC tests: small benchtop instruments (dedicated closed systems for manufacturers' test kits) Laboratory-based tests: molecular diagnostic analysers (dedicated closed systems for manufacturers' test kits)
Regulatory status	CE-marked and FDA-approved
Availability	Global
Price per test range	~US$ 25
Instrument price range	~US$ 12 000–50 000

H. Evidence for diagnostic accuracy (from application)

A report compiled by CADTH (*4*) reviewed 11 observational studies evaluating molecular tests. It found that the sensitivity of molecular assays (with culture assays as the reference test) on children or mixed populations of children and adults generally varied between 93% and 100% (with the exception of one study showing a sensitivity of 82%); the specificity varied between 91% and 99%.

Parker et al. (*5*) compared three NATs and a RADT with culture and concluded that NATs have the highest sensitivity and rapid turnaround time. The three NATs had sensitivity and specificity ranging from 95.2% to 100% and 97.4% to 100%, respectively; while positive and negative predictive values ranged from 96.7% to 100% and 95.5% to 100%, respectively.

I. Evidence for clinical usefulness and impact (from application)

Rao et al. (*6*) demonstrated utility in a paediatric primary care setting among 275 patients aged 3–18 years in the USA. They concluded that POC NATs for GAS had higher sensitivity and specificity than other available tests and resulted in appropriate antibiotic prescriptions.

Ralph et al. (*7*) did an evaluation in a hospital in a high-burden setting in northern Australia and showed high sensitivity but lower specificity (79%). They concluded that the test increases appropriate antibiotic use, although its potential impact on GAS complications (acute rheumatic fever) needs further studies.

J. Evidence for economic impact and/or cost–effectiveness (from application)

None identified.

K. Ethics, equity and human rights issues (from application)

None identified.

L. Summary of evidence evaluation

There currently is no systematic summary of the accuracy of molecular tests for GAS. Evidence on the accuracy of molecular tests is available from several studies; it suggests that they are more sensitive than antigen-based rapid tests (sensitivity > 95%), although results are variable. Studies also show that molecular tests have specificities greater than 90% based on questionable use of a second molecular test in discrepant analysis. The possible bias that this may have introduced is to overestimate specificity by up to 5%.

There is some evidence on how molecular tests impact antibiotic prescribing, but no evidence on whether molecular tests lead to better patient outcomes than treatment based on rapid antigen tests or culture.

M. Summary of SAGE IVD deliberations

GAS pharyngitis is a common and severe condition, especially among children. Prompt and adequate treatment improve the prognosis of GAS pharyngitis, so rapid and robust diagnosis is a requisite of children's primary care worldwide.

A major disadvantage of the test is that it does not form part of current international guidelines (although SAGE IVD noted that these have not been updated since 2011).

The submission also lacked key evidence on the test's performance and utility. There was no systematic review on the accuracy of GAS NAT in the submission. (Although evidence from a range of other studies suggests that the test has superior accuracy to the antigen-based rapid tests and can serve as a single confirmatory test.) No evidence was available on whether GAS NATs lead to better patient outcomes than treatment based on rapid antigen tests or culture. And the submission included little evidence of the test's utility in reducing use of antibiotics (compared with general evidence for rapid tests, which are proven to reduce antibiotic prescribing by about a third). Some SAGE IVD members argued that the test can still be of value in avoiding the unnecessary use of antibiotics in LMICs where culture facilities are not readily available. But others raised concerns about overdiagnosis, arguing that because GAS NAT has a very low limit of detection it can result in more false positives and unnecessary treatment. In short, the group agreed that the test's impact on antibiotic use remains very uncertain.

In considering POC applications of the test, the group acknowledged the importance of having a quick, reliable answer close to the patient; but they also noted that the GAS NAT is costly and may not be appropriate for low-resource settings.

N. SAGE IVD recommendations

SAGE IVD recommended excluding the GAS NAT test category as an IVD for diagnosing streptococcal pharyngitis from the third EDL.

Acknowledging that the current guidelines are nearly a decade old, however, the group further recommended that the EDL Secretariat reach out to appropriate societies to establish whether there are any plans to update the guidelines that may warrant a resubmission (if accompanied by further evidence on impact on antibiotic use).

O. References

1. The current evidence for the burden of group A streptococcal diseases. Geneva: World Health Organization; 2005 (WHO/FCH/CAH/05.07).
2. Pfoh E, Wessels MR, Goldmann D, Lee GM. Burden and economic cost of group A streptococcal pharyngitis. Pediatrics. 2008;121(2):229–234. doi:10.1542/peds.2007-0484.

3. Sims Sanyahumbi A, Colquhoun S, Wyber R, Carapetis JR. Global disease burden of group A streptococcus. In: Ferretti JJ, Stevens DL, Fischetti VA, editors. Streptococcus pyogenes: basic biology to clinical manifestations. Oklahoma City: University of Oklahoma Health Sciences Center; 2016.
4. Rapid tests for the diagnosis of group A streptococcal infection: a review of diagnostic test accuracy, clinical utility, safety, and cost effectiveness. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2018.
5. Parker KG, Gandra S, Matushek S, Beavis KG, Tesic V, et al. Comparison of 3 nucleic acid amplification tests and a rapid antigen test with culture for the detection of group A streptococci from throat swabs. J Appl Lab Med. 2019;4(2):164–169. doi:10.1373/jalm.2018.028696.
6. Rao A, Berg B, Quezada T, Fader R, Walker K, et al. Diagnosis and antibiotic treatment of group a streptococcal pharyngitis in children in a primary care setting: impact of point-of-care polymerase chain reaction. BMC Pediatr. 2019;19(1):24. doi:10.1186/s12887-019-1393-y.
7. Ralph AP, Holt DC, Islam S, Osowicki J, Carroll DE, et al. Potential for molecular testing for group A streptococcus to improve diagnosis and management in a high-risk population: a prospective study. Open Forum Infect Dis. 2019;6(4):ofz097. doi:10.1093/ofid/ofz097.

VACCINE-PREVENTABLE DISEASES

3.22. Measles immunoglobulin G antibody

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a measles IgG antibody test to the EDL as an IVD to aid in the diagnosis of clinically suspected measles infection.

B. Applicant

World Health Organization

C. WHO technical department

Immunization Analysis & Insights

D. Background (from application)

Disease condition and impact on patients

Measles, an acute illness caused by a virus in the family Paramyxoviridae, genus *Morbillivirus*, is one of the most highly infectious diseases known to humankind. Symptoms include fever (as high as 38 °C) and malaise, cough, coryza and conjunctivitis, followed by a maculopapular rash (*1*). Measles is generally a mild or moderately severe illness but can, if not diagnosed or treated, result

in complications such as pneumonia, encephalitis and death. A rare long-term sequela of measles virus infection is subacute sclerosing panencephalitis (SSPE), which is a fatal disease of the CNS that generally develops 7–10 years after infection (*1*, *2*). Measles case fatality rates vary from 0.1% in HICs to 15% in LMICs (*2*).

Does the test meet a medical need?

Accurate diagnostic tests for measles infection are essential to confirm cases and outbreaks.

A single laboratory-confirmed measles case should trigger an aggressive public health investigation and response in an elimination setting (*3*).

How the test is used

IgG seroconversion testing is used to confirm IgM-positive tests in elimination settings where incidence of measles is low. Two serum samples must be collected: the first (acute) no later than 7 days from rash onset; and the second (convalescent) 10–28 days later. A significant rise in IgG levels indicates a positive result.

Laboratory case confirmation for measles includes diagnostically significant titre change in IgG antibody level in acute or convalescent sera, or documented seroconversion (IgG negative to IgG positive).

E. Public health relevance (from application)

Prevalence

Significant gains towards measles elimination have been made with a highly effective measles vaccine. Countries in all six WHO regions have adopted measles elimination goals, and four WHO regions endorsed the Global Vaccine Action Plan to eliminate measles by 2015. Not all of the plan's goals were accomplished, but measles was successfully eliminated through comprehensive vaccination and surveillance in 61 Member States in the Region of the Americas, and the European and Western Pacific regions (*2*).

Still, in many parts of the world, the disease remains endemic. In 2015, there were 254 928 cases reported and an estimated 134 200 measles deaths globally (*1*). In 2019, the greatest number of cases were reported from the Indian subcontinent, where annual incidence rates exceeded 50 per million population. Other areas of Africa, Asia, Europe, Central and South America, and the Pacific also have large annual numbers of measles cases (*4*).

"Vaccine hesitancy" remains an obstacle to measles elimination and has been identified by WHO as one of the major threats to global health in 2019. International travel also allows measles "importation" from endemic countries and has contributed to a resurgence of the disease in high-income countries and countries previously thought to be free of the disease (*4*). For example, in

the USA, even though measles was officially eliminated in 2000, occasional outbreaks (three or more linked cases) are still reported to the CDC each year, often imported from overseas and spreading among communities with low rates of immunization (*4*).

Socioeconomic impact

Measles outbreaks can place a heavy economic burden on local and state public health institutions. USA outbreaks in 2011 cost an estimated US$ 2.7–5.3 million (*5*); 2015 outbreaks cost between US$ 0.25 million and US$ 1.35 million (*6*). Estimates from the 2013 outbreak in New York City cost the city's Department of Health and Mental Hygiene US$ 400 000 (*7*). And in 2016, it was estimated that each case of measles cost the public sector US$ 20 000 (*8*).

Outside the USA, total costs associated with measles in the Netherlands from 2013 to 2014 were estimated at US$ 4.7 million (*9*); in Italy outbreaks in 2002 and 2003 cost between €17.6 million and €22 million, respectively (*10*). Total societal costs from outbreaks in Romania in 2011 were estimated at US$ 5.5 million (*11*).

In addition, a study of 3207 lab-confirmed measles cases reported by Public Health England from January 2012 to June 2013 resulted in an estimated loss of 44.2 QALYs (*12*).

F. WHO or other clinical guidelines relevant to the test

The 2018 WHO manual for laboratory-based surveillance of measles, rubella and congenital rubella syndrome (*13*) states that measuring measles-specific IgG can be a useful additional serologic method for case classification when an equivocal result is obtained for IgM, or when a positive IgM result is questioned due to clinical or epidemiologic information that is inconsistent with a case of measles. A significant rise in IgG titre in convalescent sera from suspected cases confirms a positive IgM result.

The 2018 WHO surveillance standards (*3*) also state that IgG rise in titre between acute and convalescent sera can be used as to aid measles diagnosis.

G. Basic test characteristics (from application)

Test formats available	EIA
Specimen types	Serum, plasma, DBS, oral fluid
Equipment required	Refrigerator, calibrated pipettes, vortex, ELISA washer and reader
Regulatory status	CE-marked

Table *continued*

Availability	Global
Price per test range	Not available
Instrument price range	Not available

H. Evidence for diagnostic accuracy (from application)

No systematic reviews of measles IgG test clinical accuracy were available. One study by Dina et al. (*14*) compared three commercial assays in three different types of preselected subjects: clinical measles suspects, possible cases (IgM negative) and general population. All clinical suspects were found to be IgG positive by all methods, as were the possible cases.

I. Evidence for clinical usefulness and impact (from application)

Mosquera et al. (*15*) show the value of measles IgG combined with RT-PCR in further classifying IgM-negative cases detected during an outbreak.

J. Evidence for economic impact and/or cost–effectiveness (from application)

No data available.

K. Ethics, equity and human rights issues (from application)

None identified.

L. Summary of evidence evaluation

As an IgG response is part of the standard confirmation of a measles case, it is difficult to assess test accuracy; inevitably, the reference standard used will incorporate an IgG test in some cases. There are no studies comparing the results of this test against an independent reference standard (given that it is effectively part of the case definition). The test is embedded in WHO protocols for testing measles.

Evidence of the value of the IgG test could be better established by looking at the numbers of measles cases that were diagnosed by an IgG when other methods failed. The study from the Spanish outbreak indicates that a small number of cases were only identified by IgG testing following a negative IgM test.

Some evidence exists showing that there are measles cases detected only by use of the IgG test, particularly showing reinfection or infection in the vaccinated population.

M. Summary of SAGE IVD deliberations

Measles is a disease of public health concern, and it is important to detect outbreaks early. In clinical settings the measles IgG test is mostly used in combination with IgM or PCR tests to confirm cases. A single positive or negative result of measles IgG does not rule out the presence or absence of measles infection; so it is not diagnostic of measles on its own.

The test is known to be useful in diagnosing some complications of measles, such as SSPE, and in determining immune status in specific groups, including transplant patients.

It is also known to be a useful test for determining the immunity of a person pre- and post-vaccination, and thus could be used to study seroprevalence in a population. But the test was not submitted as a screening tool for immunization; moreover, the WHO representative confirmed that IgG is not recommended for seroprevalence.

Nevertheless, the IgG test is part of the case definition for measles and forms part of the WHO algorithm for diagnosing the disease in an elimination setting. In part because of this, there is not much data on the test's clinical accuracy or performance. And very few data were submitted on the test's usefulness. SAGE IVD noted a definite need for more evidence on how the measles IgG test has made a difference to show that it is an essential test.

Other limitations of the test include the fact that it requires paired sera and cannot be used as a POC test. It requires a laboratory with ambient environmental conditions and a regular power supply; it may also prove costly in LMICs and resource-constrained settings.

SAGE IVD noted that the application listed an EIA format for the test, but highlighted the existence of avidity assays, which could also be useful in diagnosing acute measles.

N. SAGE IVD recommendations

SAGE IVD recommended conditionally including the measles IgG antibody test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, Vaccine-preventable diseases);
- using an immunoassay format;
- to aid in the diagnosis of clinically suspected measles infection;

pending the submission of further evidence to show that the test has utility in confirming cases when used with IgM.

O. References

1. Gastanaduy PA, Redd SB, Clemmons NS, Lee AD, Hickman CJ, et al. Chapter 7: Measles. In: Roush SW, Baldy LM, Kirkconnell Hall MA, editors. Manual for the surveillance of vaccine-preventable diseases. Atlanta: US Centers for Disease Control and Prevention; 2014.
2. Orenstein WA, Cairns L, Hinman A, Nkowane B, Olivé JM, et al. Measles and rubella global strategic plan 2012–2020 midterm review report: background and summary. Vaccine. 2018;36(Suppl 1):A35–A42. doi:10.1016/j.vaccine.2017.10.065.
3. Surveillance standards for vaccine-preventable diseases, 2nd edition. Geneva: World Health Organization; 2018.
4. O'Donnell S, Davies F, Vardhan M, Nee P. Could this be measles? Emerg Med J. 2019;36:310–314. doi:10.1136/emermed-2019-208490.
5. Ortega-Sanchez IR, Vijayaraghavan SM, Barskey AE, Wallace GS. The economic burden of sixteen measles outbreaks on United States public health departments in 2011. Vaccine. 2014;32(11):1311–1317. doi:10.1016/j.vaccine.2013.10.012.
6. Ozawa S, Portnoy A, Getaneh H, Clark S, Knoll M, et al. Modeling the economic burden of adult vaccine-preventable diseases in the United States. Health Aff. 2016;11:2124–2213. doi:10.1377/hlthaff.2016.0462.
7. Rosen JB, Arciuolo RJ, Khawja AM, Fu J, Giancotti FR, et al. Public health consequences of a 2013 measles outbreak in New York City. JAMA Pediatr. 2018;172(9):811–817. doi:10.1001/jamapediatrics.2018.1024.
8. Lo NC, Hotez PJ. Public health and economic consequences of vaccine hesitancy for measles in the United States. JAMA Pediatr. 2017;171(9):887–892. doi:10.1001/jamapediatrics.2017.1695.
9. Suijkerbuijk AWM, Woudenberg T, Hahné SJM, Lochlainn LN, de Melker HE, et al. Economic costs of measles outbreak in the Netherlands, 2013–2014. Emerg Infect Dis. 2015;21(11):2067–2069. doi:10.3201/eid2111.150410.
10. Filia A, Brenna A, Panà A, Cavallaro GM, Massari M, et al. Health burden and economic impact of measles-related hospitalizations in Italy in 2002–2003. BMC Public Health. 2007;7:169. doi:10.1186/1471-2458-7-169.
11. Njau J, Janta D, Stanescu A, Pallas SS, Pistol A, Khetsuriani N, et al. Assessment of economic burden of concurrent measles and rubella outbreaks, Romania, 2011–2012. Emerg Infect Dis. 2019;25(6):1101–1109. doi:10.3201/eid2506.180339.
12. Thorrington D, Ramsay M, van Hoek AJ, Edmunds WJ, Vivancos R, et al. The effect of measles on health-related quality of life: a patient based survey. PloS One. 2014;9(9):e105153. doi:10.1371/journal.pone.0105153.
13. Manual for the laboratory-based surveillance of measles, rubella, and congenital rubella syndrome, 3rd edition. Geneva: World Health Organization; 2018.
14. Dina J, Creveuil C, Gouarin S, Viron F, Hebert A, et al. Performance evaluation of the VIDAS® measles IgG assay and its diagnostic value for measuring IgG antibody avidity in measles virus infection. Viruses. 2016;8(8):234. doi:10.3390/v8080234.
15. Mosquera MM, de Ory F, Gallardo V, Cuenca L, Morales M, et al. Evaluation of diagnostic markers for measles virus infection in the context of an outbreak in Spain. J Clin Microbiol. 2005;43(10):5117–5121. doi:10.1128/JCM.43.10.5117-5121.2005.

3.23. Measles immunoglobulin M antibody

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a measles immunoglobulin M (IgM) antibody test to the EDL as an IVD to diagnose, or aid in the diagnosis of, clinically suspected measles infection.

B. Applicant

World Health Organization

C. WHO technical department

Immunization Analysis & Insights

D. Background (from application)

Disease condition and impact on patients

Measles, an acute illness caused by a virus in the family Paramyxoviridae, genus *Morbillivirus*, is one of the most highly infectious diseases known to humankind. Symptoms include fever (as high as 38 °C) and malaise, cough, coryza and conjunctivitis, followed by a maculopapular rash (*1*).

Measles is generally a mild or moderately severe illness but can, if not diagnosed or treated, result in complications such as pneumonia, encephalitis and death. A rare long-term sequela of measles virus infection is SSPE, which is a fatal disease of the CNS that generally develops 7–10 years after infection (*1*, *2*). Measles case fatality rates vary from 0.1% in HICs to 15% in LMICs (*2*).

Does the test meet a medical need?

Accurate diagnostic tests for measles infection are essential to confirm cases and outbreaks.

A single laboratory-confirmed measles case should trigger an aggressive public health investigation and response in an elimination setting (*3*).

How the test is used

IgM testing requires the collection of serum samples, ideally between 4 and 28 days from rash onset. In elimination settings, or if the patient was vaccinated for measles within the past 6 weeks, confirmation of an IgM-positive or equivocal

test is required, by repeat IgM testing (10–28 days after first test); seroconversion testing with IgG (where the first, acute, sample is collected no later than seven days from rash onset and the second, convalescent, sample is collected 10–28 days later); RT-PCR testing; or through clinical evidence or an epidemiological link.

If samples are collected at the ideal times, an IgM-negative result rules out infection. If the sample was collected within 3 days of rash onset, IgM-negative results are confirmed with RT-PCR, and repeat IgM testing (≥ 6 days after onset of rash) is advised if the case remains suspicious for measles.

Laboratory case confirmation for measles can yield the following test results:

- Detection of anti-measles IgM antibody by EIA. This is the gold standard.
- Diagnostically significant titre change in IgG antibody level in acute or convalescent sera, or documented seroconversion (IgG negative to IgG positive), positive RT-PCR or viral isolation in cell culture.

E. Public health relevance (from application)

Prevalence

Significant gains towards measles elimination have been made with a highly effective measles vaccine. Countries in all six WHO regions have adopted measles elimination goals, and four WHO regions endorsed the Global Vaccine Action Plan to eliminate measles by 2015. Not all of the plan's goals were accomplished, but measles was successfully eliminated through comprehensive vaccination and surveillance in 61 Member States in the Region of the Americas, and the European and Western Pacific regions (*2*).

Still, in many parts of the world, the disease remains endemic. In 2015, there were 254 928 cases reported and an estimated 134 200 measles deaths globally (*1*). In 2019, the greatest number of cases were reported from the Indian subcontinent, where annual incidence rates exceeded 50 per million population. Other areas of Africa, Asia, Europe, Central and South America, and the Pacific also have large annual numbers of measles cases (*4*).

Vaccine hesitancy remains an obstacle to measles elimination and has been identified by WHO as one of the major threats to global health in 2019. International travel also allows measles importation from endemic countries and has contributed to a resurgence of the disease in high-income countries and countries previously thought to be free of the disease (*4*). For example, in the USA, even though measles was officially eliminated in 2000, occasional outbreaks (three or more linked cases) are still reported to the CDC every year, often imported from overseas and spreading among communities with low rates of immunization (*4*).

Socioeconomic impact

Measles outbreaks can place a heavy economic burden on local and state public health institutions. USA outbreaks in 2011 cost an estimated US$ 2.7–5.3 million (*5*); 2015 outbreaks cost between US$ 0.25 million and US$ 1.35 million (*6*). Estimates from the 2013 outbreak in New York City cost the city's Department of Health and Mental Hygiene US$ 400 000 (*7*). And in 2016, it was estimated that each case of measles cost the public sector US$ 20 000 (*8*).

Outside the USA, total costs associated with measles in the Netherlands from 2013 to 2014 were estimated at US$ 4.7 million (*9*); in Italy outbreaks in 2002 and 2003 cost between €17.6 million and €22 million, respectively (*10*). Total societal costs from outbreaks in Romania in 2011 were estimated at US$ 5.5 million (*11*).

In addition, a study of 3207 laboratory-confirmed measles cases reported by Public Health England from January 2012 to June 2013 resulted in an estimated loss of 44.2 QALYs (*12*).

F. WHO or other clinical guidelines relevant to the test

The 2018 WHO manual for laboratory-based surveillance of measles, rubella, and congenital rubella syndrome (*13*) recommends routine testing of all suspected cases of measles or rubella for both measles and rubella IgM, as well as initial testing of all samples for measles IgM. Detection of measles IgM in a single serum specimen is the standard method for rapid laboratory confirmation of measles.

The 2018 WHO surveillance standards for vaccine-preventable diseases (*3*) also recommend using measles IgM for laboratory confirmation and stipulate that results of IgM should be reported within 4 days of the specimen's arrival to the laboratory.

G. Basic test characteristics (from application)

Test formats available	EIA
Specimen types	Serum, plasma, DBS, oral fluid
Equipment required	Refrigerator, calibrated pipettes, vortex ELISA washer and reader for manual ELISA tests
Regulatory status	Approved by stringent regulatory authorities
Availability	Global
Price per test range	Not available
Instrument price range	Not available

H. Evidence for diagnostic accuracy (from application)

No systematic reviews of IgM test clinical accuracy were available. Four primary studies (*14–17*) were reviewed. All studies concluded that use of IgM alone does not reach 100% sensitivity for diagnosis of acute cases. The sensitivity and specificity ranges reached in these studies were:

- Bolotin et al. (*14*): 79.2% and 65.7%, respectively (21 299 tests in elimination setting of Canada).
- Tipples et al. (*15*): 87.9–96.7% and 94.6–98.7%, respectively.
- Ma et al. (*16*): 56.53% sensitivity on 0–3 day post-rash but 82.06% on days 4–28 post-rash.
- Ratnam et al. (*17*): similar conclusions on higher sensitivity in convalescent sera and in sera collected at 6–14 days. In their evaluation of 308 positive and 454 negative samples, indirect ELISAs had lower sensitivity (82.8–88.6%) and specificity (86.6–99.6%) than commercially available IgM capture ELISAs (sensitivity 92.2%, specificity 86.6%).

WHO evaluations of assays used in the Global Measles and Rubella Laboratory Network (GMRLN)noted that "confidence in the accuracy of laboratory classification provided by routine IgM testing is very high", but also noted that the assays have inherent limitations that should be considered while interpreting results (*13*).

I. Evidence for clinical usefulness and impact (from application)

The clinical utility of measles diagnostic tests relates to the ability of the tests to confirm or rule out suspected cases, and one confirmed case results in extensive outbreak control strategies.

A study in outbreak setting by Mosquera et al. (*18*) used various methods to diagnose measles infection and concluded that IgM testing supplemented by PCR or virus isolation identifies maximum patients.

J. Evidence for economic impact and/or cost–effectiveness (from application)

No data available.

K. Ethics, equity and human rights issues (from application)

None identified.

L. Summary of evidence evaluation

IgM is part of the standard diagnosis of measles; it is often part of the reference standard; and it is recommended in WHO guidelines. There are no studies

comparing the results of this test against an independent reference standard (given that it is effectively part of the case definition). The test is embedded in WHO protocols for testing for measles.

Many clinical cases are confirmed by a positive IgM. Evidence supporting the inclusion of IgM is largely based on experience that it confirms a measles diagnosis in most cases outside elimination settings.

M. Summary of SAGE IVD deliberations

Measles is a highly infectious disease of public health concern, and eradication requires concerted effort to identify outbreaks and stop chains of transmission. Despite being a common and severe condition, there is currently no test category listed for measles in the EDL.

Detection of measles IgM is part of the case definition within WHO guidelines and is already established as the gold standard method for diagnosing acute measles. It is especially critical for confirming cases in an outbreak context. SAGE IVD noted that in all settings the test should be reserved for people with clinically suspected measles and is not appropriate as a screening tool for asymptomatic patients.

The reviews submitted show that the test has good clinical accuracy and performance.

SAGE IVD members acknowledged that the test is available in some RDT formats; but this application is only considering an EIA format for use in clinical laboratories. To that end, the group noted that the test is technically easy to perform by laboratory technicians with basic training. But it does require a laboratory with ambient environmental conditions and a regular power supply; moreover, it may be costly in LMICs and resource-constrained settings (although it is still cheap compared with PCR tests).

N. SAGE IVD recommendations

SAGE IVD recommended including the measles IgM test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, Vaccine-preventable diseases);
- using an immunoassay format;
- to diagnose clinically suspected measles infection.

O. References

1. Gastanaduy PA, Redd SB, Clemmons NS, Lee AD, Hickman CJ, et al. Chapter 7: Measles. In: Roush SW, Baldy LM, Kirkconnell Hall MA, editors. Manual for the Surveillance of Vaccine-Preventable Diseases. Atlanta: US Centers for Disease Control and Prevention; 2014.

2. Orenstein WA, Cairns L, Hinman A, Nkowane B, Olivé JM, et al. Measles and rubella global strategic plan 2012–2020 midterm review report: background and summary. Vaccine. 2018;36(Suppl 1):A35–A42. doi:10.1016/j.vaccine.2017.10.065.
3. Surveillance standards for vaccine-preventable diseases, 2nd edition. Geneva: World Health Organization; 2018.
4. O'Donnell S, Davies F, Vardhan M, Nee P. Could this be measles? Emerg Med J. 2019;36:310–314. doi:10.1136/emermed-2019-208490.
5. Ortega-Sanchez IR, Vijayaraghavan SM, Barskey AE, Wallace GS. The economic burden of sixteen measles outbreaks on United States public health departments in 2011. Vaccine. 2014;32(11):1311–1317. doi:10.1016/j.vaccine.2013.10.012.
6. Ozawa S, Portnoy A, Getaneh H, Clark S, Knoll M, et al. Modeling the economic burden of adult vaccine-preventable diseases in the United States. Health Aff. 2016;11:2124–2213. doi:10.1377/hlthaff.2016.0462.
7. Rosen JB, Arciuolo RJ, Khawja AM, Fu J, Giancotti FR, et al. Public health consequences of a 2013 measles outbreak in New York City. JAMA Pediatr. 2018;172(9):811–817. doi:10.1001/jamapediatrics.2018.1024.
8. Lo NC, Hotez PJ. Public health and economic consequences of vaccine hesitancy for measles in the United States. JAMA Pediatr. 2017;171(9):887–892. doi:10.1001/jamapediatrics.2017.1695.
9. Suijkerbuijk AWM, Woudenberg T, Hahné SJM, Lochlainn LN, de Melker HE, et al. Economic costs of measles outbreak in the Netherlands, 2013–2014. Emerg Infect Dis. 2015;21(11):2067–2069. doi:10.3201/eid2111.150410.
10. Filia A, Brenna A, Panà A, Cavallaro GM, Massari M, et al. Health burden and economic impact of measles-related hospitalizations in Italy in 2002–2003. BMC Public Health. 2007;7:169. doi:10.1186/1471-2458-7-169.
11. Njau J, Janta D, Stanescu A, Pallas SS, Pistol A, Khetsuriani N, et al. Assessment of economic burden of concurrent measles and rubella outbreaks, Romania, 2011–2012. Emerg Infect Dis. 2019;25(6):1101–1109. doi:10.3201/eid2506.180339.
12. Thorrington D, Ramsay M, van Hoek AJ, Edmunds WJ, Vivancos R, et al. The effect of measles on health-related quality of life: a patient based survey. PloS One. 2014;9(9):e105153. doi:10.1371/journal.pone.0105153.
13. Manual for the laboratory-based surveillance of measles, rubella, and congenital rubella syndrome, 3rd edition. Geneva: World Health Organization; 2018.
14. Bolotin S, Lim G, Dang V, Crowcroft N, Gubbay J, et al. The utility of measles and rubella IgM serology in an elimination setting, Ontario, Canada, 2009–2014. PLoS One. 2017;12(8):e0181172. doi:10.1371/journal.pone.0181172.
15. Tipples GA, Hamkar R, Mohktari-Azad T, Gray M, Parkyn G, et al. Assessment of immunoglobulin M enzyme immunoassays for diagnosis of measles. J Clin Microbiol. 2003;41(10):4790–4792. doi:10.1128/jcm.41.10.4790-4792.2003.
16. Ma R, Lu L, Suo L, Zhangzhu J, Chen M, et al. Evaluation of the adequacy of measles laboratory diagnostic tests in the era of accelerating measles elimination in Beijing, China. Vaccine. 2019;37(29):3804–3809. doi:10.1016/j.vaccine.2019.05.058.
17. Ratnam S, Tipples G, Head C, Fauvel M, Fearon M, et al. Performance of indirect immunoglobulin M (IgM) serology tests and IgM capture assays for laboratory diagnosis of measles. J Clin Microbiol. 2000;38:99–104.
18. Mosquera MM, de Ory F, Gallardo V, Cuenca L, Morales M, et al. Evaluation of diagnostic markers for measles virus infection in the context of an outbreak in Spain. J Clin Microbiol. 2005;43(10):5117–5121. doi:10.1128/JCM.43.10.5117-5121.2005.

3.24. Measles nucleic acid test

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a human measles RT-PCR test to the EDL as an IVD for diagnosing, or aiding in the diagnosis of, clinically suspected measles infection.

B. Applicant

World Health Organization

C. WHO technical department

Immunization Analysis & Insights

D. Background (from application)

Disease condition and impact on patients

Measles, an acute illness caused by a virus in the family Paramyxoviridae, genus *Morbillivirus*, is one of the most highly infectious diseases known to humankind. Symptoms include fever (as high as 38 °C) and malaise, cough, coryza and conjunctivitis, followed by a maculopapular rash (*1*). Measles is generally a mild or moderately severe illness but can, if not diagnosed or treated, result in complications such as pneumonia, encephalitis and death. A rare long-term sequela of measles virus infection is SSPE, which is a fatal disease of the CNS that generally develops 7–10 years after infection (*1*, *2*). Measles case fatality rates vary from 0.1% in HICs to 15% in LMICs (*2*).

Does the test meet a medical need?

Accurate diagnostic tests for measles infection are essential to confirm cases and outbreaks.

A single laboratory-confirmed measles case should trigger an aggressive public health investigation and response in an elimination setting (*3*).

How the test is used

Optimal RT-PCR testing requires nasopharyngeal or throat swabs collected within the first 4 days of rash onset, or urine specimens collected within the first 7 days. A stand-alone positive RT-PCR result confirms measles infection if

there was no measles vaccination 7–14 days prior to rash onset. If vaccination is evident, follow-up sequencing is recommended to determine wild vs vaccine type strain. If the RT-PCR test is negative and no other cause of symptoms is confirmed, serology follow-up testing is recommended to rule out measles infection. If optimal IgM collection of serum sample (4–28 days from rash onset) is available, a negative IgM test result confirms no infection. If the sample was collected within 3 days of rash onset and the result is negative, repeat testing (≥ 6 days) with IgM is advised if the case remains suspicious for measles.

In elimination areas, where measles incidence is low, RT-PCR testing is recommended to confirm an IgM positive or equivocal test.

E. Public health relevance (from application)

Prevalence

Significant gains towards measles elimination have been made with a highly effective measles vaccine. Countries in all six WHO regions have adopted measles elimination goals, and four WHO regions endorsed the Global Vaccine Action Plan to eliminate measles by 2015. Not all of the plan's goals were accomplished, but measles was successfully eliminated through comprehensive vaccination and surveillance in 61 Member States in the Region of the Americas, and the European and Western Pacific regions (*2*).

Still, in many parts of the world, the disease remains endemic. In 2015, there were 254 928 cases reported and an estimated 134 200 measles deaths globally (*1*). In 2019, the greatest number of cases were reported from the Indian subcontinent, where annual incidence rates exceeded 50 per million population. Other areas of Africa, Asia, Europe, Central and South America, and the Pacific also have large annual numbers of measles cases (*4*).

Vaccine hesitancy remains an obstacle to measles elimination and has been identified by WHO as one of the major threats to global health in 2019. International travel also allows measles importation from endemic countries and has contributed to a resurgence of the disease in high-income countries and countries previously thought to be free of the disease (*4*). For example, in the USA, even though measles was officially eliminated in 2000, occasional outbreaks (three or more linked cases) are still reported to the CDC each year, often imported from overseas and spreading among communities with low rates of immunization (*4*).

Socioeconomic impact

Measles outbreaks can place a heavy economic burden on local and state public health institutions. USA outbreaks in 2011 cost an estimated US$ 2.7–5.3 million (*5*); 2015 outbreaks cost between US$ 0.25 million and US$ 1.35 million (*6*). Estimates from the 2013 outbreak in New York City cost the city's Department

of Health and Mental Hygiene US$ 400 000 (*7*). And in 2016, it was estimated that each case of measles cost the public sector US$ 20 000 (8).

Outside the USA, total costs associated with measles in the Netherlands from 2013 to 2014 were estimated at US$ 4.7 million (*9*); in Italy outbreaks in 2002 and 2003 cost between €17.6 million and €22 million, respectively (*10*). Total societal costs from outbreaks in Romania in 2011 were estimated at US$ 5.5 million (*11*).

In addition, a study of 3207 lab-confirmed measles cases reported by Public Health England from January 2012 to June 2013 resulted in an estimated loss of 44.2 QALYs (*12*).

F. WHO or other clinical guidelines relevant to the test

The 2018 WHO manual for laboratory-based surveillance of measles, rubella and congenital rubella syndrome (*13*) recommends using conventional or real-time RT-PCR to confirm cases, in combination with testing serologic or oral fluid specimens for virus-specific IgM. The manual recommends that this testing be performed in countries in coordination with the GMRLN.

The 2018 WHO surveillance standards (*3*) for vaccine-preventable diseases also recommends RT-PCR to confirm measles case.

G. Basic test characteristics (from application)

Test formats available	RT-PCR
Specimen types	Nasopharyngeal aspirates or throat swabs, oral fluid, urine
Equipment required	Refrigerator, calibrated pipettes, vortex PCR thermocycler
Regulatory status	CE-marked
Availability	Likely global
Price per test range	Not available
Instrument price range	Not available

H. Evidence for diagnostic accuracy (from application)

No systematic reviews of measles RT-PCR test clinical accuracy were available. Three primary studies were reviewed:

- Chuaa et al. (*14*) compared two platforms for measles RT-PCR and found these to be comparable.

- Ma et al. (*15*) showed that RT-PCR has lower sensitivity for detecting measles at 0–3 days in vaccinated individuals.
- Roy et al. (*16*) evaluated an in-house assay in three reference laboratories and found that their assay had 94% sensitivity for identifying five measles vaccine strains.

I. Evidence for clinical usefulness and impact (from application)

Four primary studies show the clinical value of RT-PCR for diagnosing measles:

- Mosquera et al. (*17*) used RT-PCR in an outbreak in Spain and showed its value in addition to IgM testing as it identified additional cases.
- Ma et al. (*15*) showed that RT-PCR used in China at 4–28 days post-rash was more sensitive (94.4%) than IgM (82.1%).
- Cui et al. (*18*) used RT-PCR to supplement IgM testing in a pre-elimination setting and found that RT-PCR successfully identified additional cases among IgM-negative individuals.
- Benamar et al. (*19*) similarly showed RT-PCR additionally identified 52% of IgM-negative cases in an outbreak as measles confirmed.

J. Evidence for economic impact and/or cost–effectiveness (from application)

No data available.

K. Ethics, equity and human rights issues (from application)

None identified.

L. Summary of evidence evaluation

A positive RT-PCR result for measles is considered to confirm infection, and there are no studies comparing the results of this test against an independent reference standard (given that it is effectively part of the case definition). The test is embedded in WHO protocols for testing for measles.

Substantial evidence exists showing that significant numbers of measles cases are detected only by using RT-PCR, particularly among the vaccinated population. There was also evidence that RT-PCR has a greater ability to detect measles in samples taken in the first few days after the rash appears.

Figures from the studies suggest that 4–16% of samples were IgM negative and PCR positive.

M. Summary of SAGE IVD deliberations

Measles is a disease of public health concern, and it is important to detect outbreaks early. Measles IgM testing may be the gold standard for identifying cases, but it has limitations. In particular, detection of measles IgM depends on timing of specimen collection and may need to be complemented with a PCR test. Supplementary testing may also be needed to confirm measles IgM-positive tests in a low measles incidence context (to rule out false positives).

Measles PCR is recommended in WHO guidelines, both for acute diagnosis in the early stages of disease and symptom onset, when IgM can still be negative, and to confirm IgM-positive cases in low-prevalence settings.

Because the test has become part of the case definition, it is difficult to gather data on its accuracy. In this case, SAGE IVD looked for evidence that the PCR test makes a valuable contribution in terms of the number of cases detected by PCR, particularly where IgM is negative. And the studies submitted do provide sufficient evidence of this.

But SAGE IVD noted that specialized equipment and highly skilled personnel are required to perform the PCR test and that it is more costly than either the measles IgM or IgG tests. As such, it may only be applicable where high-level PCR testing is available. In LMICs and resource-constrained settings, it may only be made available at a regional or central level.

SAGE IVD noted that if more robust assays become available that can be performed close to POC, they could reduce the turnaround time for case confirmation and help to rapidly identify cases in an outbreak situation.

N. SAGE IVD recommendations

SAGE IVD recommended including the measles nucleic acid test category in the third EDL:

- as a disease-specific IVD for use in clinical laboratories (EDL 3, Section II.b, Vaccine-preventable diseases);
- using a nucleic acid test format;
- to diagnose clinically suspected measles infection.

O. References

1. Gastanaduy PA, Redd SB, Clemmons NS, Lee AD, Hickman CJ, et al. Chapter 7: Measles. In: Roush SW, Baldy LM, Kirkconnell Hall MA, editors. Manual for the Surveillance of Vaccine-Preventable Diseases. Atlanta: US Centers for Disease Control and Prevention; 2014.
2. Orenstein WA, Cairns L, Hinman A, Nkowane B, Olivé JM, et al. Measles and rubella global strategic plan 2012–2020 midterm review report: background and summary. Vaccine. 2018;36(Suppl 1):A35–A42. doi:10.1016/j.vaccine.2017.10.065.

3. Surveillance standards for vaccine-preventable diseases, 2nd edition. Geneva: World Health Organization; 2018.
4. O'Donnell S, Davies F, Vardhan M, Nee P. Could this be measles? Emerg Med J. 2019;36:310–314. doi:10.1136/emermed-2019-208490.
5. Ortega-Sanchez IR, Vijayaraghavan SM, Barskey AE, Wallace GS. The economic burden of sixteen measles outbreaks on United States public health departments in 2011. Vaccine. 2014;32(11):1311–1317. doi:10.1016/j.vaccine.2013.10.012.
6. Ozawa S, Portnoy A, Getaneh H, Clark S, Knoll M, et al. Modeling the economic burden of adult vaccine-preventable diseases in the United States. Health Aff. 2016;11:2124–2213. doi:10.1377/hlthaff.2016.0462.
7. Rosen JB, Arciuolo RJ, Khawja AM, Fu J, Giancotti FR, et al. Public health consequences of a 2013 measles outbreak in New York City. JAMA Pediatr. 2018;172(9):811–817. doi:10.1001/jamapediatrics.2018.1024.
8. Lo NC, Hotez PJ. Public health and economic consequences of vaccine hesitancy for measles in the United States. JAMA Pediatr. 2017;171(9):887–892. doi:10.1001/jamapediatrics.2017.1695.
9. Suijkerbuijk AWM, Woudenberg T, Hahné SJM, Lochlainn LN, de Melker HE, et al. Economic costs of measles outbreak in the Netherlands, 2013–2014. Emerg Infect Dis. 2015;21(11):2067–2069. doi:10.3201/eid2111.150410.
10. Filia A, Brenna A, Panà A, Cavallaro GM, Massari M, et al. Health burden and economic impact of measles-related hospitalizations in Italy in 2002–2003. BMC Public Health. 2007;7:169. doi:10.1186/1471-2458-7-169.
11. Njau J, Janta D, Stanescu A, Pallas SS, Pistol A, Khetsuriani N, et al. Assessment of economic burden of concurrent measles and rubella outbreaks, Romania, 2011–2012. Emerg Infect Dis. 2019;25(6):1101–1109. doi:10.3201/eid2506.180339.
12. Thorrington D, Ramsay M, van Hoek AJ, Edmunds WJ, Vivancos R, et al. The effect of measles on health-related quality of life: a patient based survey. PloS One. 2014;9(9):e105153. doi:10.1371/journal.pone.0105153.
13. Manual for the laboratory-based surveillance of measles, rubella, and congenital rubella syndrome, 3rd edition. Geneva: World Health Organization; 2018.
14. Chua KYL, Thapa K, Yapa CM, Somerville LK, Chen SCA, et al. What assay is optimal for the diagnosis of measles virus infection? An evaluation of the performance of a measles virus real-time reverse transcriptase PCR using the Cepheid SmartCycler® and antigen detection by immunofluorescence. J Clin Virol. 2015;70:46–52. doi:0.1016/j.jcv.2015.07.004.
15. Ma R, Lu L, Suo L, Zhangzhu J, Chen M, et al. Evaluation of the adequacy of measles laboratory diagnostic tests in the era of accelerating measles elimination in Beijing, China. Vaccine. 2019;37(29):3804–3809. doi:10.1016/j.vaccine.2019.05.058.
16. Roy F, Mendoza L, Hiebert J, McNall RJ, Bankamp B, et al. Rapid identification of measles virus vaccine genotype by real-time PCR. J Clin Microbiol. 2016;55:735–743. doi:10.1128/JCM.01879-16.
17. Mosquera MM, de Ory F, Gallardo V, Cuenca L, Morales M, et al. Evaluation of diagnostic markers for measles virus infection in the context of an outbreak in Spain. J Clin Microbiol. 2005;43(10):5117–5121. doi:10.1128/JCM.43.10.5117-5121.2005.
18. Cui A, Mao N, Wang H, Xu S, Zhu Z, et al. Importance of real-time RT-PCR to supplement the laboratory diagnosis in the measles elimination program in China. PLoS One. 2018;13:11. e0208161. doi:10.1371/journal.pone.0208161.

19. Benamar T, Tajounte L, Alla A, Khebba F, Ahmed H, et al. Real-time PCR for measles virus detection on clinical specimens with negative IgM result in Morocco. PLoS One. 2016;11:1.e0147154. doi:10.1371/journal.pone.0147154.

3.25. Rubella immunoglobulin G antibody

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a rubella IgG antibody test to the EDL as a screening test to determine prior exposure to rubella disease or vaccine for all ages, particularly infants and pregnant women.

B. Applicant

World Health Organization

C. WHO technical department

Immunization Analysis & Insights

D. Background (from application)

Disease condition and impact on patients

Rubella is a contagious infectious disease that affects unvaccinated children and young adults. The virus is easily transmitted by respiratory droplets from an infected individual. Childhood infection is often mild and resolves without complication. But if an unvaccinated woman gets rubella in early pregnancy serious consequences can result, including miscarriage, fetal death, stillbirth and having an infant born with congenital rubella syndrome (CRS). CRS results in a group of devastating birth defects that include blindness, deafness and heart defects. More than 100 000 children are born every year with CRS, mainly in Africa, South-East Asia and the Western Pacific (*1*, *2*).

Rubella vaccine is highly effective and safe when used across a population; as a result, endemic rubella transmission has been interrupted in the Americas since 2009. Incomplete rubella vaccination programmes result in continued disease transmission, as evidenced by large outbreaks in recent years in Japan and elsewhere. Countries with high rates of susceptibility to rubella among women of childbearing age are at highest risk for CRS. This risk varies between and within countries based on epidemiological and socioeconomic differences (*3*). An economic burden study in Romania suggests rubella outbreaks have high costs and considerable impact (*4*).

Does the test meet a medical need?

If a person tests positive for IgM, a recent or current infection is detected. While the symptoms may not be severe for that particular person, if they come into contact with others, particularly women in early pregnancy, the outcome can be severe for the fetus and child. Outbreaks can occur in countries where vaccination programmes are not routine, with severe consequences. A rise in vaccine hesitancy means they can also occur in HICs where the disease has previously been eradicated. Steps such as quarantining infected individuals, and testing and vaccinating unvaccinated family members (or those who have not been previously exposed), is critical; no medications or treatments can prevent the disease other than vaccination or exposure.

IVD diagnostics are critical to manage and control the spread of the disease.

How the test is used

The test supports a final diagnosis on its own for screening for past exposure and immunity; absence of IgG indicates no exposure or vaccination.

The presence of IgG antibodies to rubella virus indicates a previous exposure either by vaccination or prior rubella infection and suggests immunity (*3*). Seroconversion of specific rubella antibodies or a significant rise of the IgG titre strongly supports the diagnosis of acute rubella infection (*3*). The quantitative determination of specific IgG is therefore used to determine the immune status to rubella and distinguish between those who have been exposed in the past and those who may have active infection (*3*).

In countries where records are not well kept, the use of IgG testing can help to identify who has natural immunity and who has immunity gained through vaccination. In eradication efforts, this test is crucial to detecting patients needing the vaccine. In countries where the virus has been rampant, natural immunity can be established to ensure vaccines are reserved for the most vulnerable patients, such as women who intend to become pregnant.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

WHO recommends that all countries that have not yet introduced rubella vaccines should consider doing so using existing, well-established measles immunization programmes. To date, four WHO regions have established goals to eliminate this preventable cause of birth defects. In 2015, the WHO Region of the Americas became the first in the world to be declared free of endemic transmission of rubella. The number of countries using rubella vaccines in their national programme continues to steadily increase. As of December 2018, 168 out of 194 countries had introduced rubella vaccines and global coverage was

estimated at 69%. Reported rubella cases declined 97%, from 670 894 cases in 102 countries in 2000 to 14 621 cases in 151 countries in 2018 (*5*, *6*).

CRS rates are highest in the WHO African and South-East Asia regions, where vaccination coverage is lowest. In April 2012, the Measles Initiative – now known as the Measles & Rubella Initiative – launched its Global Measles and Rubella Strategic Plan for 2012–2020 (*7*). The plan articulates a series of global goals, which include achieving measles and rubella elimination in at least five WHO regions.

Based on the 2018 Global Vaccine Action Plan Assessment Report by SAGE Immunization, rubella control is lagging (*8*), with 26 countries still to introduce the vaccine, while two regions (Africa and Eastern Mediterranean) have not yet set rubella elimination or control targets. SAGE Immunization recommends incorporating rubella vaccination into immunization programmes as quickly as possible to ensure additional gains in controlling rubella.

F. WHO or other clinical guidelines relevant to the test

WHO surveillance standards recommend using IgG alongside IgM testing for evaluating pregnant women exposed to rubella (*9*).

WHO offers governments and communities technical support to improve routine immunization programmes and hold targeted vaccination campaigns. The WHO GMRLN supports the diagnosis of rubella and CRS cases and the tracking of rubella virus spread.

EIAs are the most commonly used and widely available diagnostic test for rubella IgM and IgG antibodies; they are sensitive and relatively easy to perform.

G. Basic test characteristics (from application)

Test formats available	EIA
Specimen types	Serum, plasma, DBS, oral fluid
Equipment required	Refrigerators, freezers, incubators, pipettors, centrifuges, possible safety hood, vortex, microfuge
Regulatory status	CE-marked or FDA-approved
Availability	Global
Price per test range	Not available
Instrument price range	Not available

H. Evidence for diagnostic accuracy (from application)

No systematic reviews of IgG test accuracy were found.

Clinical accuracy in pre-pregnancy and perinatal screening:

Enders et al. (*10*) evaluated six assays on samples from 1090 women visiting antenatal care. Sensitivity varied from 78% to more than 90%, and specificity of all assays was 100%.

Clinical accuracy in routine screening and diagnosis of acute infection:

Dimech et al. (*11*) discuss the clinical accuracy of diagnostic assays on rubella patients and include routine screening and patients with acute infection. A total of 321 samples were evaluated (48 positive and 273 negative). The sensitivities compared with HIA ranged from 98.9% to 99.9%, and the specificity ranged from 77.1% to 95.8%.

I. Evidence for clinical usefulness and impact (from application)

Systematic reviews on the utility of rubella screening and infection are sparse; but one by Thompson et al. (*12*) concluded that serological surveys among women of childbearing age provide critical information with an impact on transmission. The review also highlights the gaps and need for more serological data to be able to eradicate rubella. This supports the importance of using diagnostic assays for future eradication of the disease.

Another systematic review by Lopez et al. (*13*) highlights the importance of establishing a CRS surveillance programme that includes IgG testing to determine the proportion of women who are not immune.

No studies were found showing the utility of rubella IgG testing for diagnosing acute infection.

J. Evidence for economic impact and/or cost–effectiveness (from application)

The cost–effectiveness of the vaccine can translate to IVD value, as treatment for CRS is very costly.

A systematic review by Babigumira et al. (*14*) of 11 cost–benefit analyses, four cost–effectiveness analyses and one cost–utility analysis concluded that CRS treatment is costly, and vaccination programmes are cost-effective.

Kanamori et al. (*15*) discuss the cost–effectiveness of rubella antibody screening among Japanese health care workers. Implementation of a seroprevalence programme (IgG assessment) among 243 new and 2664 previous health care workers was cost-effective and reduced unnecessary vaccination.

K. Ethics, equity and human rights issues (from application)

The use of rubella IVDs to ensure eventual rubella eradication is key to reducing inequity globally. The tests can identify individuals requiring the vaccine and support the need for more global access.

L. Summary of evidence evaluation

Assessment of the accuracy of different tests is challenging, as there is no independent reference standard. Cited studies may be better viewed as assessing equivalence between different tests rather than accuracy.

IgG testing has been shown to be important in vaccination campaigns, both to identify who requires vaccination and to assess the success of campaigns. These include national programmes and institute programmes such as in-hospital staff. There is thus evidence of the clinical utility of the IgG test.

M. Summary of SAGE IVD deliberations

The elimination of rubella is a global priority and requires a concerted effort to identify outbreaks and stop chains of transmission. The rubella IgG test can help by identifying immunity to rubella and preventing CRS, and is particularly useful for screening pregnant women. IgG testing forms part of the case definition for rubella and is an important decision-making tool that is widely used in vaccination strategies, both to identify whom to vaccinate and to assess the success of individual campaigns.

This class of IVDs offer good quality, safety and performance for detecting IgG antibodies to rubella virus. A broad range of products are available at all levels, from small, low-tech laboratories to high-throughput commercial laboratories.

Some specific questions remain about the test's use in very low incidence settings, but these do not negate the overall value of the test for preventing CRS.

WHO guidelines suggest that rubella IgG testing could also be useful as an additional serologic method for case classification, although the submission to the EDL 3 did not include any data to support this test purpose.

N. SAGE IVD recommendations

SAGE IVD recommended the inclusion of the rubella IgG antibody test category in the third EDL:

- as a disease-specific IVD for clinical laboratories (EDL 3; Section II.b, Vaccine-preventable diseases);
- using an immunoassay format;
- to screen for prior exposure to rubella infection or vaccination, particularly in pregnant women.

The group further recommended securing the requisite evidence to enable an additional test purpose for the rubella IgG antibody test category of diagnosing acute infection.

O. References

1. Disease and epidemiology. In: EMRO/Rubella [website]. Cairo: WHO Regional Office of the Eastern Mediterranean; 2020 (http://www.emro.who.int/health-topics/rubella/disease-and-epidemiology.html, accessed 11 March 2020).
2. More about measles, rubella and congenital rubella syndrome (CRS) worldwide. In: CDC/Global Health [website]. Atlanta: US Centers for Disease Control and Prevention; 2020 (https://www.cdc.gov/globalhealth/measles/facts.htm, accessed 11 March 2020).
3. Lambert N, Strebel P, Orenstein W, Icenogle J, Poland GA. Rubella. Lancet. 2015;385(9984):2297–2307. doi:10.1016/S0140-6736(14)60539-0.
4. Njau J, Janta D, Stanescu A, Pallas SS, Pistol A, et al. Assessment of economic burden of concurrent measles and rubella outbreaks, Romania, 2011–2012. Emerg Infect Dis. 2019;25(6):1101–1109. doi:10.3201/eid2506.180339.
5. Rubella. In: WHO/Newsroom [website]. Geneva: World Health Organization; 2020 (https://www.who.int/news-room/fact-sheets/detail/rubella, accessed 11 March 2020).
6. Lanzieri T, Redd S, Abernathy E, Icenogle J. Chapter 14: Rubella. In: Roush SW, Baldy LM, Kirkconnell Hall MA, editors. Manual for the surveillance of vaccine-preventable diseases. Atlanta: US Centers for Disease Control and Prevention; 2020.
7. Measles & Rubella Initiative [website]. Washington DC: American Red Cross; 2020 (https://measlesrubellainitiative.org/, accessed 11 March 2020).
8. Immunization today and in the next decade: 2018 assessment report of the Global Vaccine Action Plan. Geneva: World Health Organization; 2018 (WHO reference number: WHO/IVB/18.11).
9. Rubella. In: Surveillance standards for vaccine-preventable diseases, 2nd edition. Geneva: World Health Organization; 2018.
10. Enders M, Bartelt U, Knotek F, Bunn K, Strobel S, et al. Performance of the Elecsys Rubella IgG assay in the diagnostic laboratory setting for assessment of immune status. Clin Vaccine Immunol. 2013;20(3):420–426. doi:10.1128/CVI.00688-12.
11. Dimech W, Panagiotopoulos L, Francis B, Laven N, Marler J, et al. Evaluation of eight anti-rubella virus immunoglobulin G immunoassays that report results in international units per milliliter. J Clin Microbiol. 2008;46(6):1955–1960. doi:10.1128/JCM.00231-08.
12. Thompson KM, Odahowski CL. Systematic review of measles and rubella serology studies. Risk Anal. 2016;36(7):1459–1486. doi:10.1111/risa.12430.
13. Lopez AL, Raguindin PF, Silvestre MA, Fabay XC, Vinarao AB, et al. Rubella and congenital rubella syndrome in the Philippines: a systematic review. Int J Pediatr. 2016. doi:10.1155/2016/8158712.
14. Babigumira JB, Morgan I, Levin A. Health economics of rubella: a systematic review to assess the value of rubella vaccination. BMC Public Health. 2013;13:406. doi:10.1186/1471-2458-13-406.
15. Kanamori H, Tokuda K, Ikeda S, Endo S, Ishizawa C, et al. Prevaccination antibody screening and immunization program for healthcare personnel against measles, mumps, rubella, and varicella in a Japanese tertiary care hospital. Tohoku J Exp Med. 2014;234(2):111–116. doi:10.1620/tjem.234.111.

3.26. Rubella immunoglobulin M antibody

All content that is taken from the application has been summarized and copy-edited for sense and clarity of language. The original submission and the evidence review are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

A. Proposal

The application proposed adding a rubella IgM antibody test to the EDL as an IVD to determine active rubella infection or a recent infection for all ages when the disease is suspected.

B. Applicant

World Health Organization

C. WHO technical department

Immunization Analysis & Insights

D. Background (from application)

Disease condition and impact on patients

Rubella is a contagious infectious disease that affects unvaccinated children and young adults. The virus is easily transmitted by respiratory droplets from an infected individual. Childhood infection is often mild and resolves without complication. But if an unvaccinated woman gets rubella in early pregnancy serious consequences can result, including miscarriage, fetal death, stillbirth and having an infant born with CRS. CRS results in a group of devastating birth defects that include blindness, deafness and heart problems. More than 100 000 children are born every year with CRS, mainly in Africa, South-East Asia and the Western Pacific (*1*, *2*).

Rubella vaccine is highly effective and safe when used across a population; as a result, endemic rubella transmission has been interrupted in the Americas since 2009. Incomplete rubella vaccination programmes result in continued disease transmission, as evidenced by large outbreaks in recent years in Japan and elsewhere. Countries with high rates of susceptibility to rubella among women of childbearing age are at highest risk for CRS. This risk varies between and within countries based on epidemiological and socioeconomic differences (*3*). An economic burden study in Romania suggests rubella outbreaks have high costs and considerable economic impact (*4*).

Does the test meet a medical need?

If a person tests positive for IgM, a recent or current infection is detected. While the symptoms may not be severe for that particular person, if they come into contact with others, particularly women in early pregnancy, the outcome can be severe for the fetus and child. Outbreaks can occur in countries where vaccination programmes are not routine, with severe consequences. A rise in vaccine hesitancy means they can also occur in HICs where the disease has previously been eradicated. Steps such as quarantining infected individuals and testing and vaccinating unvaccinated family members (or those who have not been previously exposed) is critical; no medications or treatments can prevent the disease other than vaccination or exposure.

IVDs are critical to manage and control the spread of the disease. In an elimination setting, where every single case of rubella must be identified, sensitive, specific, high-quality standardized tests are required to support swift public health interventions.

How the test is used

The test supports a final diagnosis on its own to detect active disease or recent exposure.

Given the high incidence of rubella, and the fact that several countries have not eliminated it (despite having an elimination goal), laboratory investigation by IgM testing for potential rubella cases and CRS is essential, with a huge positive public health impact.

Laboratory confirmation of a suspect rubella case heavily relies on specific IgM serology.

E. Public health relevance (from application)

Prevalence and socioeconomic impact

WHO recommends that all countries that have not yet introduced rubella vaccines should consider doing so using existing, well-established measles immunization programmes. To date, four WHO regions have established goals to eliminate this preventable cause of birth defects. In 2015, the WHO Region of the Americas became the first in the world to be declared free of endemic rubella. The number of countries using rubella vaccines in their national programme continues to steadily increase. By December 2018, 168 out of 194 countries had introduced rubella vaccines and global coverage was estimated at 69%. Reported rubella cases declined 97%, from 670 894 cases in 102 countries in 2000 to 14 621 cases in 151 countries in 2018 (*5*, *6*).

CRS rates are highest in the WHO African and South-East Asia regions, where vaccination coverage is lowest. In April 2012, the Measles Initiative – now

known as the Measles & Rubella Initiative – launched its Global Measles and Rubella Strategic Plan for 2012–2020 (*7*). The plan articulates a series of global goals, which include achieving measles and rubella elimination in at least five WHO regions.

Based on the 2018 Global Vaccine Action Plan Assessment Report by SAGE Immunization, rubella control is lagging (*8*), with 26 countries still to introduce the vaccine, while two regions (Africa and Eastern Mediterranean) have not yet set rubella elimination or control targets. SAGE Immunization recommends incorporating rubella vaccination into immunization programmes as quickly as possible to ensure additional gains in controlling rubella.

F. WHO or other clinical guidelines relevant to the test

WHO rubella surveillance standards (*9*) define a laboratory-confirmed case of rubella as a suspected case of rubella that has been confirmed by an IgM positive result by EIA; a significant rise in serum rubella IgG titre; or rubella PCR positive. These standards also recommend using rubella IgM testing in elimination settings to confirm cases).

WHO recommends collecting specimens on first contact with the case without waiting for the ideal window, to avoid loss to follow-up. The follow-up serum sample for IgM testing should be collected after day 5 post-rash onset for rubella IgM retesting. Samples should still be collected on first contact with the case.

However, IgM detection by EIA for rubella is more sensitive when collected 6–28 days after the onset of rash. A second serum sample may be required for additional testing under the following circumstances:

- Detection of virus-specific RNA by RT-PCR is either unavailable or the results were inconclusive.
- The first serum specimen was collected ≤ 3 days after rash onset and is negative for measles IgM, or is negative in serum collected ≤ 5 days for rubella IgM by EIA.
- Repeat testing of the initial serum specimen fails to resolve an equivocal result for IgM.

WHO offers governments and communities technical support to improve routine immunization programmes and hold targeted vaccination campaigns. The WHO GMRLN supports the diagnosis of rubella and CRS cases and the tracking of rubella virus spread.

EIAs are the most commonly used and widely available diagnostic test for rubella IgM and IgG antibodies; they are sensitive and relatively easy to perform.

G. Basic test characteristics (from application)

Test formats available	EIA
Specimen types	Whole blood, serum, DBS, oral fluid
Equipment required	Refrigerators, freezers. These are dependent on the test used: incubators or water baths, pipettors, timer, centrifuges, possible safety hood, vortex, microfuge, microwell plate washer and reader.
Regulatory status	CE-marked or FDA-approved
Availability	Global
Price per test range	Not available
Instrument price range	Not available

H. Evidence for diagnostic accuracy (from application)

No systematic reviews of IgM test clinical accuracy were found. One primary study (*10*) comparing three immunoassays, based on 57 samples from individuals with recent rubella and 220 samples from those without infection, showed sensitivities of the assays to range from 84.2% to 96.5%, and the specificities to range from 96.8% to 99.9%.

I. Evidence for clinical usefulness and impact (from application)

Despite much progress in tackling rubella, it remains an important pathogen and public health concern around the world. For example, the rubella epidemic in Japan, with more than 11 000 rubella cases and at least 13 CRS cases in the first 6 months of 2013, highlights the fact that a partial vaccination strategy can lead to major outbreaks (*3*). In the Japanese outbreak, 70% of the rubella cases were among males aged 20–39 years, indicating the weakness of the commonly used strategy to only give the rubella vaccine to adolescent girls. In 2012, Poland and Romania also experienced rubella outbreaks that predominantly affected males as a result of a vaccination strategy that initially focused on vaccination of females. For this reason, a global commitment to rubella control, elimination and eventual eradication must be in place.

Rubella virus is a candidate for global elimination because humans are the only known host, safe and highly effective vaccines (> 95% following a single dose) exist, accurate diagnostic assays exist and sustained interruption of endemic transmission has been demonstrated in the Americas since 2009.

Data from different African countries (*11*) and WHO surveillance data on vaccine preventable diseases suggest that rubella virus is common. They also point to a large number of susceptible women of childbearing age, highlighting

the potential risk of acquiring CRS. Diagnostic tests will have a high impact on intervention with vaccines and patient outcomes.

A French survey (*12*) found that the proportion of 18–32-year-olds susceptible to both measles and rubella infections remained high in France even after promotion campaigns about vaccination during and after the major 2009–2011 epidemic, further suggesting the importance of diagnostic intervention for follow-up.

An Australian serological test survey (*13*) before and after a vaccination campaign found that susceptibility to rubella varied considerably between age groups and by sex. Before the vaccination campaign, 40% of children aged 10–12 years were susceptible to rubella; they were too old to have received rubella vaccine in infancy and too young to have received an adolescent dose. These differences disappeared after the campaign, when only 6% of preschool and 5% of primary school-aged children remained susceptible, suggesting the importance of diagnostic tools.

IgM assays can detect rubella active infection during an outbreak to support patient management, particularly of pregnant women and CRS cases.

J. Evidence for economic impact and/or cost–effectiveness (from application)

A cost analysis for IVDs is not available, but cost–effectiveness of the vaccine can translate to IVD value as treatment for CRS is very costly.

K. Ethics, equity and human rights issues (from application)

None identified.

L. Summary of evidence evaluation

Rubella IgM testing is a routine part of investigating possible infections, as indicated in WHO measles and rubella surveillance standards.

Little evidence was provided that quantifies the usefulness of the rubella IgM test. However, the GMRLN tests approximately 100 000 serum specimens annually using an IgM EIA, signalling the availability of data supporting the use of the test.

M. Summary of SAGE IVD deliberations

Rubella IgM testing is a routine part of investigating possible infections, as indicated in WHO rubella surveillance standards. Screening for rubella serostatus is recommended, in certain countries, as part of standard prenatal screening if a pregnant woman has no record of past immunity and no proof of immunization. The use of rubella IVDs to ensure eventual eradication is key to reducing inequity globally.

The submission does not include data quantifying the usefulness of the test. But SAGE IVD acknowledged that IgM testing is used as a standard diagnostic tool across the world and forms an integral part of the case definition for rubella. Furthermore, while not included in the submission, the recent literature includes a body of recommendations on assay accuracy that document and analyse decades of work with IgM assays.

Sensitive and specific assays for detection are available to match a wide range of laboratory and testing capacities.

N. SAGE IVD recommendations

SAGE IVD recommended the inclusion of the rubella IgM antibody test category in the third EDL:

- as a disease-specific IVD for clinical laboratories (EDL 3; Section II.b, Vaccine-preventable diseases);
- using an immunoassay format;
- to diagnose active rubella infection or recent exposure.

The group further recommended that the rubella IgM antibody listing in EDL 3 should be clearly accompanied by a link to WHO guidelines on rubella vaccines, including:

- The immunological basis for immunization series. Module 11: Rubella. Geneva: World Health Organization; 2008 (https://apps.who.int/iris/bitstream/handle/10665/43922/9789241596848_eng.pdf?sequence=1)
- WHO. Rubella vaccines: WHO position paper. Weekly Epidemiol Record. 2011;29(86):301–316 (https://www.who.int/wer/2011/wer8629.pdf?ua=1).

O. References

1. Disease and epidemiology. In: EMRO/Rubella [website]. Cairo: WHO Regional Office of the Eastern Mediterranean; 2020 (http://www.emro.who.int/health-topics/rubella/disease-and-epidemiology.html, accessed 11 March 2020).
2. More about measles, rubella and congenital rubella syndrome (CRS) worldwide. In: CDC/Global Health [website]. Atlanta: US Centers for Disease Control and Prevention; 2020 (https://www.cdc.gov/globalhealth/measles/facts.htm, accessed 11 March 2020).
3. Lambert N, Strebel P, Orenstein W, Icenogle J, Poland GA. Rubella. Lancet. 2015;385(9984):2297–2307. doi:10.1016/S0140-6736(14)60539-0.
4. Njau J, Janta D, Stanescu A, Pallas SS, Pistol A, et al. Assessment of economic burden of concurrent measles and rubella outbreaks, Romania, 2011–2012. Emerg Infect Dis. 2019;25(6):1101–1109. doi:10.3201/eid2506.180339.

5. Rubella. In: WHO/Newsroom [website]. Geneva: World Health Organization; 2020 (https://www.who.int/news-room/fact-sheets/detail/rubella, accessed 11 March 2020).
6. Lanzieri T, Redd S, Abernathy E, Icenogle J. Chapter 14: Rubella. In: Roush SW, Baldy LM, Kirkconnell Hall MA, editors. Manual for the surveillance of vaccine-preventable diseases. Atlanta: US Centers for Disease Control and Prevention; 2020.
7. Measles & Rubella Initiative [website]. Washington DC: American Red Cross; 2020 (https://measlesrubellainitiative.org/, accessed 11 March 2020).
8. Immunization today and in the next decade: 2018 assessment report of the Global Vaccine Action Plan. Geneva: World Health Organization; 2018 (WHO/IVB/18.11).
9. Rubella. In: Surveillance standards for vaccine-preventable diseases, 2nd edition. Geneva: World Health Organization; 2018.
10. Dimech W, Panagiotopoulos L, Marler J, Laven N, Leeson S, et al. Evaluation of three immunoassays used for detection of anti-rubella virus immunoglobulin M antibodies. Clin Diagn Lab Immunol. 2005;12(9):1104–1108. doi:10.1128/CDLI.12.9.1104-1108.2005.
11. Mirambo MM, Majigo M, Aboud S, Groß U, Mshana SE. Serological makers of rubella infection in Africa in the pre vaccination era: a systematic review. BMC Res Notes. 2015;8:716. doi:10.1186/s13104-015-1711.
12. Antona D, Morel P, Jacquot C, Fonteneau L, Dina J, et al. Measles and rubella seroprevalence in a population of young adult blood donors, France 2013. Epidemiol Infect. 2019;147:e109. doi:10.1017/S0950268819000050.
13. Gilbert GL, Escott RG, Gidding HF, Turnbull FM, Heath TC, et al. Impact of the Australian measles control campaign on immunity to measles and rubella. Epidemiol Infect. 2001;127(2):297–303. doi:10.1017/s0950268801005830.

4. Applications for revisions to the EDL

The EDL Secretariat proposed a number of minor corrections to the EDL to correct proofing errors and harmonize the language used to describe the test purpose of all tests in the list, in line with current definitions (see Section 3). All were accepted by the SAGE IVD. Other changes were requested by SAGE IVD members. The most significant ones are described in this section.

In addition to these minor corrections, SAGE IVD received 26 applications for revisions to 24 tests in the EDL. These came from a range of stakeholders, including WHO, partners, industry, academia and professional societies.

An Edits Working Group made up of a selection of SAGE IVD members reviewed each proposed change before the online meetings; a summary of their deliberations was presented to the full SAGE IVD, after which SAGE IVD made its recommendations.

All applications for revisions are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

Section Ib. Disease-specific IVDs for use in community settings and health facilities without laboratories

CLINICAL CHEMISTRY

4.1. Changes requested by SAGE IVD

Following a thorough review by SAGE IVD. the below changes were implemented:

- Tests that were exclusively used for the diagnosis of diabetes mellitus were moved to Section II.b under a diabetes-specific section.
- Whole blood lactate was removed from this section and transferred to Section II.a. SAGE IVD argued that handheld analysers used for this purpose were not widely available in primary care settings, especially in LMICs.

HEPATITIS B

4.2. Hepatitis B e antigen

A. Proposed revision

Médecins Sans Frontières proposed removing the hepatitis B e antigen (HbeAg) RDT because to their knowledge there are none available that are quality assured and have good performance.

B. Summary of working group review

The listing is in line with WHO recommendations to use the HBeAg test to defer treatment of chronic hepatitis B in settings where hepatitis B virus (HBV) DNA is not available (*1*). Still, a note of clarification could be added to the test purpose to reflect that these tests should only be used where HBV DNA NAT is not available.

C. SAGE IVD decision: revision rejected, with edits

Add clarification note to test purpose: (only for use where an HBV DNA test is not available).

D. References

1. Guidelines for the prevention, care and treatment of persons with chronic hepatitis B infection: Mar 201515. Geneva: World Health Organization; 2015.

HIV INFECTION

4.3. Qualitative HIV virological nucleic acid test

A. Proposed revision

Médecins Sans Frontières proposed moving the qualitative HIV virological NAT for diagnosis of HIV infection in infants under 18 months of age out of Section I.b because POC tests for early infant diagnosis (EID) are expensive (both the instrument and complexity of supply management); and because rising maternal-to-child transmission rates are not due to of a lack of POC tests for EID at Tier 1 but rather suboptimal prevention of mother-to-child transmission (PMTCT) programmes.

B. Summary of working group review

The POC NATs for EID are an important aspect of HIV EID, as they have shorter turnaround times and reduce loss to follow-up. Data from recent clinical trials in LMICs (*1*) have shown that these can be scaled up and implementation leads to early diagnosis, reducing loss to follow-up. The rationale presented is reductive: early diagnosis and loss to follow-up are important programme performance measures, and better infant diagnosis and care will likely also affect PMTCT programme performance.

C. SAGE decision: revision rejected

Keep HIV virologic NAT listing for EID in Section I.b of the EDL.

D. References

1. Jani IV, Meggi B, Loquiha O, Tobaiwa O, Mudenyanga C, et al. Effect of point-of-care early infant diagnosis on antiretroviral therapy initiation and retention of patients. Aids. 2018;32(11):1453–1463. doi:10.1097/QAD.0000000000001846.

TUBERCULOSIS

4.4. Lipoarabinomannan antigen

A. Proposed revision

Stop-TB proposed moving the test to Section I.b of the EDL and changing the test purpose to "to aid in the diagnosis of TB in seriously ill HIV-positive inpatients, and in the diagnosis of TB in HIV-positive adult outpatients with signs and symptoms of TB". The rationale given for the revision was that WHO Global TB Programme policy guidance for managing advanced HIV disease and rapid start of ART describe the TB lipoarabinomannan (LAM) test as a POC test that is recommended for use for inpatients living with HIV with CD4 < 100 cells/mm³ or who are seriously ill. The guidance indicates that this recommendation also applies to HIV-positive adult outpatients who are seriously ill regardless of CD4 count or with unknown CD4 count, based on generalizing data from inpatients. WHO also announced at the IAS conference (Mexico, July 2019) that the recommendations for use in outpatient settings was going to be expanded to indicate HIV-positive adults with signs and symptoms of TB. Therefore, given the recommendations apply to outpatients and CD4 testing is not a requirement for using the test, there is no need for a clinical laboratory to run this RDT; the EDL should indicate its category as for use in facilities without laboratories. Furthermore, the test purpose in the EDL should not be restricted to inpatients.

B. Summary of working group review

In keeping with the WHO recommendation for using the LAM test and the policy update (*1*), the test should be used in facilities without laboratories. The test may not be used for those without signs and symptoms of TB and CD4 counts of 100–200 cells/mm3. In seriously ill patients or those with CD4 counts of 100 or less, the test may be used regardless of signs and symptoms of TB.

C. SAGE IVD decision: revision accepted

Move the TB LAM antigen test to Section I.b of the EDL.

D. References

1. Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV: policy update 2019. Geneva: World Health Organization; 2019 (WHO/CDS/TB/2019.16).

Section II.a. General IVDs for use in clinical laboratories

BACTERIOLOGY, MYCOLOGY AND PARASITOLOGY

4.5. Changes requested by SAGE IVD

Following a review by the SAGE IVD, an overall reorganization of this subsection was undertaken. The following changes were implemented:

- The section was renamed "Clinical Microbiology". The SAGE IVD agreed on the change of title for this subsection as it captures and better reflects all the test categories within this section.
- The microscopy section was renamed "Staining procedures", as the SAGE IVD considered that microscopy was better suited as the assay format rather than the test category.
- The urinalysis test strips were removed from this section as they are already listed in Section I.a and in line with the principle of the structure of the EDL where the tests in Section I are "also assumed to be available, in combination with the extended list in Section II, at health care facilities with laboratories".
- The test purposes for several test categories were reworded to reflect current guidelines, including: ALT, AST, CRP, albumin and uric acid.

CLINICAL CHEMISTRY

4.6. Changes requested by SAGE IVD

Following a review by SAGE IVD, an overall reorganization of this subsection was undertaken. The following changes were implemented:

- Disaggregation of panels (basic metabolic panel and comprehensive metabolic panel): members of SAGE IVD agreed that the exact content of these panels varied, depending on the geographical context and resources available. The group's intention was to avoid the use of unnecessary diagnostic tests to preserve limited resources. The content of the panels was maintained in the list as individual test categories.
- Relevant test categories were grouped under kidney function tests.
- Relevant test categories were grouped as a liver profile.
- Tests that were used exclusively for the diagnosis of diabetes mellitus were moved to Section II.b under a diabetes-specific section. This change was also implemented in Section I.

4.7. C-reactive protein

A. Proposed revision

Médecins Sans Frontières proposed adding "monitor response to treatment" to the test purpose, but gave no rationale for the change.

B. Summary of working group review

Although no evidence was provided in the application for a revision, the SAGE IVD working group reviewed scientific literature for the indications for use and found that several studies and reviews show CRP to have a role in monitoring inflammatory conditions and infections (*1–4*). The working group emphasized that the CRP used to diagnose cardiovascular illnesses is high-sensitivity CRP that is not currently listed in the EDL.

C. SAGE IVD decision: revision accepted

Add "to monitor response to treatment" to the test purpose. After further discussion, SAGE IVD further decided to remove specific examples from the test purpose.

D. References

1. Peltola H. C-reactive protein for rapid monitoring of infections of the central nervous system. Lancet. 1982;319(8279):980–983. doi:10.1016/s0140-6736(82)91989-4.
2. Póvoa P. C-reactive protein: a valuable marker of sepsis. Intensive Care Med. 2002;28(3):235–243. doi:10.1007/s00134-002-1209-6.
3. Pääkkönen M, Kallio MJ, Kallio PE, Peltola H. Sensitivity of erythrocyte sedimentation rate and C-reactive protein in childhood bone and joint infections. Clin Orthop Relat Res. 2010;468(3):861–866. doi:10.1007/s11999-009-0936-1.
4. Póvoa P, Teixeira-Pinto AM, Carneiro AH, Portuguese Community-Acquired Sepsis Study Group. C-reactive protein, an early marker of community-acquired sepsis resolution: a multi-center prospective observational study. Crit Care. 2011;15(4):R169. doi:10.1186/cc10313.

4.8. Creatinine

A. Proposed revision

Médecins Sans Frontières proposed adding "monitor kidney function" to the test purpose because the test is used to monitor kidney function in general, but this use is not specific to severe infections.

B. Summary of working group review

This was a typographical error that requires correction.

C. SAGE IVD decision: revision accepted

Following an extensive discussion, the final test purpose was edited to say: "to assess kidney function through estimated glomerular filtration rate (eGFR), urine albumin : creatinine ratio (ACR), and urine protein : creatinine ratio".

4.9. Glucose-6-phosphate dehydrogenase activity

A. Proposed revision

Médecins Sans Frontières proposed moving this test from the clinical chemistry subsection to the haematology one, although they gave no rationale for the change.

B. Summary of working group review

Glucose-6-phosphate dehydrogenase (G6PD) activity is a biochemical assay that detects a genetic metabolic disorder (G6PD deficiency) with consequences for RBCs (leading to haemolytic anaemia). But it is not a direct haematologic test that detects or measures any blood elements. So it should remain under the clinical chemistry category; it could, however, potentially refer to its indications related to haemolytic anaemia, particularly after exposure to certain foods (such as fava beans), medications or infections associated with RBC destruction.

C. SAGE IVD decision: revision rejected, with edits

Keep G6PD in Section II.a, Clinical chemistry of the EDL. SAGE IVD noted that G6PD is also included in Section II.b of the EDL, under malaria, and agreed to amalgamate the two by moving the malaria entry up to Section II.a and revising the test purpose to be more generic, using malaria as an example. So, in addition to the test purpose for screening newborns for G6PD deficiency, the test can also be used to determine G6PD activity (normal, intermediate, deficient) for a decision to administer oxidant drugs, for example 8-aminoquinoline group drugs for radical cure of *P. vivax* malaria.

HAEMATOLOGY

4.10. Changes requested by SAGE IVD

SAGE IVD requested the addition of a test purpose to the D-dimer test. The final test purpose is now:

- to diagnose disseminated intravascular coagulation; and
- to aid in the diagnosis of deep vein thrombosis, pulmonary embolism.

Its function as part of the current recommendations for clinical management of COVID-19 patients was delayed for discussion in the next edition of the list.

4.11. Complete blood count automated

A. Proposed revision

IPOPI proposed adding PID as an indication for this test because it would allow non-PID experts to have all the disease-specific tests that are required to diagnose a PID on a single page of the EDL.

B. Summary of working group review

The use of clinical criteria and several other diagnostic test results may similarly inform the suspecting clinician of the possibility of a PID disorder. But not all these tests are listed, because EDL does not provide diagnostic algorithms. The results of a complete blood count (CBC) test can contribute to numerous diagnoses and is not specific to the diagnosis of PID.

C. SAGE IVD decision: revision rejected

Keep CBC automated in Section II.a of the EDL, with no change.

CLINICAL PATHOLOGY (NEW SECTION)

4.12. Changes requested by SAGE IVD

SAGE IVD requested the creation of an additional subsection for clinical pathology, which now includes the following test categories:

- Urine microscopy
- Body fluid microscopy.

These changes were made as the listing for microscopy in EDL 2 under the Bacteriology, Mycology and Parasitology section contained examples that were not relevant in that section.

SEROLOGY

The serology subsection under Section II.a was renamed "Pregnancy testing", as the only test listed under this section was the hCG. This change was intended to harmonize the listing of this analyte in Sections I and II. In addition, the test purpose related to its use for the diagnosis of germ cell neoplasms was removed, as it is already mentioned in its listing under the Cancer section.

Section II.b. Disease-specific IVDs for use in clinical laboratories

CANCER

4.13. Alpha-fetoprotein

A. Proposed revision

WHO proposed adding "diagnosis and staging of hepatoblastoma" to the test purpose. The Children's Hepatic tumors International Collaboration (CHIC) developed a risk stratification system for use in international clinical trials on the basis of prognostic features present at diagnosis (*1*, *2*), which is the basis for five risk groupings of paediatric hepatoblastoma. The information on alpha-fetoprotein (AFP) is essential for the risk: AFP ≤ 100 ng/mL at diagnosis is a strong independent prognostic factor and is associated with poor prognosis. For patients with higher-risk hepatoblastoma, an escalation of the treatments is mandated.

The risk-stratified staging of hepatoblastoma has been formulated through an intergroup effort, comprising the International Childhood Liver Tumours Strategy Group, Children's Oncology Group, the German Society for Paediatric Oncology and Haematology, and the Japanese Study Group for Paediatric Liver Tumours, based on data from eight completed hepatoblastoma trials (1605 cases) treated between 1988 and 2008. At univariate analysis, the relative risk of death was elevated in patients presenting with AFP < 100 ng/mL (RR 4.3; $P < 0.0001$) vs children with very high AFP (in excess of 1 million ng/ml). The intrinsic prognostic role of AFP (low) has the same clinical implications of the presence of a metastatic disease (R 3.8).

The use of AFP is in combination with abdominal ultrasound to screen patients with genetic syndromes associated with higher risk of hepatoblastoma (e.g. Beckwith–Wiedemann uniparental disomy is outside the scope).

The *WHO Classification of Tumours* also supports this use (*3*).

B. Summary of working group review

The working group accepted the rationale provided and agreed that the test should be used to aid in the diagnosis of hepatoblastoma and staging of hepatoblastoma.

C. SAGE IVD decision: revision accepted

Add to test purpose: "to aid in the diagnosis and staging of hepatoblastoma".

D. References

1. Meyers RL, Maibach R, Hiyama E, Häberle B, Krailo M, et al. Risk-stratified staging in paediatric hepatoblastoma: a unified analysis from the Children's Hepatic tumors International Collaboration. Lancet Oncol. 2017;18(1):122–131. doi:10.1016/S1470-2045(16)30598-8.

2. Czauderna P, Haeberle B, Hiyama E, Rangaswami A, Krailo M, et al. The Children's Hepatic tumors International Collaboration (CHIC): novel global rare tumor database yields new prognostic factors in hepatoblastoma and becomes a research model. Eur J Cancer. 2016;52:92–101. doi:10.1016/j.ejca.2015.09.023.
3. Digestive system tumours. In: WHO Classification of Tumours, 5th edition, volume 1. Geneva: World Health Organization; 2019.

4.14. Basic panel for immunohistochemistry testing for diagnosis of lymphoma

A. Proposed revision

WHO proposed adding the CD138, kappa and lambda chains, and paired box 5 (PAX5) markers to the antigen panel, and provided evidence to support the use of each one, as follows:

- **CD138/syndecan-1:** This is a proteoglycan strongly expressed on multiple myeloma cells. CD138 is present on the surface membrane of 95% of plasma cells in paraffin wax sections and negative on other haemopoietic cells, endothelial cells and other lymphomas. The use of antibodies to CD138 enables assessment of malignant plasmacytosis in the bone marrow, taking into account occasional heterogeneity in tumour antigen expression. CD138 is the gold standard marker for identifying plasma cells (*1*). As highly specific of plasma cells, the presence of a neoplastic clone CD138 positive in bone marrow biopsy sample generally suggests multiple myeloma (*2*, *3*).
- **Kappa and lambda:** Detection of clonality with kappa and lambda immunohistochemistry (IHC) analysis in bone marrow biopsy specimens has sensitivity and specificity for the diagnosis of multiple myeloma (monoclonal) vs reactive plasmacytosis (polyclonal) of 100% and 97.8%, respectively (*4*–*6*).
- **PAX5** is a B cell lineage-specific activator protein and a member of the PAX family of transcription factors (expressed at early, but not late stages of B cell differentiation). The expression of PAX5 has been detected in 95–100% of Reed–Sternberg cells of Hodgkin's lymphoma and is useful in the differential diagnosis of B cell neoplasms, especially with T/null cell anaplastic large cell lymphoma (*7*, *8*).

B. Summary of working group review

The working group accepted the evidence and agreed with the rationale for adding CD138/syndecan-1, kappa and lambda IHC analysis of bone marrow

for the diagnosis of myeloma, and PAX5 IHC staining for the differentiation of Hodgkin's lymphoma.

C. SAGE IVD decision: revision accepted

Add CD138, kappa and lambda chains, and PAX5 to the panel of IHC antigens for diagnosis of lymphoma.

D. References

1. Wei A, Juneja S. Bone marrow immunohistology of plasma cell neoplasms. J Clinic Pathol. 2003;56(6):406–411. doi:10.1136/jcp.56.6.406.
2. Wijdenes J, Vooijs WC, Clément C, Post J, Morard F, et al. A plasmocyte selective monoclonal antibody (B-B4) recognizes syndecan-1. B J Haematol. 1996;94(2):318–323. doi:10.1046/j.1365-2141.1996.d01-1811.x.
3. Rawstron AC, Orfao A, Beksac M, Bezdickova L, Brooimans RA, et al. Report of the European Myeloma Network on multiparametric flow cytometry in multiple myeloma and related disorders. Haematologica. 2008;93(3):431–438. doi:10.3324/haematol.11080.
4. Shafieno Y, Izadi B, Khazaei S, Foroughikia M, Hookari S, et al. The diagnostic value of kappa/lambda ratios determined by immunohistochemistry in the differentiation of plasma cell myeloma from reactive plasmacytosis. Clin Cancer Investig J. 2019;8(5):177. doi:10.4103/ccij.ccij_65_19.
5. Ajise OE, Roshal M, Wang L, Sukhram GN, Smith KM, et al. Clinical utility of morphology, immunohistochemistry, flow cytometry, and FISH analysis in monitoring of plasma cell neoplasms in the bone marrow. J Hematopathol. 2016;9(1):9–18. doi:10.1007/s12308-015-0264-1.
6. Dass J, Arava S, Mishra PC, Dinda AK, Pati HP. Role of CD138, CD56, and light chain immunohistochemistry in suspected and diagnosed plasma cell myeloma: a prospective study. South Asian J Cancer. 2019;8(1):60. doi:10.4103/sajc.sajc_64_17.
7. Torlakovic E, Torlakovic G, Nguyen PL, Brunning RD, Delabie J. The value of anti-pax-5 immunostaining in routinely fixed and paraffin-embedded sections: a novel pan pre-B and B-cell marker. Am J Surg Pathol. 2002;26(10):1343–1350. doi:10.1097/00000478-200210000-00011.
8. Feldman AL, Law ME, Inwards DJ, Dogan A, McClure RF, et al. PAX5-positive T-cell anaplastic large cell lymphomas associated with extra copies of the PAX5 gene locus. Mod Pathol. 2010 Apr;23(4):593. doi:10.1038/modpathol.2010.4.

4.15. Basic panel of immunohistochemical markers for diagnosis of solid tumours

A. Proposed revision

WHO proposed adding a series of markers to the panel, specifically: placental alkaline phosphatase (PLAP) for choriocarcinoma; CD30 for embryonal carcinoma; AFP and PLAP for yolk sac tumour; and Oct3/4, NANOG, CD117/c-kit and SALL4 for seminoma. It provided evidence on the use of each marker, as follows:

- **Oct3/4 and NANOG:** Transcription factors are used for embryonal carcinoma and seminoma. These markers are negative in yolk sac tumour (*1*–*3*).
- **SALL4:** The stem cell marker has been shown to stain all subtypes of germ cell tumours with high sensitivity. In the pivotal experience on SALL4, 22 seminomas, 7 dysgerminomas, 22 embryonal carcinomas, and 14 of 15 yolk sac tumours displayed strong and diffuse SALL positivity in > 90% of tumour cells (80% of tumour cells were strongly positive in the remaining yolk sac tumour). Five of 7 choriocarcinomas and 9 of 18 teratomas were also variably positive for SALL4. In contrast, only 10 (oesophageal, gastric and colonic adenocarcinomas) of 170 metastatic somatic tumours demonstrated focally weak SALL4 reactivity (< 25% tumour cells) (*4*). The experience was replicated on a larger series of 110 patients with germ cell tumours (*5*). All classic seminomas and embryonal carcinomas demonstrated strong SALL4 and OCT4 staining in more than 90% of tumour cells. The only non-neoplastic cells in the testis stained with SALL4 were spermatogonia and a few primary spermatocytes. SALL4 is more sensitive than AFP and glypican-3 for yolk sac tumours.
- **CD30 and CD117/c-kit:** Immunoreactivity is present in seminoma/dysgerminoma but is negative in both embryonal carcinoma and yolk sac tumour. A panel for distinguishing seminoma from embryonal carcinoma should include CD30. Few seminomas (less than 5%) and more than 95% of embryonal carcinoma (*6*) present membrane staining with CD30. Staining for hCG is usually supportive but not pathognomonic of a component of choriocarcinoma. The finding of hCG-positive syncytiotrophoblastic giant cells in seminoma is described, associated with an elevated serum hCG.
- **PLAP:** PLAP is very sensitive for germ cell tumours (positive in > 95% of seminomas), but it lacks specificity. Combining PLAP with a panel for germ cell neoplasms is useful to refine the differential diagnosis, especially between seminoma and non-seminoma tumours, and germ cell tumours and carcinomas.
- **AFP:** AFP is traditionally used to stain yolk sac tumours. But background staining is frequently high, and focal staining can occasionally be seen in the non-yolk sac tumour components. The most specific staining profile for yolk sac tumour is the presence of SALL4 and AFP, and an absence of OC3/4 and NANOG.
- **Wilms' tumour antigen 1 (WT1):** WT1 is a protein associated with multiple solid tumours, both adult (ovarian cancer) and paediatric (Wilms' tumour) neoplasms. For paediatric cancer, the role of

WT1 in pathology diagnostics is especially useful for metastatic or monophasic nephroblastoma. Staining of WT1 in Wilms' tumour is positive in more than 75% of samples, in contrast with non-nephroblastomas, where it is positive in less than 25% of the cases (*7*–*9*). The role of WT1 in the differential diagnosis of kidney masses in children is thus essential.

B. Summary of working group review

The working group accepted the evidence provided and agreed that SALL4, PLAP, CD117/c-kit, Oct3/4 and NANOG reactivity should be listed as markers of germ cell tumours, and WT1 should be listed as a marker of solid tumours (paediatric nephroblastoma).

C. SAGE IVD decision: revision accepted

Add SALL4, PLAP, Oct3/4, NANOG, CD30 and CD117/c-kit, and WT1 to the IHC panel for diagnosis of solid tumours.

D. References

1. Santagata S, Ligon KL, Hornick JL. Embryonic stem cell transcription factor signatures in the diagnosis of primary and metastatic germ cell tumors. Am J Surg Pathol. 2007;31(6):836–845. doi:10.1097/PAS.0b013e31802e708a.
2. Cheng L, Sung MT, Cossu-Rocca P, Jones TD, MacLennan GT, et al. OCT4: biological functions and clinical applications as a marker of germ cell neoplasia. J Pathol. 2007;211(1):1–9. doi:10.1002/path.2105.
3. Jones TD, Ulbright TM, Eble JN, Baldridge LA, Cheng L. OCT4 staining in testicular tumors: a sensitive and specific marker for seminoma and embryonal carcinoma. Am J Surg Pathol. 2004;28(7):935–940. doi:10.1097/00000478-200407000-00014.
4. Cao D, Humphrey PA, Allan RW. SALL4 is a novel sensitive and specific marker for metastatic germ cell tumors, with particular utility in detection of metastatic yolk sac tumors. Cancer. 2009;115(12):2640–2651. doi:10.1002/cncr.24308.
5. Cao D, Li J, Guo CC, Allan RW, Humphrey PA. SALL4 is a novel diagnostic marker for testicular germ cell tumors. Am J Surg Pathol. 2009;33(7):1065–1077. doi:10.1097/PAS.0b013e3181a13eef.
6. Lau SK, Weiss LM, Chu PG. D2-40 immunohistochemistry in the differential diagnosis of seminoma and embryonal carcinoma: a comparative immunohistochemical study with KIT (CD117) and CD30. Mod Pathol. 2007;20(3):320–325. doi:10.1038/modpathol.3800749.
7. Goyal S, Mishra K, Sarkar U, Sharma S, Kumari A. Diagnostic utility of Wilms' tumour-1 protein (WT-1) immunostaining in paediatric renal tumours. Indian J Med Res. 2016;143(Suppl 1):S59. doi:10.4103/0971-5916.191776.
8. Nakatsuka SI, Oji Y, Horiuchi T, Kanda T, Kitagawa M. Immunohistochemical detection of WT1 protein in a variety of cancer cells. Mod Pathol. 2006;19(6):804–814. doi:10.1038/modpathol.3800588.
9. Ramani P, Cowell JK. The expression pattern of Wilms' tumour gene (WT1) product in normal tissues and paediatric renal tumours. J Pathol. 1996;179(2):162–168. doi:10.1002/(SICI)1096-9896(199606)179:2<162::AID-PATH545>3.0.CO;2-0.

4.16. Essential flow cytometry panel of antibodies for leukaemia

A. Proposed revision

WHO proposed adding HLA DR, CD5, CD23 and CD43 markers to the flow cytometry panel, and provided evidence for the use of each one as follows:

- **CD5, CD23 and CD43:** These markers are useful in the diagnosis of chronic lymphocytic leukaemia (CLL)/small lymphocytic lymphoma (SLL). WHO, the International Workshop on Chronic Lymphocytic Leukaemia (iwCLL) and NCCN diagnostic criteria for CLL are based on the morphology and immunophenotype of neoplastic B cells with co-expression of CD19, CD5 and CD23, with weak CD20 and monoclonal surface immunoglobulin expression (*1–3*). In 2018, the European Research Initiative on CLL (ERIC) & European Society for Clinical Cell Analysis (ESCCA) reached a consensus on the diagnosis of CLL, identifying "required" diagnostic markers (such as CD19, CD5, CD20, CD23, kappa and lambda) and "recommended" markers that are potentially useful for differential diagnosis, such as CD43. Based on cases referred for entry into a clinical trial and using each centre's current practice, the proposed ERIC/ESCCA consensus would have an NPV of 92% with a PPV, specificity and sensitivity for the diagnosis of CLL of > 99% (*4*).
- **Human leukocyte antigen (HLA)-DR:** HLA class II molecules are expressed on acute myeloid leukaemia (AML) blasts at diagnosis in most cases of AML, with the exception of acute promyelocytic leukaemia (APL), which is characterized by lack of HLA-DR antigen expression (*5–9*). Optimal diagnostic criteria for untreated APL (CD34– or HLA-DR– and CD11b– and CD11c–) result in 100% sensitivity and 98% specificity (*10*).

B. Summary of working group review

The working group accepted the evidence provided and agreed that all suggested markers should be included as essential flow cytometry antibodies for the diagnosis and differentiation of leukaemia.

C. SAGE IVD decision: revision accepted

Add HLA DR, CD5, CD23 and CD43 to the essential flow cytometry panel of antibodies for leukaemia. In addition, the name of the test was modified to "Basic flow cytometry panel of antibodies for leukaemia" to harmonize it with other panel names.

D. References

1. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, et al. WHO classification of tumours of haematopoietic and lymphoid tissues, 4th edition. In: WHO classification of tumours, volume 2. Geneva: World Health Organization; 2008.
2. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute – Working Group 1996 guidelines. Blood. 2008;111(12):5446–5456. doi:10.1182/blood-2007-06-093906.
3. Wierda WG, Byrd JC, Abramson JS, Bilgrami SF, Bociek G, et al. NCCN guidelines insights: chronic lymphocytic leukemia/small lymphocytic lymphoma, version 2.2019. J Natl Compr Canc Netw. 2019;17(1):12–20. doi:10.6004/jnccn.2019.0002.
4. Rawstron AC, Kreuzer KA, Soosapilla A, Spacek M, Stehlikova O, et al. Reproducible diagnosis of chronic lymphocytic leukemia by flow cytometry: An European Research Initiative on CLL (ERIC) & European Society for Clinical Cell Analysis (ESCCA) Harmonisation project. Cytometry B Clin Cytom. 2018;94(1):121–128. doi:10.1002/cyto.b.21595.
5. Wetzler M, McElwain BK, Stewart CC, Blumenson L, Mortazavi A, et al. HLA-DR antigen-negative acute myeloid leukemia. Leukemia. 2003;17(4):707–715. doi:10.1038/sj.leu.2402865.
6. Hanson CA, Gajl-Peczalska KJ, Parkin JL, Brunning RD. Immunophenotyping of acute myeloid leukemia using monoclonal antibodies and the alkaline phosphatase-antialkaline phosphatase technique. Blood. 1987;70(1): 83–89.
7. Scott CS, Patel D, Drexler HG, Master PS, Limbert HJ, Roberts BE. Immunophenotypic and enzymatic studies do not support the concept of mixed monocytic-granulocytic differentiation in acute promyelocytic leukaemia (M3): a study of 44 cases. B J Haematol. 1989;71(4):505–509. doi:10.1111/j.1365-2141.1989.tb06310.x.
8. De Rossi G, Avvisati G, Coluzzi S, Fenu S, LoCoco F, et al. Immunological definition of acute promyelocyte leukemia (FAB M3): A study of 39 cases. Eur J Haematol. 1990;45(3):168–171. doi:10.1111/j.1600-0609.1990.tb00446.x.
9. Stone RM, Mayer RJ. The unique aspects of acute promyelocytic leukemia. J Clin Oncol. 1990;8(11):1913–1921. doi:10.1200/JCO.1990.8.11.1913.
10. Horna P, Zhang L, Sotomayor EM, Lancet JE, Moscinski LC. Diagnostic immunophenotype of acute promyelocytic leukemia before and early during therapy with all-trans retinoic acid. Am J Clin Pathol. 2014;142(4):546–552. doi:10.1309/AJCPPOKEHBP53ZHV.

4.17. Basic panel of immunohistochemical markers for diagnosis of solid tumours

A. Proposed revision

Abbott proposed removing IHC and flow cytometry, respectively, from the diagnostic test name for these two listings and adding fluorescent in situ hybridization (FISH) as an assay format, although no rationale or evidence was provided to support the change.

B. Summary of working group review

While FISH is acknowledged to be a valuable diagnostic tool, it has significant cost, training and infrastructure requirements; moreover, evidence on its

performance for specific indications still needs to be reviewed. This should be done through full submissions for FISH antigens for specific test purposes.

C. SAGE IVD decision: revision rejected

Keep the test format as IHC testing.

4.18. Tyrosine-protein kinase receptor or human epidermal growth factor receptor 2 overexpression

A. Proposed revision

Abbott proposed adding FISH as a test format, although no rationale or evidence was provided to support the change.

B. Summary of working group review

While FISH is acknowledged to be a valuable diagnostic tool, it has significant cost, training and infrastructure requirements; moreover, evidence on its performance for specific indications still needs to be reviewed. This should be done through full submissions for FISH antigens for specific test purposes.

C. SAGE IVD decision: revision rejected

Keep the test format as IHC testing.

4.19. Lactate dehydrogenase activity

A. Proposed revision

The University of New South Wales, Australia, proposed reviewing the LDH activity test purpose with particular relevance to heart disease, liver disease, thrombotic thrombocytopenic purpura (TTP), *Pneumocystis* infection and other malignancies, arguing that assay of LDH activity in serum or plasma is a long-established chemical pathology test for many conditions. LDH is widely distributed in the body, and damage to tissues such as heart, liver, kidney, skeletal muscle and red blood cells may cause cells to release LDH, elevating readings in serum or plasma. This means that while LDH may be a sensitive assay for tissue damage, the LDH could come from various sources. The relative proportion of different LDH isoenzymes has been used to give a clue to the tissue origin of LDH in the circulation, but this step adds to the complexity of testing. Biomarkers with more specific tissue distribution are available and have supplanted LDH in a number of settings. Apart from malignancy, there are very few systematic reviews or meta-analyses on the diagnostic role of LDH:

- **Role in heart disease:** LDH was an early biomarker for myocardial infarction, but was supplanted by creatine kinase and other assays,

which in turn were supplanted by troponins T and I. A recent review of diagnostic biomarkers of acute coronary syndrome gave troponins T and I the top score of 4 for both sensitivity and specificity, whereas LDH scored 2 for sensitivity and 1 for specificity (*1*). Troponin assays are readily available, so there is little role for LDH as a cardiac biomarker.

- **Role in liver disease:** LDH activity is included in clinical chemistry platforms that measure various liver function tests. But in approach to a patient with liver disease, LDH is insensitive and non-specific (*2*). LDH now has little value in assessment of liver function.
- **Role in TTP:** TTP is a rare and serious condition in which the diagnostic gold standard is a marked reduction of ADAMTS13 (a disintegrin and metalloprotease with a thrombospondin type 1 motif, member 13) activity in plasma. In patients diagnosed with TTP, LDH activity is prognostically useful, with elevated LDH associated with poorer outcomes (3). Assays for ADAMTS13 activity are complex and are typically done in highly specialized laboratories. This role of LDH could be reviewed after a full assessment of other diagnostics for TTP.
- **Role in *Pneumocystis* infection:** The *Pneumocystis* PCR submission to EDL 3 states that the PCR may be used in conjunction with LDH, oxygen saturation and radiology in diagnosis. LDH is a well-established biomarker that is sensitive for *Pneumocystis* pneumonia, but it is not specific (*4*). Assessment of BDG in the circulation provides greater specificity in the diagnosis of *Pneumocystis* pneumonia compared with LDH (*5*). If the BDG assay is widely available, there is no need to consider adding *Pneumocystis* pneumonia to the test purposes of LDH. But if there is limited availability of the BDG assay, LDH could be considered as an alternative.
- **Role in malignancy:** The LDH entry states that LDH has a role in the prognosis and monitoring of malignant disorders. But it also has a role in diagnosis. In the investigation of lymphoma, the BMJ Best Practice website includes LDH in the list of first investigations (*6*). LDH elevation is an indication of the proliferative rate of the lymphoma, which may be an important diagnostic and prognostic factor. The National Comprehensive Cancer Network also includes LDH in the list of essential tests in the workup of lymphoma (*7*). And the National Academy of Clinical Biochemistry recommends LDH for diagnosis, prognosis and monitoring of testicular cancer (*8*).

The applicant suggested there is evidence to support the addition of "diagnosis" to the use of LDH in haematological malignancies (lymphoma) and germ cell tumours.

B. Summary of working group review

The working group accepted the evidence provided and agreed that the role of LDH should be expanded to include diagnosis, monitoring and prognosis of haematological malignancies (lymphoma) and germ cell tumours. It noted that the role of LDH in diagnosing heart and liver disease is limited, and agreed that while LDH may be a prognostic factor in TTP, a full assessment of other diagnostic and prognostic IVDs should be performed before including this role in the EDL. For *Pneumocystis* pneumonia, the availability of better IVDs such as BDG should be evaluated.

C. SAGE IVD decision: revision accepted, with edits

Keep the test purpose as it is but add a note stating that LDH activity is also used as a marker in heart disease, liver disease, *Pneumocystis* infection, TTP and other malignancies. SAGE IVD recommended circulating the LDH activity EDL listing to relevant specialist organizations for a further revision (for EDL 4) to incorporate any additional conditions that may have been overlooked.

D. References

1. Danese E, Montagnana M. An historical approach to the diagnostic biomarkers of acute coronary syndrome. Ann Transl Med. 2016;4(10):194. doi:10.21037/atm.2016.05.19.
2. Murali AR, Carey WD. Liver test interpretation approach to the patient with liver disease: a guide to commonly used liver tests. Lyndhurst: Cleveland Clinic Center for Continuing Education; 2017 (http://www.clevelandclinicmeded.com/medicalpubs/diseasemanagement/hepatology/guide-to-common-liver-tests/#top, accessed 8 March 2020).
3. Staley EM, Cao W, Pham HP, Kim CH, Kocher NK, et al. Clinical factors and biomarkers predict outcome in patients with immune-mediated thrombotic thrombocytopenic purpura. Haematologica. 2019;104(1):166–175. doi:10.3324/haematol.2018.198275.
4. Quist J, Hill AR. Serum lactate dehydrogenase (LDH) in Pneumocystis carinii pneumonia, tuberculosis, and bacterial pneumonia. Chest. 1995;108(2):415–418. doi:10.1378/chest.108.2.415.
5. Esteves F, Calé SS, Badura R, de Boer MG, Maltez F, et al. Diagnosis of pneumocystis pneumonia: evaluation of four serologic biomarkers. Clin Microbiol Infect. 2015;21(4):379. doi:10.1016/j.cmi.2014.11.025.
6. Non-Hodgkin's Lymphoma. In: BMJ Best Practice [website]. London: British Medical Journal; 2019 (https://bestpractice.bmj.com/topics/en-gb/312, accessed 8 March 2020).
7. Zelenetz AD, Abramson JS, Advani RH, Andreadis CB, Bartlett N, et al. Non-Hodgkin's lymphomas. J Natl Compr Canc Netw. 2011;9(5):484–560. doi:10.6004/jnccn.2011.0046.

8. Sturgeon CM, Duffy MJ, Stenman UH, Lilja H, Brünner N, et al. National Academy of Clinical Biochemistry laboratory medicine practice guidelines for use of tumor markers in testicular, prostate, colorectal, breast, and ovarian cancers. Clin Chem. 2008;54(12):e11–79. doi:10.1373/clinchem.2008.105601.

HEPATITIS B

4.20. Quantitative HBV virological nucleic acid test

A. Proposed revision

Abbott proposed adding DBS to the list of specimen types, although no rationale or evidence was provided to support the change.

B. Summary of working group review

WHO guidelines on hepatitis B and C testing provide a conditional recommendation for using DBS for HBV NATs (*1*), and stipulate that programmes should use only assays with manufacturer validation of DBS. However, current technologies do not have a claim for use of the test using DBS as a specimen type.

C. SAGE IVD decision: revision rejected

Keep list of specimen types as is and modify based on changes to the commercially available products.

D. References

1. Guidelines on hepatitis B and C testing. Geneva: World Health Organization; 2017.

HEPATITIS C

4.21. Qualitative or quantitative HCV virological nucleic acid test

A. Proposed revision

Abbott proposed adding DBS to the list of specimen types, although no rationale or evidence was provided to support the change.

B. Summary of working group review

WHO guidelines on hepatitis B and C testing provide a conditional recommendation for using DBS for HCV NATs (*1*), and stipulate that programmes should use only assays with manufacturer validation of DBS. The working group suggested that results of an ongoing evaluation on DBS performance and protocols for HCV (*2*) could be considered before listing as a recommended specimen type.

C. SAGE IVD decision: revision accepted

SAGE IVD initially recommended keeping the list of specimen types as is. However, upon further discussion with the WHO relevant team, DBS was added to the list as this specimen type has been recommended by WHO since 2017 and commercially available products now carry a claim for DBS.

D. References

1. Guidelines on hepatitis B and C testing. Geneva: World Health Organization; 2017.
2. Hepatitis C & HIV. In: FIND [website]. Geneva: Foundation for Innovative New Diagnostics; 2020 (https://www.finddx.org/hcv-hiv/, accessed 8 March 2020).

HIV INFECTION

4.22. Quantitative HIV virological nucleic acid test

A. Proposed revisions

Abbott proposed removing the condition for manufacturer validation in diagnosis of HIV infection in infants < 18 months of age, arguing that laboratories should be allowed to validate this off-label use. Abbott further proposed adding a condition to the specimen types stipulating that plasma collection is only required where phlebotomists are available, although no rationale or evidence was provided to support this change.

B. Summary of working group review

The working group emphasized that the intended use validated by the manufacturer is critical to the ways devices are used. In addition, individual laboratory validation is prone to introducing non-standardized methods that vary widely. Verifying manufacturer claims is an important part of quality assurance; but manufacturers should generate data to validate the use of their test system in infants and be explicit about this claim. Given that most quantitative assays are currently only validated for viral load measurements, only assays with claims for diagnosis or as an aid in diagnosis should be considered for this test category.

The working group noted that IVDs listed in Section II are described as suitable for facilities with access to laboratories and such facilities usually have phlebotomists available, so there is no need for a new note to stipulate that plasma collection is only required where phlebotomists are available.

C. SAGE IVD decision: revisions rejected

Keep the condition for manufacturer validation for diagnosis of HIV infection in infants < 18 months of age. Do not add any condition to the specimen types

HUMAN PAPILLOMAVIRUS INFECTION

4.23. Human papillomavirus nucleic acid test

A. Proposed revisions

Abbott proposed adding liquid cytology to the list of specimen types, and changing "fluid" to "vessel" for cervical cell collection to account for acceptable collection devices not containing fluid. No rationale or evidence was provided to support the changes.

B. Summary of working group review

A change from cervical cells to liquid cytology is not necessary, as the specimen type requires that cervical cells be included in the sample. But, as some new and prequalified products do not depend on liquid/ fluid transport media (*1*), the specimen type may be changed to "fluid/vessel".

C. SAGE IVD decision: revisions rejected, with edits

Include a minor edit to the specimen type: cervical cells collected in test-specific transport fluid/vessel.

D. References

1. Public reports of WHO prequalified IVDs. In: WHO/In vitro diagnostics and laboratory technology [website]. Geneva: World Health Organization; 2020 (https://www.who.int/diagnostics_laboratory/evaluations/pq-list/public_report_hpv/en/, accessed 8 March 2020).

PRIMARY IMMUNODEFICIENCIES

4.24. Lymphocyte subtype enumeration

A. Proposed revision

IPOPI proposed adding CD3 (T cells) and CD19 (B cells) to the list of lymphocyte subtype enumeration markers. T cells are CD3+ lymphocytes and are sometimes the only cell surface marker used to count for T cells. CD4 and CD8 cell surface markers are T cell subtypes, helper and cytotoxic, respectively. The most severe PIDs affect the T cell compartment and may harbour normal relative values of CD4 and/or CD8. B cells are counted using either CD19 or CD20 cell surface markers (1–4). Some laboratories use one or the other, so both should be mentioned in the diagnostic toolkit for B cell defects, leading to agammaglobulinaemia or hypogammaglobulinaemia.

B. Summary of the working group review

The working group accepted the evidence submitted and agreed that both CD3 and CD19 should be included as markers for the specified indication.

C. SAGE IVD decision: revision accepted

Add CD3 and CD19 as T and B cell markers to the test name.

D. References

1. Finak G, Langweiler M, Jaimes M, Malek M, Taghiyar J, et al. Standardizing flow cytometry immunophenotyping analysis from the Human ImmunoPhenotyping Consortium. Sci. Rep. 2016;6:20686. doi:10.1038/srep20686.
2. Ma CS, Tangye SG. Flow cytometric-based analysis of defects in lymphocyte differentiation and function due to inborn errors of immunity. Front. Immunol. 2019;10:2108. doi:10.3389/fimmu.2019.02108.
3. van Dongen JJM, van der Burg M, Kalina T, Perez-Andres M, Mejstrikova E, et al. EuroFlow-based flowcytometric diagnostic screening and classification of primary immunodeficiencies of the lymphoid system. Front. Immunol. 2019;10:1271. doi:10.3389/fimmu.2019.01271.
4. Madkaikar MR, Shabrish S, Kulkarni M, Aluri J, Dalvi A, et al. Application of flow cytometry in primary immunodeficiencies: experience from India. Front. Immunol. 2019;10:1248. doi:10.3389/fimmu.2019.01248.

SEXUALLY TRANSMITTED INFECTIONS

4.25. Qualitative test for *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infections

A. Proposed revision

Abbott proposed adding ocular swab to the list of specimen types for sampling for infant blindness, although no evidence was provided to support this indication.

B. Summary of working group review

The working group reviewed some available evidence on the performance of a few assays for detecting trachoma in children 0–9 years (*1*, *2*). But more evidence on field performance is needed because the manufacturer validation data is not available for this specimen type. The group suggested waiting for the results of the WANTAIM trial (*3*) (POC test use on 2000 eye swabs from neonates) and other forthcoming data before listing this specimen type.

C. SAGE IVD decision: revision rejected

Keep list of specimen types as is.

D. References

1. Jenson A, Dize L, Mkocha H, Munoz B, Lee J, et al. Field evaluation of the Cepheid GeneXpert Chlamydia trachomatis assay for detection of infection in a trachoma endemic community in Tanzania. PLoS Negl Trop Dis. 2013;7(7):e2265. doi:10.1371/journal.pntd.0002265.

2. Chan M, Jiang B, Ng SY, Tan TY. Novel cost-effective quality control approach for the Cepheid Xpert CT/NG assay for the detection of Chlamydia Trachomatis and Neisseria Gonorrhoeae. J Microbiol Methods. 2016;125:87–90. doi:10.1016/j.mimet.2016.04.010.

3. Vallely AJ, Pomat WS, Homer C, Guy R, Luchters S, et al. Point-of-care testing and treatment of sexually transmitted infections to improve birth outcomes in high-burden, low-income settings: study protocol for a cluster randomized crossover trial (the WANTAIM Trial, Papua New Guinea). Wellcome Open Res. 2019;4:53. doi:10.12688/wellcomeopenres.15173.2.

TUBERCULOSIS

4.26. *Mycobacterium tuberculosis* DNA nucleic acid test

A. Proposed revisions

Abbott proposed removing the condition for simultaneous rifampin detection from the test purpose, but gave no rationale or evidence to support the change.

The Stop-TB Partnership further proposed moving the test to Section I of the EDL and changing the test format to POC NAT. The organization stated that the only WHO-recommended NAT for TB in the EDL uses the GeneXpert platform, which is also a platform for use of a WHO-prequalified NAT for diagnosing HIV infection in infants < 18 months of age (early infant diagnosis [EID]). While the NAT for TB is categorized in the EDL as being for use in clinical laboratories, the test for HIV EID is described as a POC test and categorized for use in health facilities without laboratories. This represents a significant discordance, given the use of the TB and HIV EID tests on the same platform and given the TB test and its sample preparation also have minimal training and biosafety requirements. Furthermore, the GeneXpert platform family includes a new system, GeneXpert Edge, which is portable and battery operated and allows for even further decentralization; this system received an approved change request by WHO prequalification in January 2019.

B. Summary of working group review

As regards Abbott's proposal, the working group was clear that simultaneous detection of resistance provides an additional value, which is why mentioning it separately is justified. A recent WHO policy update (*1*) evaluated assays that sequentially detect rifampin resistance after detecting *M. tuberculosis* DNA for diagnosis and concluded that the detection of rifampin resistance is a priority. The test purpose should be updated in line with WHO recommendations to simultaneously or sequentially detect rifampin resistance.

As regards Stop-TB's request, the working group acknowledged that recently released platforms are better aligned for POC use, but recommended restricting NAT use to laboratory settings pending availability of data on the acceptability and implementation of new platforms as well as process indicators of reduced time to treatment and impact on earlier diagnosis, and, lastly, cost–effectiveness.

The implementation of HIV NAT for EID has been evaluated and determined to be "accurate, feasible and acceptable in hospital settings as a POC test, with benefits of early ART for all positive infants at birth facilities" (*2*). In rural health clinics without laboratories in a clinical trial setting, use of POC NAT for EID diagnosis "enabled clinics to more rapidly diagnose and provide treatment to HIV-infected infants, reduced opportunities for pretreatment loss to follow-up and enabled a larger proportion of infants to receive test results and initiate antiretroviral therapy", leading to "early retention in care" (*3*). Such data are not available for TB diagnosis.

C. SAGE IVD decision: revisions rejected, with edits

Update test purpose to indicate either simultaneous or sequential rifampin resistance, but keep the test in Section II.b of the EDL.

D. References

1. Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: rapid communication. Policy update. Geneva: World Health Organization; 2020.
2. Spooner E, Govender K, Reddy T, Ramjee G, Mbadi N, et al. Point-of-care HIV testing best practice for early infant diagnosis: an implementation study. BMC Public Health. 2019;19:731. doi:10.1186/s12889-019-6990-z.
3. Jani IV, Meggi B, Loquiha O, Tobaiwa O, Mudenyanga C, et al. Effect of point-of-care early infant diagnosis on antiretroviral therapy initiation and retention of patients. AIDS. 2018;32(11):1453–1463. doi:10.1097/QAD.0000000000001846.

4.27. *Mycobacterium tuberculosis* DNA loop-mediated isothermal amplification

A. Proposed revision

Abbott proposed removing the LAMP format from the *M. tuberculosis* DNA test to diagnose active TB (and restricting the listing to the NAT only), to accommodate broader acceptable technology description.

B. Summary of working group review

TB-LAMP is based on a specific NAT amplification method, so a separate listing is justified.

C. SAGE IVD decision: revision rejected

Keep LAMP assay format.

4.28. *Mycobacterium tuberculosis* DNA mutations associated with resistance

A. Proposed revision

Abbott proposed removing the molecular line probe assay (LPA) format from the *M. tuberculosis* DNA test to detect resistance to first- and second-line anti-TB medicines (and restricting the listing to the NAT only), to accommodate a more broadly acceptable technology description.

B. Summary of working group review

LPA is a specific NAAT type that is based on PCR and reverse hybridization, so a separate listing is justified. The difference between effects of first-line and second-line drugs justifies their separate listing.

C. SAGE IVD decision: revision rejected

Keep LPA assay formats.

Section II.c. Disease-specific IVDs for blood screening laboratories

OTHER TRANSFUSION-TRANSMITTED ORGANISMS

4.29. IVDs to screen for pathogens in blood donations

A. Proposed change

Médecins Sans Frontières proposed adding malaria to the list of pathogens, although no rationale or evidence was provided to support the change.

B. Summary of working group review

To screen blood for malaria, some countries recommend universal screening while others recommend selective screening. In all cases, careful donor selection criteria and deferral strategies should be implemented. Diagnostic screening requires local epidemiological evidence. The applied assays differ by epidemiological context. For endemic countries, detection by thick-film microscopy is recommended, and where unavailable, antigen detection tests may be used (*1*). In non-endemic areas, antibody detection is preferred.

C. SAGE IVD decision: revision accepted

Include malaria on the list of pathogens to screen for, in accordance with WHO guidelines.

D. References

1. Screening donated blood for transfusion-transmissible infections: recommendations. Geneva: World Health Organization; 2010.

5. Applications for Do Not Do recommendations in the EDL

WHO technical departments submitted two applications for IVDs to be added to the new list of Do Not Do recommendations in the EDL, which recommend against use of named tests. Both applications were submitted based on the latest WHO guidelines and policies, which are informed by systematic reviews of quality studies, so an abbreviated review process (using external experts) was used. SAGE IVD then evaluated the reviews alongside the available evidence to support the WHO guideline in question and made its decision on whether or not to recommend against use of the proposed IVDs.

All applications for revisions are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

5.1. Commercial serodiagnostic tests to diagnose tuberculosis

A. Applicant

World Health Organization

B. WHO technical department

Global TB Programme

C. Background

Serological diagnostic tests (antibody recognition of antigens of *M. tuberculosis* by the humoral immune response) in a user-friendly format could potentially enhance capacity for TB diagnosis to lower levels of health services. But systematic reviews of studies performed on such tests for TB have revealed that these tests are unreliable. Several commercial serological tests for TB are being marketed, despite growing evidence of poor sensitivity and specificity of these tests for the diagnosis of TB (*1*).

D. WHO guidance

A 2011 WHO policy update (*1*) reviewed evidence on the diagnostic accuracy of commercial serological tests for pulmonary and extrapulmonary TB, using the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) approach to evidence evaluation.

E. Evidence supporting WHO guidance

For pulmonary TB, 67 studies on adult populations were included in the review (32 of these studies were performed in LMICs). Sensitivities and specificities of all commercial tests evaluated were highly variable, giving ranges of 0–100%

and 31–100%, respectively. For extrapulmonary TB 25 studies on adult populations were reviewed (10 from LMICs). Sensitivity and specificity were similarly highly variable for all commercial tests, 0% to 100% and 59% to 100%, respectively. The review concluded that these tests are imprecise and do not lead to patient-important outcomes. High false-positive and false-negative rates make the tests harmful, with negative impacts on patient safety. The policy statement recommends against using these tests to diagnose pulmonary or extrapulmonary TB.

F. Summary of SAGE IVD deliberations

The WHO policy against TB serology assays has been in place for nearly a decade and is unlikely to change. It is clear that these tests do not increase diagnostic certainty, do not support clinical management and do not provide valuable prognostic information. There is great potential to cause harm through overdiagnosis and incurred costs, including harmful therapy instituted based on test results.

These tests should not be used to diagnose or monitor TB. Inclusion in the EDL negative list is intended to guide countries against procuring or using these tests. But this does not mean that research to develop a good serodiagnostic test should discontinue. As the WHO policy clearly states, targeted research to identify more accurate serological tests is strongly encouraged, based on adequate study design, including quality principles such as representative suspect populations, prospective follow-up and adequate, explicit blinding.

G. SAGE IVD recommendations

SAGE IVD recommended against using commercial serodiagnostic tests for diagnosis of TB, and requested its addition to the Do Not Do recommendations of the third EDL (EDL 3, Section III).

H. References

1. WHO. Commercial serodiagnostic tests for diagnosis of tuberculosis: policy statement. Geneva: World Health Organization; 2011.

5.2. Western blot and line immunoassays to diagnose HIV infection

A. Applicant

World Health Organization

B. WHO technical department

Department of HIV/AIDS

C. Background

Current HIV testing algorithms rely on highly sensitive and specific assays. Algorithms before 2014 used the western blot (WB) assay, which has lower sensitivity in early HIV infection. Using this assay in algorithms may also delay the start of ART, with consequent loss to follow-up. Continued investment in HIV WB assays diverts economic and human resources from more accurate assays.

D. WHO guidance

WHO consolidated guidelines on HIV testing services (*1*) recommend that WB and line immunoassays should not be used in national HIV testing strategies and algorithms (strong recommendation, low-quality evidence). To support scale-up of HIV testing, prevention and treatment, countries should move away from WB and line immunoassays in favour of simpler RDTs and EIAs.

E. Evidence supporting WHO guidance

Clinical accuracy data for the performance of HIV WB assays from six studies (*2–7*) revealed that using WB in an algorithm does not impact the sensitivity and specificity of HIV diagnosis. But WB assays produced a high number of indeterminate results that were not observed with EIAs.

Data from five studies (*8–12*) reveal that using WB assays in HIV testing algorithms also led to longer turnaround times and greater loss to follow-up when compared with WB-free algorithms.

F. Summary of SAGE IVD deliberations

HIV WB and line immunoassays are problematic for several reasons. Compared with RDTs, which are cheap, easy to perform and give a same-day diagnosis, WB and line assays are expensive, require specialist equipment and personnel, and take much longer to confirm the patient's HIV status. WB and line assays are also more prone to inconclusive results. Both of these characteristics lead to delays in treatment and loss to follow-up.

Shifting away from WB and line assays in HIV programmes would facilitate faster initiation of ART and streamline the offer of pre-exposure prophylaxis. It would also support the expansion and decentralization of HIV testing services, including community-based testing.

Inclusion in the EDL negative list is intended to guide countries against procuring and using these tests.

G. SAGE IVD recommendations

SAGE IVD recommended against using HIV western blot and HIV line immunoassays for the diagnosis of HIV infection in adults, adolescents, children and infants over 18 months of age.

As such, the group recommended including both tests in the Do Not Do recommendations of the third EDL (EDL 3, Section III).

H. References

1. Consolidated guidelines on HIV testing services for a changing epidemic: policy brief. Geneva: World Health Organization; 2019 (WHO/CDS/HIV/19.31).
2. Conway DP, Holt M, McNulty A, Couldwell DL, Smith DE, et al. Multi-centre evaluation of the Determine HIV Combo assay when used for point of care testing in a high risk clinic-based population. PLoS One. 2014;8:e94062. doi:10.1371/journal.pone.0094062. Erratum in: PLoS One. 2014;9(7):e103399. doi:10.1371/journal.pone.0103399.
3. Martin EG, Salaru G, Paul SM, Cadoff EM. Use of a rapid HIV testing algorithm to improve linkage to care. J Clin Virol. 2011;52S:S11–5. doi:10.1016/j.jcv.2011.09.014.
4. Pando MA, Coloccini RS, Schvachsa N, Pippo M, Alfie LG, et al. Improved strategies for HIV diagnosis among men who have sex with men (MSM) in Buenos Aires, Argentina, a population with a high prevalence and incidence of HIV infection. HIV Med. 2012;13(9):564–567. doi:10.1111/j.1468-1293.2012.01011.x.
5. Shima-Sano T, Yamada R, Sekita K, Hankins RW, Hori H, et al. A human immunodeficiency virus screening algorithm to address the high rate of false-positive results in pregnant women in Japan. PLoS One. 2010;23:e9382. doi:10.1371/journal.pone.0009382.
6. Styer L, Sullivan T, Parker M. Evaluation of an alternative supplemental testing strategy for HIV diagnosis by retrospective analysis of clinical HIV testing data. J Clin Virol. 2011;52S:S35–S40. doi:10.1016/j.jcv.2011.09.009.
7. Walensky RP, Arbelaez C, Reichmann WM, Walls RM, Katz JN, et al. Revising expectations from rapid HIV tests in the emergency department. Ann Intern Med. 2008;149(3):153–160. doi:10.7326/0003-4819-149-3-200808050-00003.
8. Chen DJ, Yao JD. Comparison of turnaround time and total cost of HIV testing before and after implementation of the 2014 CDC/APHL Laboratory Testing Algorithm for diagnosis of HIV infection. J Clin Virol. 2017;91:69–72. doi:10.1016/j.jcv.2017.04.004. Erratum in: J Clin Virol. 2017;93:45. doi:10.1016/j.jcv.2017.06.003.
9. Delaney K, Rurangirwa J, Facente S, Dowling T, Janson M, et al. Using a multitest algorithm to improve the positive predictive value of rapid HIV testing and linkage to HIV care in nonclinical HIV test sites. J Acquir Immune Defic Syndr. 2016;71(1):78–86. doi:10.1097/QAI.0000000000000807.
10. Nasrullah M, Wesolowski L, Ethridge S, Cranston K, Pentella M, et al. Acute infections, cost and time to reporting of HIV test results in three U.S. state public health laboratories. J Infect. 2016;73(2):164–172. doi:10.1016/j.jinf.2016.05.006.
11. Obare F, Fleming P, Anglewicz P, Thornton R, Martinson F, et al. Acceptance of repeat population-based voluntary counselling and testing for HIV in rural Malawi. Sex Transm Infect. 2009;85(2):139–144. doi:10.1136/sti.2008.030320.
12. Stevinson K, Martin E, Marcell S, Paul SM. Cost effectiveness analysis of the New Jersey rapid testing algorithm for HIV testing in publicly funded testing sites. J Clin Virol. 2011;52 Suppl 1:S29–S33. doi:10.1016/j.jcv.2011.09.012.

6. Applications for reversing conditional listings in the EDL

EDL 2 includes nine IVDs that were listed conditionally, pending further evidence of test performance, clinical utility or acceptability. The original applicants of these IVDs were invited to submit further evidence and supporting data through a structured questionnaire. Five provided satisfactory additional evidence, which was assessed by previous SAGE IVD reviewers and discussed by SAGE IVD through the 2020 remote process for deliberations (see Section 1.2.2).

Original submissions are available in full at: https://www.who.int/groups/who-strategic-advisory-group-of-experts-on-in-vitro-diagnostics under the third meeting of the SAGE IVD.

HIV INFECTION

6.1. *Histoplasma* antigen test

A. EDL listing

The Histoplasma antigen test was conditionally listed in EDL 2 as an IVD to aid in the diagnosis of disseminated histoplasmosis, pending submission within 1 year of more evidence on the performance of the test.

B. Applicant

Global Action Fund for Fungal Infections

C. WHO technical department

Not applicable

D. Summary of evidence requested by SAGE IVD 2019

SAGE IVD requested the WHO technical department on HIV infection to include advice on use of the test in their guidelines.

E. Summary of evidence provided

A 2019 systematic review by Caceres et al. *(1)* and two primary studies by Caceres et al. (*2*, *3*) were provided as evidence for test performance.

The PAHO/WHO guidelines on histoplasmosis have been published (*4*).

F. Summary of SAGE IVD review

More papers have been published since the first submission in 2018. Highlights from these include the following:

- Falci and Pasqualotto (*5*) point to a lack of diagnostics and the importance of implementation.

- Samauyoa at al. (*6*) find that antigen *Histoplasma* test for urine is the most suitable for opportunistic infections in HIV but they require sensitivity and specificity.
- Persaud et al. (*7*) show that urine antigen tests are being used in routine laboratories.
- Falci et al. (*8*) show that *Histoplasma* urine antigen tests in Brazil increase the diagnosis of disseminated *Histoplasma* by 53%.
- Caceres et al. (*1*) find 95% sensitivity and specificity in antigen detection, with laboratory assays based on detecting circulating *Histoplasma* antigen having the best analytical performance.
- Torres-González et al. (*9*) find that *Histoplasma* urine antigen has a very high specificity and is easy to perform outside the laboratory, and thus is faster than other methods. But it lacks overall sensitivity when compared to centralized latter generations of the test. Still, likelihood ratios from the data show that HIV patients with progressive disseminated histoplasmosis are 27 times more likely to present a positive *Histoplasma* urine antigen test. In theory, this is a useful test for diagnosing disseminated histoplasmosis in patients with a high level of suspicion in endemic regions for *Histoplasma capsulatum*. The test may also favour early targeted antifungal treatment, improving the prognosis of these frail HIV patients. Importantly, co-infections are frequent, and a negative result does not discard progressive disseminated histoplasmosis.
- Caceres et al. (*3*) showed, in two laboratories in Latin America, high sensitivity and high specificity in commercial antigen-capture ELISA for diagnosing histoplasmosis in people living with HIV. Semiquantitative and quantitative ELISAs showed similar results. This finding could decrease testing costs, facilitating use of the IMMY monoclonal ELISA in LMICs.

G. SAGE IVD recommendation

SAGE IVD accepted the evidence provided and recommended reversing the conditional listing on the Histoplasma antigen test in EDL 3.

H. References

1. Caceres DH, Knuth M, Derado G, Lindsley MD. Diagnosis of progressive disseminated histoplasmosis in advanced HIV: a meta-analysis of assay analytical performance. J Fungi. 2019;5(3):E76. doi:10.3390/jof5030076.
2. Cáceres DH, Gómez BL, Tobón AM, Chiller TM, Lindsley MD. Evaluation of a histoplasma antigen lateral flow assay for the rapid diagnosis of progressive disseminated histoplasmosis in Colombian patients with AIDS. Mycoses. 2020;63(2):139–144. doi:10.1111/myc.13023.

3. Cáceres DH, Samayoa BE, Medina NG, Tobón AM, Guzmán BJ, et al. Multicenter validation of commercial antigenuria reagents to diagnose progressive disseminated histoplasmosis in people living with HIV/AIDS in two Latin American countries. J Clin Microbiol. 2018;56(6). doi:10.1128/JCM.01959-17.
4. Guidelines for diagnosing and managing disseminated histoplasmosis among people living with HIV. Washington DC: Pan American Health Organization and World Health Organization; 2020 (https://iris.paho.org/handle/10665.2/52304, accessed 16 July 2020).
5. Falci DR, Pasqualotto AC. Clinical mycology in Latin America and the Caribbean: a snapshot of diagnostic and therapeutic capabilities. Mycoses. 2019;62:368–373. doi:10.1111/myc.12890.
6. Samouya B, Aguirre L, Bonilla O, Medina N, Lau-Bonilla D, et al. The Diagnostic Laboratory Hub: a new health care system reveals the incidence and mortality of tuberculosis, histoplasmosis, and cryptococcosis of PWH in Guatemala. Open Forum Infect Dis. 2019;7(1):ofz534. doi:10.1093/ofid/ofz534.
7. Persaud SP, Lawton T, Burnham CAD, Anderson NW. Comparison of urine antigen assays for the diagnosis of Histoplasma capsulatum infection. J Appl Lab Med. 2019;4(3):370–382. doi:10.1373/jalm.2018.028910.
8. Falci DR, Monteiro AA, Caurio CFB, Magalhães TCO, Xavier MO, et al. Histoplasmosis, an underdiagnosed disease affecting people living with HIV/AIDS in Brazil: results of a multicenter prospective cohort study using both classical mycology tests and Histoplasma urine antigen detection. Open Forum Infect Dis. 2018;6(4):ofz073. doi:10.1093/ofid/ofz073.
9. Torres-González P, Niembro-Ortega MD, Martínez-Gamboa A, Ahumada-Topete VH, Andrade-Villanueva J, et al. Diagnostic accuracy cohort study and clinical value of the Histoplasma urine antigen (ALPHA Histoplasma EIA) for disseminated histoplasmosis among HIV infected patients: a multicenter study. PLoS Negl Trop Dis. 2018;12(11):e0006872. doi:10.1371/journal.pntd.0006872.

NEGLECTED TROPICAL DISEASES

6.2. Kato-Katz test for soil-transmitted helminths and intestinal schistosomes

A. EDL listing

The Kato-Katz test was conditionally listed in EDL 2 as an IVD for the surveillance and diagnosis of soil-transmitted helminthiasis and schistosomiasis caused by *Schistosoma mansoni*, *S. intercalatum*, *S. japonicum* and *S. mekongi*, pending submission of evidence of the test's validity.

B. Applicant

World Health Organization

C. WHO technical department

Control of Neglected Tropical Diseases

D. Evidence requested by SAGE IVD 2019

SAGE IVD requested evidence on the test's validity, applicability or precision, and asked to see comparisons with newer tests and evidence for their use.

E. Summary of evidence provided

A 2016 systematic review by Danso-Appiah et al. (*1*) looked at the accuracy of the Kato-Katz test for detecting schistosomiasis compared with a POC circulatory cathodic antigen (CCA) test.

A second 2016 systematic review by Kittur et al. (*2*) compared the prevalence and intensity of infection of *S. mansoni* detected by the CCA and Kato-Katz test and found that below 50% prevalence the CCA assay is much more sensitive than the Kato-Katz one. But the authors found that the data are inadequate to precisely define the relationship between CCA and Kato-Katz at lower levels of Kato-Katz prevalence. They concluded that more studies are needed directly comparing the two assays in low-prevalence areas to inform decision-making.

A 2019 study by Cools et al. (*3*) compared performance of the Kato-Katz test with two microscopy tests (MINI Flotac, FecPAKG2) and a quantitative PCR (qPCR) test. The study, which was mentioned in the previous submission, found that the Kato-Katz test performed better than newer microscopy tests but was not as sensitive as qPCR. The authors concluded that the PCR test should be used for decisions about discontinuing deworming programmes.

F. Summary of SAGE IVD review

A new systematic review by Cools et al. (*3*), which was not previously available, provides acceptable performance criteria for soil-transmitted helminths, and favourably compares Kato-Katz for surveillance and diagnosis to new tests.

G. SAGE IVD recommendation

SAGE IVD accepted the evidence provided and recommended reversing the conditional listing on the Kato-Katz test in EDL 3.

H. References

1. Danso-Appiah A, Minton J, Boamah D, Otchere J, Asmah RH, et al. Accuracy of point-of-care testing for circulatory cathodic antigen in the detection of schistosome infection: systematic review and meta-analysis. Bull World Health Organ. 2016;94(7):522–533A. doi:10.2471/BLT.15.158741.
2. Kittur N, Castleman JD, Campbell CH, King CH, Colley DG. Comparison of Schistosoma mansoni prevalence and intensity of infection, as determined by the circulating cathodic antigen urine assay or by the Kato-Katz fecal assay: a systematic review. Am J Trop Med Hyg. 2016;94(3):605–610. doi:10.4269/ajtmh.15-0725.
3. Cools P, Vlaminck J, Albonico M, Ame S, Ayana M, et al. Diagnostic performance of a single and duplicate Kato-Katz, Mini-FLOTAC, FECPAKG2 and qPCR for the detection and quantification of soil-transmitted helminths in three endemic countries. PLoS Negl Trop Dis. 2019;13(8):e0007446. doi:10.1371/journal.pntd.0007446.

PRIMARY IMMUNODEFICIENCIES

6.3. Enumeration of lymphocyte subtypes: CD4, CD8, CD20 and CD16/56 cells, B cells and NK cells

A. EDL listing

The enumeration of lymphocyte subtypes was conditionally listed in EDL 2 as an IVD to aid in the diagnosis of primary and secondary immunodeficiencies, pending submission of evidence for its clinical usefulness and use in LMICs within 1 year.

B. Applicant

International Patient Organisation for Primary Immunodeficiencies (IPOPI)

C. WHO technical department

Not applicable

D. Evidence requested by SAGE IVD 2019

SAGE IVD requested evidence on the test's use in LMICs as well as an evaluation of evidence that patients were treated with the associated EML drugs as a result of use of the test. It further requested evidence on the test's diagnostic accuracy or safety, and on guidelines for its use.

E. Summary of evidence provided

Evidence of clinical usefulness in LMICs was offered in a 2013 study of PID diagnosis in Mexico by Guaní-Guerra et al. (*1*), which described the status of PID diagnosis in the state of Guanajuato. IPOPI also did a survey on this use of the test in LMICs: 26 countries responded that the test is available and used; only two countries said that fewer than 50% of suspected PID cases received a test; 13 countries said that 80% or more of their suspected cases received this test.

Evidence of diagnostic accuracy was provided in two studies by the same authors in Europe (*2*, *3*), which show the performance of flow cytometric analysis against patients with genetically demonstrated PID.

Two examples of guidelines for using the test were provided:

- practice parameters for diagnosing and managing PIDs issued by a joint task force representing the American Academy of Allergy, Asthma & Immunology, the American College of Allergy, Asthma & Immunology, and the Joint Council of Allergy, Asthma & Immunology (*4*); and
- interdisciplinary guidelines by the Association of the Scientific Medical Societies in Germany (*5*).

The first of these (practice parameters) also provide a link to the EML by referring to listed antibiotics for recurring infection, and a broad set of other drugs.

F. Summary of SAGE IVD review

The resubmission contains more comprehensive (and more recent data) on the importance and utility of lymphocyte subtype analysis in diagnosing PID. Immunoglobulin analysis is a good diagnostic aid for PID, but it only identifies antibody-deficient immune deficiencies. The additional data from lymphocyte studies helps categorize these deficiencies but also identifies the important group of cellular deficiencies.

Oliveira and Fleisher (*6*) discuss laboratory evaluation, while Bonilla et al. (*4*) discuss practice parameters; both illustrate the importance and techniques of measurement. Two publications from the Euroflow consortium (*2*, *3*) show the consistency of flow cytometry in both screening for and diagnosis of PID and its importance in classifying the specific type of immunodeficiency so that appropriate treatment can be instituted. These measures do not have sensitivity and specificity measurements by their very nature, but there is concordance across laboratories.

Bainbridge et al. (*7*) similarly show consistency in measurement, with a high level of performance being maintained for more than a decade. The authors discuss the importance of combining education and awareness raising among general and specialist clinicians with up-to-date diagnostics.

Studies by the applying group and others show that access to flow cytometry in the diagnosis of PIDs is patchy across LMICs (*1*, *8*, *9*). This is unsatisfactory. Flow cytometry is now well established for other conditions and widely available in many LMICs and could be expanded to include PID studies; this is probably already the case in many countries. It may not be cost-effective to buy flow cytometers purely for PID studies (unless perhaps in a large specialist centre), but extending the use of equipment that is already available should be encouraged.

G. SAGE IVD recommendation

SAGE IVD accepted the evidence provided and recommended reversing the conditional listing on the enumeration of lymphocyte subtypes (CD4, CD8, CD20 and CD16/56 cells, B cells and NK cells) test in EDL 3.

H. References

1. Guaní-Guerra E, García-Ramírez UN, Jiménez-Romero AJ, Velázquez-Ávalos JM, Gallardo-Martínez G, et al. Primary immunodeficiency diseases at reference and high-specialty hospitals in the state of Guanajuato, Mexico. BioMed Res Int. 2013;187254. doi:10.1155/2013/187254.

2. van Dongen JJM, van der Burg M, Kalina T, Perez-Andres M, Mejstrikova E, et al. EuroFlow-based flow cytometric diagnostic screening and classification of primary immuno-deficiencies of the lymphoid system. Front Immunol. 2019;10:1271. doi:10.3389/fimmu.2019.01271.
3. van der Burg M, Kalina T, Perez-Andres M, Vlkova M, Lopez-Granados E, et al. The EuroFlow PID orientation tube for flow cytometric diagnostic screening of primary immunodeficiencies of the lymphoid system. Front Immunol. 2019;10:246. doi:10.3389/fimmu.2019.00246.
4. Bonilla FA, Khan DA, Ballas ZK, Chinen J, Frank MM, et al. Practice parameter for the diagnosis and management of primary immunodeficiency. J Allergy Clin Immunol. 2015;136(5):1186–1205. doi:10.1016/j.jaci.2015.04.049.
5. Farmand S, Baumann U, von Bernuth H, Borte M, Foerster-Waldl E, et al. Interdisciplinary AWMF guideline for the diagnostics of primary immunodeficiency. Klin Padiatr. 2011;223(6):378–385. doi:10.1055/s-0031-1287835.
6. Oliveira JB, Fleisher TA. Laboratory evaluation of primary immunodeficiencies. J Allergy Clin Immunol. 2010;125(2 Suppl 2):S297–S305. doi:10.1016/j.jaci.2009.08.043.
7. Bainbridge J, Rountree W, Louzao R, Wong J, Whitby L, et al. Laboratory accuracy improvement in the UK NEQAS Leucocyte Immuno-phenotyping Immune Monitoring Program: an eleven-year review via longitudinal mixed effects modeling. Cytometry B Clin Cytom. 2018;94(2):250–256. doi:10.1002/cyto.b.21531.
8. Pilania RK, Chaudhary H, Jindal AK, Rawat A, Singh S. Current status and prospects of primary immunodeficiency diseases in Asia. Genes Dis. 2019;7(1):3–11. doi:10.1016/j.gendis.2019.09.004.
9. Eley B, Esser M. Investigation and management of primary immunodeficiency in South African children. S Afr Med J. 2014;104(11):7304. doi:10.7196/SAMJ.8946.

6.4. Immunoglobulin plasma levels

A. EDL listing

The immunoglobulin plasma levels (IgG, IgA, IgM) test was conditionally listed in EDL 2 as an IVD to identify patients with low immunoglobulin levels and monitor replacement, pending submission of evidence of its clinical usefulness, with testing in various regions in field surveys, within 1 year.

B. Applicant

International Patient Organisation for Primary Immunodeficiencies (IPOPI)

C. WHO technical department

Not applicable

D. EDL edition listing

EDL 2, 2019

E. Evidence requested by SAGE IVD 2019

SAGE IVD requested more evidence of the test's use in LMICs, as well as an evaluation of evidence that patients were treated with the associated EML drugs

as a result of use of the test. It further asked for evidence on diagnostic accuracy or safety and on guidelines for its use.

F. Summary of evidence provided

Makhija et al. (*1*) and Sewell et al. (*2*) explain the clinical contexts where low immunoglobulins may be found and why.

Three guidelines on managing PIDs worldwide (*3*–*5*) provide evidence of the test's utility in LMICs, as do the 2014 global epidemiological estimates of PIDs published by Modell et al. (*6*).

Evidence on diagnostic accuracy includes UK NEQAS proficiency data (no reference provided) and package inserts for CE-marked tests for IgG, IgM and IgA with performance data.

The practice parameters for diagnosing and managing PIDs issued by a joint task force representing the American Academy of Allergy, Asthma & Immunology, the American College of Allergy, Asthma & Immunology, and the Joint Council of Allergy, Asthma & Immunology (*7*) provide a link to the EML by referring to antibiotics for recurring infection and a broad set of other drugs.

G. Summary of SAGE IVD review

The applicant indicated that no systematic review of analytical accuracy exists for these tests; but the organization has tried to provide an alternative through international external quality assessment data. IPOPI also justified the addition of these assays by citing international studies of immunodeficiencies and international guidelines developed for immunodeficiency identification and treatment, including in resource-limited environments. Treatment for immunoglobulin deficiencies is already listed in the EML: these tests are basic tests required for diagnosis.

H. SAGE IVD recommendation

SAGE IVD accepted the evidence provided and recommended reversing the conditional listing on the immunoglobulin plasma levels (IgG, IgM, IgA) test in EDL 3.

I. References

1. Makhija M, O'Gorman MRG. Common in vitro tests for allergy and immunology. Allergy Asthma Proc. 2012;33 Suppl 1:108–111. doi:10.2500/aap.2012.33.3564.
2. Sewell WAC, Khan S, Doré PC. Early indicators of immunodeficiency in adults and children: protocols for screening for primary immunological defects. Clin Exp Immunol. 2006;145(2):201–203. doi:10.1111/j.1365-2249.2006.03126.x.
3. Bonilla FA, Barlan I, Chapel H, Costa-Carvalho BT, Cunningham-Rundles C, et al. International Consensus Document (ICON): common variable immunodeficiency disorders. J Allergy Clin Immunol Pract. 2016;4(1):38–59. doi:10.1016/j.jaip.2015.07.025.

4. Chapel HM. Consensus on diagnosis and management of primary antibody deficiencies. BMJ. 1994;308(6928):581–585. doi:10.1136/bmj.308.6928.581.
5. Bousfiha A, Jeddane L, Ailal F, Benhsaien I, Mahlaoui N, et al. Primary immunodeficiency diseases worldwide: more common than generally thought. J Clin Immunol. 2013;33:1–7. doi:10.1007/s10875-012-9751-7.
6. Modell F, Knaus M, Modell F, Roifman C, Orange J, et al. Global overview of primary immunodeficiencies: a report from Jeffrey Modell Centers worldwide focused on diagnosis, treatment, and discovery. Immunol Res. 2014;60:132–144. doi:10.1007/s12026-014-8498-z.
7. Bonilla FA, Khan DA, Ballas ZK, Chinen J, Frank MM, et al. Practice parameter for the diagnosis and management of primary immunodeficiency. J Allergy Clin Immunol. 2015;136(5):1186–1205. doi:10.1016/j.jaci.2015.04.049.

6.5. Zika virus immunoglobulin M

A. EDL listing

The Zika virus IgM test was conditionally listed in EDL 2 as an IVD to aid in the diagnosis of suspected Zika virus (ZIKV) infection, pending provision within 1 year of further evidence.

B. Applicant

World Health Organization

C. WHO technical department

Health Emergencies, Infectious Hazards Management, High-threat Pathogens

D. EDL edition listing

EDL 2, 2019

E. Evidence requested by SAGE IVD 2019

SAGE IVD requested evidence of a clear recommendation in WHO guidelines for the use of these assays; and an updated review of the available data.

F. Summary of evidence provided

The CDC updated its guidelines on Zika testing in November 2019, specifying scenarios in which IgM tests should be used (*1*). WHO has yet to publish updated guidelines.

Several additional publications were provided as further evidence; but the only review article, submitted by Peters and Stevenson (*2*), paid little attention to serological testing, as acknowledged by the applicant. Most of the other publications submitted (*3*–*7*) highlight problems with the tests that were already identified in the original submission. They do, however, contain more data from commercial assays compared with the first submission, as well as

better characterized specimens. Sharp et al. (*8*) provide a good overview and review of dengue and Zika testing.

The availability of an international standard (*9*, *10*) is a great step forward, but further review revealed that it is not relevant to IgM assays; rather, it showed usefulness with neutralization assays. Nonetheless, there are some positive developments. The FDA has now cleared three IgM assays for sale in the USA, including two instrument-based ones and one plate-based EIA (which is a second version EIA that shows significant improvement in specificity over the first). The availability of these changes the landscape of Zika IgM IVD clearance in the USA in a way that may encourage other IVD manufacturers to pursue FDA clearance for their devices.

G. Summary of evidence evaluation by SAGE IVD reviewer

Evidence of a recommendation in WHO guidelines on when to use Zika IgM assays is still lacking. The CDC already had guidance recommending the use of IgM assays in certain scenarios when the conditional inclusion of Zika IgM assays was recommended; the fact that they have been updated does not fulfil the evidence request by SAGE IVD.

The applicants have fulfilled the request for an updated review of available data. Even though the publications submitted still point to concerns with the assays, the quality of these publications was acceptable. Previously there were few studies that allowed comparison between assays because they used different assays and characterized specimens in different ways. Assays and regulatory mechanisms also appear to be improving, which will lead to better quality diagnostics over time.

H. SAGE IVD recommendation

SAGE IVD accepted the evidence provided for an updated review of the available data; but the group noted that the WHO guidelines are still pending and so recommended maintaining the Zika virus IgM test as a conditional listing in EDL 3.

I. References

1. NEW Zika and dengue testing guidance. In: CDC/Testing guidance [website]. Atlanta: US Centers for Disease Control and Prevention; 2019 (https://www.cdc.gov/zika/hc-providers/testing-guidance.html, accessed 7 April 2020).
2. Peters R, Stevenson M. Zika virus diagnosis: challenges and solutions. Clin Microbiol Infect. 2019;25(2):142–146. doi:10.1016/j.cmi.2018.12.002.
3. Zaidi MB, Cedillo-Barron L, González Y Almeida ME, Garcia-Cordero J, Campos FD, et al. Serological tests reveal significant cross-reactive human antibody responses to Zika and dengue viruses in the Mexican population. Acta Trop. 2020;201:105201. doi:10.1016/j.actatropica.2019.105201.

4. Van der Beken Y, De Geyter D, Van Esbroeck M. Performance evaluation of the Diasorin LIAISON® XL Zika capture IgM CLIA test. Diagn Microbiol Infect Dis. 2019;95(2):144–148. doi:10.1016/j.diagmicrobio.2019.05.004.
5. Basile AJ, Ao J, Horiuchi K, Semenova V, Steward-Clark E, et al. Performance of InBios ZIKV Detect™ 2.0 IgM Capture ELISA in two reference laboratories compared to the original ZIKV Detect™ IgM Capture ELISA. J Virol Methods. 2019;271:113671. doi:10.1016/j.viromet.2019.05.011.
6. Matheus S, Talla C, Labeau B, de Laval F, Briolant S, et al. Performance of 2 commercial serologic tests for diagnosing Zika virus infection. Emerg Infect Dis. 2019;25(6):1153–1160. doi:10.3201/eid2506.180361.
7. Granger D, Theel ES. Evaluation of a rapid immunochromatographic assay and two enzyme-linked immunosorbent assays for detection of IgM-class antibodies to Zika virus. J Clin Microbiol. 2019;57(3). doi:10.1128/JCM.01413-18.
8. Sharp TM, Fischer M, Muñoz-Jordán JL, Paz-Bailey G, Staples JE, et al. Dengue and Zika virus diagnostic testing for patients with a clinically compatible illness and risk for infection with both viruses. MMWR Recomm Rep. 2019;68(1):1–10. doi:10.15585/mmwr.rr6801a1.
9. Fares-Gusmao R, Chancey C, Volkova E, Grinev A, Sippert E, et al. Production and characterization of Zika virus RNA reference reagents as a response to a public health emergency. Transfusion. 2018;58(9):2171–2174. doi:10.1111/trf.14870.
10. Mattiuzzo G, Knezevic I, Hassall M, Ashall J, Myhill S, et al. Harmonization of Zika neutralization assays by using the WHO International Standard for anti-Zika virus antibody. NPJ Vaccines. 2019;4:42. doi:10.1038/s41541-019-0135-3.

7. Summary and recommendations

SAGE IVD advises the Director-General of WHO in the area of IVDs. It provides technical advice on global policies and strategies related to priority, essential and neglected IVDs, and oversees maintenance of the EDL. The EDL is developed through a broad consensus-building process, including expert review and public consultation.

SAGE IVD's third meeting, which was scheduled for 23–27 March 2020, was cancelled because of the global COVID-19 pandemic. Instead of meeting face-to-face, SAGE IVD members conducted their business online, through email exchanges and a series of biweekly virtual meetings. Acknowledging that this way of working cannot hope to effectively replace the face-to-face SAGE IVD meeting, members agreed to advance their work as far as possible while remaining realistic about what they can achieve and without compromising the quality of its decision-making.

The bulk of the SAGE IVD virtual sessions were dedicated to reviewing submissions to EDL 3 and agreeing the test purpose, format and setting for each one recommended for inclusion.

The group also discussed the development and use of the EDL so far, and considered the list's future, making recommendations on issues such as how to improve the quality of submissions, how to link the list to the EML and other selection tools, and how best to support countries to both contribute to and implement the EDL effectively. SAGE IVD also made its usual recommendations on priorities for next year's list.

7.1. IVDs adopted for EDL 3

In total, SAGE IVD considered:

- 26 applications for the addition of new IVD categories to the EDL;
- 26 applications for other revisions to the EDL;
- 5 applications to reverse the conditional listing of specific IVD categories in the EDL; and
- 2 applications for Do Not Do recommendations that advise countries against the procurement and use of specific IVD categories.

Following review and deliberation of each application, SAGE IVD adopted several IVD categories for inclusion in EDL 3, as listed below.

7.1.1. Additions and revisions

Section I.a: General IVDs for community settings and health facilities without laboratories

- Erythrocyte sedimentation rate (ESR) (*new*)

Section I.b: Disease-specific IVDs for community settings and health facilities without laboratories

Sickling disorders

- Sickle cell testing, rapid diagnostic test format (*new*)

Streptococcal pharyngitis

- Group A *Streptococcus* antigen, rapid diagnostic test format (*new*)

Neglected tropical diseases

- *Trypanosoma cruzi* immunoglobulin G antibody, rapid diagnostic test format (*new*)

Tuberculosis

- Lipoarabinomannan antigen (*revision*)

Section II.a: General IVDs for clinical laboratories

Clinical chemistry

- Creatinine (*revision*)
- C-reactive protein (*revision*)
- Glucose-6-phosphate dehydrogenase activity (*revision*)

Haematology

- Cerebrospinal fluid profile (*new*)

Section II.b: Disease-specific IVDs for clinical laboratories

Cancer

- Alpha-fetoprotein immunoassay (*revision*)
- Epidermal growth factor receptor (*new*)
- Immunohistochemistry for solid tumours (*revision*)

- Immunohistochemistry for lymphoma (*revision*)
- Essential flow cytometry for leukaemia (*revision*)
- Lactate dehydrogenase activity (*revision*)

COVID-19

- SARS-CoV-2 antigen detection test (*new*)
- SARS-CoV-2 nucleic acid test (*new*)

Endocrine disorders (new subsection)

- Cortisol (*new*)
- Estradiol (*new*)
- Follicle-stimulating hormone (*new*)
- Luteinizing hormone (*new*)
- Progesterone (*new*)
- Prolactin (*new*)

HIV infection

- *Histoplasma* antigen (*reversal of conditional listing*)

Human papillomavirus (HPV) infection

- Human papillomavirus nucleic acid test (*revision*)

Aspergillosis

- *Aspergillus* antigen (*new*)
- *Aspergillus* immunoglobulin G antibody (*new*)

Pneumocystis pneumonia

- *Pneumocystis jirovecii* nucleic acid test (*new*)

Neglected tropical diseases

- Kato-Katz faecal smear (*reversal of conditional listing*)
- *Trypanosoma cruzi* immunoglobulin G antibody, immunoassay (*new*)
- Visceral leishmaniasis direct agglutination test (*new*)

Primary immunodeficiencies

- Immunoglobulin plasma levels (IgG, IgA, IgM) (*reversal of conditional listing*)
- Lymphocyte subtype enumeration (*reversal of conditional listing, and revision*)

Vaccine-preventable diseases

- Measles immunoglobulin G (*new, conditional*)
- Measles immunoglobulin M (*new*)
- Measles nucleic acid test (*new*)
- Rubella immunoglobulin M antibody (*new*)
- Rubella immunoglobulin G antibody (*new*)

Streptococcal pharyngitis

- Group A *Streptococcus* antigen (*new*)

Section II.c: Disease-specific IVDs for blood screening laboratories

Other transfusion-transmitted organisms

- IVDs to screen for pathogens in blood donations (*revision*)

7.1.2. Do Not Do recommendations

In line with the latest WHO guidelines, SAGE IVD issued three Do Not Do recommendations for inclusion in the EDL, specifically advising countries against the procurement and use of:

- Western blot assays to diagnose HIV infection
- Line immunoassays to diagnose HIV infection
- Commercial serodiagnostic tests to diagnose TB

7.1.3. Rejections

After careful review, SAGE IVD rejected three applications for additions to the EDL:

- Group A *Streptococcus* nucleic acid test as an IVD to diagnose streptococcal pharyngitis
- Serum and urine protein electrophoresis as an IVD to diagnose multiple myeloma

- Serum and urine protein electrophoresis as an IVD for diagnosing primary immunodeficiencies

SAGE IVD rejected 11 applications for revisions to the following tests:

- Basic panel of immunohistochemical markers for diagnosis of solid tumours
- Complete blood count automated
- Estrogen and progesterone receptors
- Essential flow cytometry panel of antibodies for leukaemia
- Hepatitis B e antigen
- Qualitative and quantitative HIV virological nucleic acid test
- Qualitative test for *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infections
- Quantitative HBV virological nucleic acid tests
- *Mycobacterium tuberculosis* DNA
- Tyrosine-protein kinase receptor or human epidermal growth factor receptor 2 overexpression

SAGE IVD also rejected an application to reverse the conditional listing on:

- Zika virus immunoglobulin M

7.2. IVDs recommended for EDL 4

7.2.1. New submissions

As part of its IVD-specific deliberations, SAGE IVD issued four requests for full submissions for consideration as additions to EDL 4 in 2021:

- Serum and urine protein electrophoresis, immunofixation and serum-free light-chain test for multiple myeloma
- Testosterone

Individual group members further identified a number of tests as potentially useful additions to the EDL, subject to the usual evidence evaluation and review. These included tests for:

- acute tropical infections, such as scrub typhus (IgM) and leptospirosis
- autoimmune diseases
- blood disorders, such as anaemia (peripheral smear)
- cancers, for example myeloma tests and tumour markers (e.g. ca-125)

- chronic infections, including brucellosis, melioidosis (POC test on fluids)
- drugs of abuse (tox)
- emergencies, general technologies for preparedness
- intestinal parasites, such as stool microscopy saline and iodine mounts
- muscle disorders
- neglected tropical diseases
- renal diseases (urine microscopy)
- rheumatoid diseases (rheumatoid factor)
- thyroid diseases, including free T4 and free T3
- additional vaccine-preventable diseases (typhoid, rotavirus, diphtheria, pertussis, yellow fever, polio, pneumonia,)
- additional general IVDs, including for coagulation, platelet aggregation, bleeding time, whole blood clotting time and clot retraction time.

The group further recommended reviewing all entries in EDL 2 and adding CSF profile as a specimen type as and where relevant in the next edition of the EDL (EDL 4).

7.2.2. Additional evidence

SAGE IVD requested additional evidence or information to be provided on four tests that it recommended for inclusion in EDL 3:

- Measles immunoglobulin G antibody (conditional listing): further evidence required to show that the test has utility in confirming cases when used with IgM.
- Rubella immunoglobulin G antibody (full listing): further evidence requested to enable an additional test purpose for diagnosing acute infection.
- Luteinizing hormone (full listing): further information requested on the applications of LH in urine to support an edit in EDL 4.
- Estradiol (full listing): further information requested on the applications of estradiol in urine to support an edit in EDL 4.

7.3. General recommendations

SAGE IVD made a series of general recommendations related to EDL products, processes and strategy, as listed below. Progress on the suggested actions will be reported to the SAGE IVD at its 4th meeting in 2021.

7.3.1. EDL format

- Include a conflict of interest statement in the submission form.
- Consider adding a notes column to the EDL, for comments or qualifications of importance.
- Review Section II.c to include donated organs.

7.3.2. EDL processes

- Extend the time allowed for experts to serve as members to the SAGE IVD.
- Ensure a participatory approach to developing the EDL that provides partners and countries with the information they need to make a contribution and engages more stakeholders in disseminating the EDL and encouraging its use.
- Vet applications more rigorously early on in the submission process to ensure tests are ultimately judged on test utility, not submission quality.
- Increase the focus on cost–effectiveness during SAGE IVD deliberations.
- Establish a formal methodology for reviewing specialist tests when there are gaps in expertise within the SAGE IVD.
- Develop a process for "fast-track review" to enable rapid review of innovative tests in between annual SAGE IVD sessions to tackle emergencies (e.g. COVID-19 diagnostics).

7.3.3. EDL-related products

- Continue development of eEDL, with a focus on making it is easy to use and informative for decision-makers.
- Finalize guidance for country implementation of the EDL.
- Finalize technical specifications to support procurement of specific products in line with the EDL.
- Continue working to standardize nomenclature within the EDL and harmonize it with ICD-11.

- Continue contributing to WHO's UHC compendium, identifying IVD tests for each intervention and disease in the database.
- Consider developing a companion tool to the EDL on rational use of diagnostics.

7.3.4. EDL strategy

- Establish an approach to prioritization that can support country implementation of the EDL, including considering the value of developing a core and complementary list like the EML does.
- Review the structure and scope of the EDL, including evaluating the pros and cons of mirroring the EML structure and the added value of including non-in vitro diagnostics and population-level diagnostics, for example for fluoride, arsenic, air quality and bacterial contamination.

Five working groups were established to follow up the recommendations on EDL-related products and strategy.

Annex 1

The third WHO model list of essential in vitro diagnostics (EDL 3)

Third Model List of Essential In Vitro Diagnostics (EDL)

The EDL is presented by health care facility level in two tiers:

I. Community and health settings without laboratories, with two sections:
   a. General IVDs for community and health settings without laboratories
   b. Disease-specific IVDs for community and health settings without laboratories

II. Health care facilities with clinical laboratories, with three sections:
   a. General IVDs for clinical laboratories
   b. Disease-specific IVDs for clinical laboratories
   c. Disease-specific IVDs for blood screening laboratories

III. Do Not Do recommendations

Note: The specimen types listed for each diagnostic test category comprise possible specimens for that category; however, not all test brands within each category will be validated for all the specimen types listed. It is assumed that products within each category will be used strictly in accordance with the manufacturer's instructions for use.

Assay formats listed are generic and may be based on different methodologies, e.g. immunoassays are available in various forms – manual microplate assays and automated platforms – with various types of chemical detection (e.g. turbidimetry, chemiluminescence and electrochemiluminescence assays). Selection of products and assay formats should be based on local processes.

I. Community settings and health facilities without laboratories

These lists contain tests for community settings and health facilities that include health posts and centres, doctors' offices, outreach clinics, ambulatory care and home-based and self-testing. If laboratory facilities are available in community settings, please refer to the IVDs described in Section II. If laboratory facilities are not available, specimens may be collected, transported to and processed at a higher tier of the health system. The tests in this section of the EDL are also assumed to be available, in combination with the extended list in Section II, at health care facilities with laboratories, though assay formats may vary.

I.a General IVDs for use in community settings and health facilities without laboratories

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Blood typing	A, B and O blood groups and Rhesus (Rh) factor	To determine A, B and O groups and Rh type	Slide agglutination test	Capillary whole blood Venous whole blood [1]
Clinical chemistry	Albumin	To detect or monitor kidney disease	Dipstick	Urine
	Bilirubin	To detect or monitor liver disease and bile duct disorders	Dipstick	Urine
	Glucose	To diagnose hypoglycaemia	Glucose meter	Capillary whole blood
	Ketones	To aid in the diagnosis of ketosis, e.g. in uncontrolled diabetes, starvation, pregnancy or in diabetic ketoacidosis	Dipstick	Urine
	Urinalysis test strips	To aid in the diagnosis of urinary tract infections (UTIs) by detection of leukocyte esterase derived from white blood cells or nitrites as a result of bacteria in urine	Dipstick	Urine

[1] If a phlebotomist is available.

I.a General IVDs for use in community settings and health facilities without laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Haematology	Erythrocyte sedimentation rate (ESR)	To detect inflammation as an indicator of various conditions when C-reactive protein (CRP) is not available	Westergren	Venous whole blood[2]
	Haemoglobin (Hb)	• To diagnose and monitor anaemia • To monitor the safety of certain drugs (e.g. zidovudine for HIV infection) • To screen potential blood donors	Haemoglobinometer	Capillary whole blood Venous whole blood[2]
		• Clinical marker for certain severe infections (e.g. malaria, viral haemorrhagic fevers) • To aid in the diagnosis of intravascular haemolysis, renal conditions, rhabdomyolysis (myoglobinuria)[3]	Dipstick	Urine
Pregnancy testing	Human chorionic gonadotrophin (hCG)	To aid in the early detection of pregnancy	Rapid diagnostic test (RDT) (dipstick and cassette), latex agglutination	Urine (early morning)

[2] If a phlebotomist is available.

[3] This test does not differentiate between myoglobin and haemoglobin.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Chagas disease	*Trypanosoma cruzi* IgG antibody	• For surveillance of *T. cruzi* infection • To screen girls, women of childbearing age and pregnant women without previous treatment for *T. cruzi* infection • To screen children and other at-risk populations • To aid in the diagnosis of chronic *T. cruzi* infection (Chagas disease) (only in settings where laboratory-based methods are not available)	RDT[4]	Capillary whole blood Venous whole blood[5] Serum	N/A	Guidelines for the diagnosis and treatment of Chagas disease (2018) https://www.who.int/publications/i/item/9789275120439 https://www.who.int/health-topics/chagas-disease#tab=tab_1

[4] A negative result does not exclude infection.

[5] If a phlebotomist is available.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cholera	*Vibrio cholerae* antigen	For initial detection or exclusion of a cholera outbreak (Not for use in case management)	RDT	Stool Rectal swab	N/A	Interim technical note: the use of cholera rapid diagnostic tests (2016) https://www.who.int/health-topics/cholera#tab=tab_1

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Coronavirus disease (COVID-19)	SARS-CoV-2 antigen[6]	To diagnose COVID-19 in settings where NAT is unavailable or where prolonged turnaround times preclude clinical utility[7] To aid in the diagnosis of COVID-19 in the early symptomatic phases of illness, or in asymptomatic individuals with known contact with a confirmed case[7]	RDT Handheld or small benchtop instrument for POC use	Upper respiratory specimens (e.g. nasopharyngeal or nasal swab)	Emergency Use Listing (EUL) https://extranet.who.int/pqweb/vitro-diagnostics/coronavirus-disease-covid-19-pandemic-%E2%80%94-emergency-use-listing-procedure-eul-open	Antigen-detection in the diagnosis of SARS-CoV-2 infection using rapid immunoassays Interim guidance (11 September 2020) https://www.who.int/publications/i/item/antigen-detection-in-the-diagnosis-of-sars-cov-2infection-using-rapid-immunoassays Guidance on SARS CoV-2 testing is reviewed regularly based on available evidence. For up to date guidance, see: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance-publications?publicationtypes=f85a3610-b102-4287-a6df-f3bc0b2e9f7c

[6] Listing was based on available evidence and interim WHO guidelines which are subject to change. Regulatory oversight of most commercially available tests was limited to emergency use authorizations at the time of listing.

[7] A negative test does not rule out infection and should not determine clinical care.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Diabetes mellitus	Glucose	• To aid in the diagnosis of diabetes mellitus (if blood glucose testing is not available); to screen for type 2 diabetes mellitus[8]	Dipstick	Urine	N/A	HEARTS-D: diagnosis and management of type 2 diabetes (2020) https://www.who.int/publications/i/item/who-ucn-ncd-20.1 https://www.who.int/health-topics/diabetes#tab=tab_1
		• To diagnose and monitor[9] type 1 and type 2 diabetes mellitus • To diagnose impaired fasting glucose/impaired glucose tolerance • To screen for type 2 diabetes mellitus and impaired fasting glucose/ impaired glucose tolerance	Glucose meter	Capillary whole blood		

[8] Screening for diabetes is possible with various methods, both biochemical and non-biochemical, of varying sensitivity and specificity. Urine glucose testing can be used if the aim of screening is to identify people who have more severe hyperglycaemia within the diabetes spectrum.

[9] If HbA1c testing is not available.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Diabetes mellitus *continued*	Haemoglobin A1c (HbA1c)	To diagnose and monitor diabetes mellitus	Handheld and small analysers	Capillary whole blood	N/A	HEARTS-D: diagnosis and management of type 2 diabetes (2020) https://www.who.int/publications/i/item/who-ucn-ncd-20.1 https://www.who.int/health-topics/diabetes#tab=tab_1
Hepatitis B virus (HBV) infection	Hepatitis B surface antigen (HBsAg)	To screen for HBV infection, or to aid in the diagnosis of chronic and acute HBV infection: infants > 12 months of age, children, adolescents and adults	RDT	Capillary whole blood Venous whole blood [10]	Public reports of WHO-prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=63	Guidelines on hepatitis B and C testing (February 2017) https://apps.who.int/iris/handle/10665/254621 https://www.who.int/news-room/fact-sheets/detail/hepatitis-b

[10] If a phlebotomist is available.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Hepatitis B virus (HBV) infection *continued*	Hepatitis B e antigen (HBeAg)	Staging to assess the need for HBV treatment in chronic HBV infection and as a criterion for use of antivirals in the mother to prevent mother-to-child transmission (if HBV DNA test is not available) (Only for use where an HBV DNA test is not available)	RDT	Capillary whole blood Venous whole blood[11]	N/A	
Hepatitis C virus (HCV) infection	Antibodies to hepatitis C virus (HCV) (anti-HCV)	To screen for or to aid in the diagnosis of viraemic HCV infection: infants > 18 months of age, children, adolescents and adults	RDT	Oral fluid Capillary whole blood Venous whole blood[11]	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=59	Guidelines on hepatitis B and C testing (February 2017) https://apps.who.int/iris/handle/10665/254621 https://www.who.int/news-room/fact-sheets/detail/hepatitis-c

[11] If a phlebotomist is available.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
HIV infection	HIV 1/2 antibody (anti-HIV Ab)	HIV self-testing to aid in the diagnosis of HIV infection	RDT	Oral fluid Capillary whole blood	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=60	Consolidated guidelines on HIV testing services (December 2019) https://www.who.int/publications/i/item/978-92-4-155058-1 Guidelines on HIV self-testing and partner notification (2016) https://apps.who.int/iris/handle/10665/251655 WHO implementation tool for pre-exposure prophylaxis (PrEP) of HIV infection, module 10 for testing providers (2017) https://apps.who.int/iris/handle/10665/258516 WHO HIV molecular diagnostics toolkit to improve access to viral load testing and infant diagnosis: Toolkit (July 2019) https://apps.who.int/iris/handle/10665/325961/ https://www.who.int/health-topics/hiv-aids/#tab=tab_1
		To screen for or to aid in the diagnosis of HIV infection: adults, adolescents, children and infants > 18 months of age	RDT	Oral fluid Capillary whole blood Venous whole blood[12]	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=58	
	Combined HIV antibody/p24 antigen (anti-HIV/p24 Ag)	To screen for or to aid in the diagnosis of HIV infection: adults, adolescents, children and infants > 8 months of age	RDT	Capillary whole blood Venous whole blood[12]		

[12] If a phlebotomist is available.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
HIV infection *continued*	Qualitative HIV nucleic acid test (NAT)	To diagnose HIV infection in infants < 18 months of age	Point-of-care NAT	Capillary whole blood Venous whole blood[13] Dried blood spots (DBS)	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=61	
	CD4 cell enumeration	• For staging of advanced HIV disease • To monitor response to antiretroviral therapy (in settings where HIV viral load is not available)	Point-of-care flow cytometry platform	Capillary whole blood Venous whole blood[13]	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=66	Consolidated guidelines on HIV testing services (December 2019) https://www.who.int/publications/i/item/978-92-4-155058-1 Guidelines for managing advanced HIV disease and rapid initiation of antiretroviral therapy (2017) https://apps.who.int/iris/handle/10665/255884/ https://www.who.int/health-topics/hiv-aids/#tab=tab_1

[13] If a phlebotomist is available.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*						
Disease	**Diagnostic test**	**Test purpose**	**Assay format**	**Specimen type**	**WHO prequalified or recommended products**	**WHO supporting documents**
HIV infection *continued*	Cryptococcal antigen	To screen for and diagnose cryptococcal meningitis in people with advanced HIV disease	RDT	Capillary whole blood Venous whole blood[14]	N/A	Guidelines for the diagnosis, prevention, and management of cryptococcal disease in HIV-infected adults, adolescents and children (2018) https://apps.who.int/iris/handle/10665/260399 Guidelines for managing advanced HIV disease and rapid initiation of antiretroviral therapy (2017) https://apps.who.int/iris/handle/10665/255884

[14] If a phlebotomist is available.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Influenza	Influenza A and B antigen	To aid in the diagnosis of seasonal influenza infection (Not recommended for surveillance testing)	RDT Instrument-based point-of-care immunoassay	Nasal swab Nasopharyngeal swab Nasopharyngeal aspirate or wash	N/A	Use of influenza rapid diagnostic tests (2010) https://apps.who.int/iris/handle/10665/44304/ Manual for the laboratory diagnosis and virological surveillance of influenza (2011) https://apps.who.int/iris/handle/10665/44518 Global Epidemiological Surveillance Standards for Influenza (2014) https://www.who.int/influenza/resources/documents/WHO_Epidemiological_Influenza_Surveillance_Standards_2014.pdf https://www.who.int/health-topics/influenza-seasonal#tab=tab_1
	Influenza A and B nucleic acid test	To diagnose seasonal influenza infection	Point-of-care NAT	Nasal swab Nasopharyngeal swab Nasopharyngeal aspirate or wash	N/A	

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Malaria	*Plasmodium* spp. antigens; species-specific (e.g. HRP2) and/or pan-species specific (e.g. pan-pLDH)	To diagnose one or more human malaria parasite species (*P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*)	RDT	Capillary whole blood Venous whole blood[15]	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=64	WHO guidelines for the treatment of malaria, third edition (2015) https://apps.who.int/iris/handle/10665/162441 Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 8 (2016–2018) https://apps.who.int/iris/handle/10665/276190 WHO good practices for selecting and procuring rapid diagnostic tests for malaria (2011) https://apps.who.int/iris/handle/10665/44530 Information note on recommended selection criteria for procurement of malaria rapid diagnostic tests https://apps.who.int/iris/handle/10665/259870 https://www.who.int/health-topics/malaria#tab=tab_1

[15] If a phlebotomist is available.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Streptococcal pharyngitis	Group A *Streptococcus* antigen	To aid in the diagnosis of Group A streptococcal pharyngitis	RDT	Throat swab	N/A	N/A
Sickling disorders	Sickle cell testing	To screen for or to aid in the diagnosis of sickle cell disease, C trait (SCT) and other variant sickling disorders	RDT	Capillary whole blood Venous whole blood[16]	N/A	N/A
Syphilis	Antibodies to *Treponema pallidum*	To diagnose or to aid in the diagnosis of *T. pallidum*	RDT	Capillary whole blood Venous whole blood[16]	N/A	WHO laboratory diagnosis of sexually transmitted infections, including human immunodeficiency virus (2013) https://apps.who.int/iris/handle/10665/85343 https://www.who.int/news-room/fact-sheets/detail/sexually-transmitted-infections-(stis)
	Combined antibodies to *T. pallidum* and HIV-1/2	To diagnose or to aid in the diagnosis of HIV-1/2 and/or *T. pallidum*	RDT	Capillary whole blood Venous whole blood[16]	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=57	Consolidated guidelines on HIV testing services (December 2019) https://www.who.int/publications/i/item/978-92-4-155058-1

[16] If a phlebotomist is available.

I.b Disease-specific IVDs for use in community settings and health facilities without laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products [17]	WHO supporting documents
Tuberculosis (TB)	Tuberculin skin test (TST) (Mantoux test)	To diagnose latent TB infection	Intradermal test	N/A		Latent TB infection: updated and consolidated guidelines for programmatic management (2018) https://apps.who.int/iris/handle/10665/260233 https://www.who.int/health-topics/tuberculosis#tab=tab_1
	Lipoarabino-mannan (LAM) antigen	To aid in the diagnosis of TB in seriously ill HIV-positive inpatients and in HIV-positive adult outpatients with signs and symptoms of TB	RDT	Urine	The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV: policy update (2015) http://apps.who.int/iris/handle/10665/193633	Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV: policy update 2019 https://apps.who.int/iris/handle/10665/329479 https://www.who.int/health-topics/tuberculosis#tab=tab_1
Visceral leishmaniasis (kala-azar)	Recombinant K39 (rK39) antigen	To aid in the diagnosis of clinically suspected visceral leishmaniasis	RDT	Serum [18] Capillary whole blood Venous whole blood [18]	N/A	WHO Technical Report Series 949 https://apps.who.int/iris/handle/10665/44412 https://www.who.int/teams/control-of-neglected-tropical-diseases

[17] All TB tests are evaluated and guidelines developed by the WHO global TB programme.

[18] If a phlebotomist is available.

II. Health care facilities with clinical laboratories

These lists contain additional tests for district, regional, provincial or specialized hospitals or laboratories, and national reference laboratories. It is assumed that trained laboratory technologists, specialist expertise, and laboratory infrastructure and equipment are available at the appropriate level. All diagnostic tests available in community settings and health facilities as described in Section I are assumed to be available at higher levels, as appropriate. The list comprises sections for:

a. General IVDs for use in clinical laboratories
b. Disease-specific IVDs for use in clinical laboratories
c. Disease-specific IVDs for blood screening laboratories

II.a General IVDs for use in clinical laboratories

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type	WHO supporting documents
Anatomical pathology[19]	Histopathology	To assess tissue for infection, neoplasia, inflammatory and degenerative disorders	Macroscopic assessment of tissue and selection of areas for microscopic examination Microscopy of tissue sections mounted on slides and stained most commonly with haematoxylin and eosin in the first instance, then treated with a variety of special stains, selected case by case to identify pathogens and other abnormal features	Surgical resection Biopsy Core biopsy Cell block	WHO priority medical devices for cancer management https://apps.who.int/iris/handle/10665/255262 Basic histopathology and anatomical pathology services for developing countries with variable services https://apps.who.int/iris/handle/10665/119675 WHO Guide for establishing a pathology laboratory in the context of cancer control https://www.who.int/publications/i/item/guide-for-establishing-a-pathology-laboratory-in-the-context-of-cancer-control
	Cytology (cytopathology)	To assess cells for infection, neoplasia, inflammatory and degenerative disorders	Microscopy of stained cells on slides	Cervical specimen for Papanicolaou (Pap) smear Body fluids: e.g. cerebrospinal fluid (CSF), urine, pleural and peritoneal fluids, fine-needle aspirate (FNA) of lymph nodes, spleen and other tissues, bone marrow aspirate, sputum, bronchial brushings, bronchoalveolar lavage (BAL) and skin samples	

[19] Note: The tests described in this section require specialized anatomical pathology laboratories and trained anatomical pathologists.

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type	WHO supporting documents
Anatomical pathology[20] *continued*	Immunohisto-chemistry (IHC)	To assess cells for specific markers to identify infection, neoplasia, inflammatory and degenerative disorders	Microscopy of histopathology tissue sections mounted on slides and stained with antibodies to specific markers Refer to EDL sections on disease-specific tests for individual assays	Surgical resection Biopsy Core biopsy Cell block	
	Post-mortem examination	To determine the cause of death and correlation with pre-mortem clinical features and investigations	Macroscopic assessment and microscopy of tissue sections. Procedures selected case by case	Tissue from cadaver	International guidelines for the determination of death – Phase I https://www.who.int/patientsafety/montreal-forum-report.pdf

[20] Note: The tests described in this section require specialized anatomical pathology laboratories and trained anatomical pathologists.

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Clinical microbiology	Staining procedures	For the presumptive identification of pathogens and for determination of microbial morphology	Microscopic examination of slides which may use different types of microscopes and stains	Disease-appropriate specimens (e.g. sputum, venous whole blood, urine, stool, body fluids, cerebrospinal fluid or cultures)
	Culture	Initial step in detection and identification of bacterial and fungal species for selection of appropriate antimicrobial regimens	Culture on growth media plates or broth in an incubator followed by recovery of isolates and species identification (traditional manual techniques or automated equipment)	Disease-appropriate specimens (e.g. urine, stool, sputum, body fluids, e.g. cerebrospinal fluid, etc.)
	Blood culture	To detect bacterial and fungal bloodstream infections (sepsis)	Blood culture bottle in an incubator followed by recovery of isolates (traditional manual techniques or automated equipment)	Venous whole blood
	Genus and species identification of bacteria and fungi	To identify the genus or species of bacteria or fungi from microbial isolates	A range of biochemical tests that may be performed manually or on automated equipment	Microbial isolates

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Clinical microbiology *continued*	Antimicrobial susceptibility testing	Final step in selection of appropriate antibiotics after species identification and interpretation by EUCAST[21] and CLSI guidelines[22] Note: WHO regards the development of antimicrobial resistance (AMR) a high-priority global health issue. See WHO Global Antimicrobial Resistance Surveillance System (GLASS): https://www.who.int/activities/facilitating-global-surveillance-of-antimicrobial-resistance	Antimicrobial susceptibility testing of isolates may be done manually (by disc diffusion, gradient tests and broth microdilution), or by automated platforms	Microbial isolates

[21] EUCAST, European Committee on Antimicrobial Susceptibility Testing: Breakpoint tables for interpretation of MICs and zone diameters. Version 9.0.

[22] CLSI, Clinical and Laboratory Standards Institute: CLSI M100 performance standards for antimicrobial susceptibility testing, 29th edition.

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Clinical chemistry[23]	**Tests under this subsection may be requested together as part of a liver profile**[24]			
	Alanine aminotransferase (ALT)	To aid in the diagnosis of liver disease as a marker of liver injury.	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma
	Aspartate aminotransferase (AST)	To aid in the diagnosis of liver disease as a marker of liver injury.	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma
	Albumin	To aid in the diagnosis and monitoring of diseases affecting protein metabolism (synthesis, loss, intake, absorption) e.g. liver disease, kidney disease, severe malnutrition, malabsorption, burns, etc.)[25]	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma
	Alkaline phosphatase (ALP)	To aid in the diagnosis of hepatobiliary and bone disorders	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma
	Gamma-glutamyl transferase (GGT)	• To assess hepatobiliary function • To distinguish between bone and hepatobiliary causes of raised ALP	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma

[23] Combinations of tests of clinical chemistry are often referred to as basic and comprehensive metabolic panels. However, the content of these may vary according to laboratory and health care system resources, disease profiles and patient needs.

[24] Often termed in common usage as "liver function tests", which may include other tests such as tests of coagulation, like prothrombin time. These individual tests, however, reflect different disease processes in the liver and do not all assess liver synthetic, metabolic or excretory function.

[25] Serum albumin levels may also decrease acutely in systemic inflammation, e.g. sepsis.

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Clinical chemistry *continued*	Globulin	To determine the globulin fraction levels which may indicate underlying infections, chronic inflammatory diseases or haematologic malignancies[26]	Calculation	N/A
	Total protein	To measure total protein in blood and body fluids	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma Body fluids
	Total bilirubin	To detect hyperbilirubinemia as an aid in diagnosis and monitoring of diseases of the liver, biliary duct or pancreas, haemolysis, and other causes like neonatal hyperbilirubinemia	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma
	Direct bilirubin and indirect bilirubin	To measure direct (conjugated) bilirubin and to estimate indirect (unconjugated) bilirubin as an aid in the differential diagnosis of hyperbilirubinaemia	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma
	Kidney function tests (tests under this subsection may be requested as part of a panel that will vary depending on the context)			
	Albumin	To monitor kidney function	Optical methods on semi-automated or automated chemistry analysers	Urine
	Blood urea nitrogen (BUN)	To assess kidney function and aid in the classification of acute kidney injury *Note: When used for emergency or critical care, results are time-sensitive.*	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma

[26] Calculated as total protein minus albumin. As such, many proteins are included in this calculation and include the sum of alpha 1, alpha 2, beta and gamma globulin (primarily immunoglobulin) fraction.

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Clinical chemistry *continued*	Total calcium	To identify hypercalcaemia or hypocalcaemia, to assess calcium metabolism, to monitor total calcium levels in patients with underlying disease such as certain kinds of cancer (e.g. multiple myeloma, breast cancer and lung cancer), kidney disease, parathyroid disorder or malabsorption.	Semi-automated or automated chemistry analyser	Serum Plasma
	Ionized calcium	To measure free (ionized) calcium in situations in which there are changes in the concentrations of certain proteins (such as albumin) and/or changes in physiological status such as acid–base disorders. To diagnose and monitor hypercalcemia or hypocalcaemia.	Blood gas analysers, including portable analysers for emergency and critical care	Arterial whole blood Venous whole blood Capillary whole blood
	Creatinine	To assess kidney function through estimated glomerular filtration rate (eGFR), urine albumin : creatinine ratio (ACR) and urine protein : creatinine ratio. *Note: When used for emergency or critical care, results are time-sensitive.*	Electrochemical or optical methods on semi-automated or automated chemistry analysers	Serum Urine
	Electrolytes (sodium, potassium, chloride and bicarbonate[27])	To monitor fluid, electrolyte and acid–base balance *Note: When used for emergency or critical care, results are time-sensitive.*	Electrochemical or optical methods on semi-automated or automated chemistry analyser	Serum Plasma

[27] Bicarbonate is sometimes measured as total carbon dioxide.

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Clinical chemistry *continued*	Magnesium	To detect hypomagnesaemia in patients with underlying conditions (i.e. malabsorption, malnutrition), to detect hypermagnesaemia, to aid in the monitoring of kidney function	Semi-automated chemistry analyser	Serum Plasma
	Phosphate	To monitor phosphorus levels in diseases of the kidney, parathyroid, vitamin D metabolism and in tumour lysis syndrome	Optical methods on semi-automated or automated chemistry analysers	Serum Plasma
	Blood pH and gases	To assess lung function, metabolic or kidney disorders and monitor oxygen therapy (includes blood pH, partial pressure of O_2 and carbon dioxide, electrolytes and calculated anion gap) *Note: When used for emergency or critical care, results are time-sensitive.*	Blood gas analysers, including portable analysers for emergency and critical care	Arterial whole blood
	C-reactive protein (CRP)	• To detect inflammation as an indicator of various conditions • To monitor response to treatment *Note: When used for emergency or critical care, results are time-sensitive.*	RDT Latex agglutination assay Immunoassay	Venous whole blood Serum Plasma
	Whole blood lactate	To assess metabolic acidosis, diabetic ketoacidosis, sepsis and dehydration *Note: When used for emergency or critical care, results are time-sensitive.*	Chemistry analyser, blood gas analyser and handheld analyser	Venous whole blood

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Clinical chemistry *continued*	Glucose	To diagnose hypoglycaemia *Note: When used for emergency or critical care, results are time-sensitive.*	Optical and electrochemical methods on semi-automated or automated chemistry analysers and handheld analyzers.	Serum Plasma Capillary whole blood
	Glucose-6-phosphate dehydrogenase (G6PD)	• To screen newborns for G6PD deficiency • To determine G6PD activity (normal, intermediate, deficient) for a decision to administer oxidant drugs, e.g. 8-aminoquinoline drugs for radical cure of *P. vivax* malaria [28]	Semi-quantitative fluorescent spot test	Venous whole blood
	Lipase or amylase	To assess acute pancreatitis and other pancreatic disorders *Note: Lipase result is time-sensitive for emergency and critical care.*	Optical methods, automated chemistry analyser if available	Serum Plasma Peritoneal fluid (amylase)
	Lipid profile	To assess risk of cardiovascular disease (CVD) by measuring cholesterol, triglycerides, high-density lipoprotein (HDL) and low-density lipoprotein (LDL) [29]	Optical methods, automated chemistry analyser if available	Plasma Serum
	Procalcitonin	To guide antibiotic therapy or its discontinuation in sepsis and lower respiratory tract infections (For use only in tertiary care facilities and above)	RDT	Serum Plasma
			Point-of-care immunoassay	Venous whole blood Capillary whole blood Plasma
			Immunoassay	Serum Plasma

[28] Guidelines for the treatment of malaria, 3rd edition: http://apps.who.int/iris/10665/162441.

[29] While low-density lipoprotein (LDL) can be measured, it is routinely calculated.

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Clinical chemistry *continued*	Troponin T/I	To diagnose myocardial infarction *Note: When used for emergency or critical care, results are time-sensitive.*	Immunoassay (handheld or large automated instrument)	Venous whole blood Serum Plasma
	Uric acid	• To aid in the diagnosis and to monitor treatment of gout • To aid in the diagnosis of tumour lysis syndrome associated with acute kidney injury by renal urate deposition during chemotherapy administration	Optical methods, automated chemistry analyser if available	Serum Plasma
	Urine chemistry	To detect and quantify substances in urine associated with metabolic disorders, renal dysfunction or UTIs *Note: When used for emergency or critical care, results are time-sensitive.*	Automated chemical analyser	Urine

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Haematology	Blood cross-matching	To determine blood compatibility for blood transfusions *Note: When used for emergency or critical care, results are time-sensitive.*	Slide and/or tube agglutination tests	Capillary whole blood Venous whole blood
	Complete blood count (CBC), automated	To evaluate overall health and to detect a wide range of disorders, including anaemia, infections, leukaemias, and red blood cell (RBC), white blood cell (WBC) and platelet abnormalities, and primary immune disorders To diagnose and monitor chemotherapy-associated myelotoxicity *Note: When used for emergency or critical care, results are time-sensitive.*	Automated haematology analyser, total and differential counts of WBCs, RBCs, platelets, Hb and haematocrit (Hct)	Capillary whole blood Venous whole blood
	Basic cerebrospinal fluid (CSF) Profile (CSF leukocyte count, CSF differential leukocyte count and CSF protein and glucose)	CSF leukocyte count: • To aid in the diagnosis of bacterial, mycobacterial, fungal and viral meningitis [30]	Haemocytometer / automated haematology analysers with body fluid mode	Cerebrospinal fluid
		CSF leukocyte differential count: • To aid in the diagnosis of bacterial, mycobacterial, fungal and viral meningitis [30]	Wright–Giemsa-stained smears /automated haematology analysers with body fluid mode	
		CSF protein and glucose: • To aid in the diagnosis of bacterial, mycobacterial, fungal and viral meningitis [30]	Automated/semi-automated chemistry analyser	

[30] Definitive diagnosis requires microbiological confirmation – Gram staining, culture, antigen testing, nucleic acid testing.

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Haematology *continued*	D-Dimer	• To diagnose disseminated intravascular coagulation • To aid in the diagnosis of deep vein thrombosis, pulmonary embolism	Immunoassay	Citrate plasma
	Direct antiglobulin test (DAT), also known as direct Coombs test	• To aid in the diagnosis of the cause of immune haemolytic anaemias • To investigate a blood transfusion reaction • To diagnose haemolytic disease of the newborn (HDNB)	Haemagglutination	Venous whole blood
	Fibrinogen	To diagnose disseminated intravascular coagulation	Hand-held or automated coagulation analyser (fibrinogen activity) Immunoassay (fibrinogen antigen)	Citrate plasma
	Haematocrit (Hct)	To diagnose and monitor anaemia *Note: When used for emergency or critical care, results are time-sensitive.*	Micro-haematocrit method (if automated haematology analyser not available)	Capillary whole blood Venous whole blood
			Haematology analyser (preferred)	

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Haematology *continued*	Haemoglobin (Hb)	• To diagnose and monitor anaemia and polycythaemia • To monitor the safety of certain drugs (e.g. zidovudine for HIV infection) • To screen potential blood donors • Clinical marker for certain severe infections (e.g. malaria, viral haemorrhagic fevers) • To aid in the diagnosis of intravascular haemolysis, renal conditions, rhabdomyolysis (myoglobinuria) *Note: When used for emergency or critical care, results are time-sensitive.*	Optical methods, haemoglobinometer, if automated haematology analyser not available Haematology analyser (preferred)	Capillary whole blood Venous whole blood
	Indirect antiglobulin test (IAT), also known as indirect Coombs test or red blood cell antibody screen	• To screen for antibodies to red blood cells before a blood transfusion or during pregnancy • To aid in the diagnosis of haemolytic anaemia and blood transfusion reactions	Haemagglutination	Serum
	Iron studies: Iron Ferritin Transferrin Calculated total iron-binding capacity (TIBC) or transferrin saturation	To diagnose iron deficiency and overload	Optical methods (iron and TIBC) Immunoassay (ferritin and transferrin)	Serum Plasma

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Haematology *continued*	Partial thromboplastin time (PTT), also known as activated partial thromboplastin time (APTT)	• To diagnose bleeding or thrombotic disorders • To monitor anticoagulant therapy	Handheld or automated coagulation analyser	Citrate plasma
	Peripheral blood film examination	To detect red blood cell, white blood cell and platelet abnormalities, malignancies and parasites, and for white blood cell differential count	Microscopic examination of Romanowsky-stained blood films	Capillary whole blood Venous whole blood
	Platelet count	• To diagnose thrombocytopenia or thrombocytosis • Marker to manage severe infections associated with bleeding or sepsis (e.g. viral haemorrhagic fever, meningococcaemia) and certain haematological disorders *Note: When used for emergency or critical care, results are time-sensitive.*	Haemocytometer, if automated haematology analyser is not available Haematology analyser (preferred)	Capillary whole blood Venous whole blood
	Prothrombin time and international normalized ratio (PT/INR)	• To detect or diagnose bleeding or thrombotic disorders (PT) • To monitor performance of anticoagulant medications (INR) *Note: When used for emergency or critical care, results are time-sensitive.*	Hand-held or automated coagulation analyser	Citrate plasma

II.a General IVDs for use in clinical laboratories *continued*

Discipline	Diagnostic test	Test purpose	Assay format	Specimen type
Haematology *continued*	White blood cell count, total	To aid in the diagnosis of infections and leukaemias	Haemocytometer, if automated haematology analyser not available	Capillary whole blood Venous whole blood
			Haematology analyser (preferred)	
	Sickle cell testing	To aid in the diagnosis of sickle cell anaemia, sickle cell trait and other sickling disorders	Sodium metabisulfite slide test	Venous whole blood
			Haemoglobin solubility	
		To diagnose sickle cell anaemia, sickle cell trait and other sickling disorders	Haemoglobin electrophoresis	Venous whole blood
Clinical pathology	Urine microscopy	To aid in the diagnosis of kidney and urologic diseases by detecting the presence of cells (white blood cells, red blood cells, epithelial cells), casts and crystals in urine sediment To detect the presence of pathogens	Microscopic examination May require staining procedures for microbial pathogens (e.g. Gram stain, modified Ziehl–Neelsen stain)	Urine
	Body fluid microscopy	To aid in the diagnosis of inflammatory, infectious and neoplastic diseases involving body fluids (e.g. pleural, peritoneal, synovial, pericardial) by detecting the presence or absence of cells (white blood cells, red blood cells, mesothelial cells) along with cell count and differential count	Microscopic examination May require staining procedures for microbial pathogens and cytological examination for neoplastic cells	Body fluids
Pregnancy testing	Human chorionic gonadotrophin (hCG)	To detect and/or confirm pregnancy	Optical method	Serum
			Immunoassay	

II.b Disease-specific IVDs for use in clinical laboratories

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Aspergillosis	*Aspergillus* IgG antibody	To aid in the diagnosis of chronic pulmonary aspergillosis	RDT Immunoassay	Serum Plasma	N/A	
	Aspergillus antigen test	To aid in the diagnosis of invasive aspergillosis in immunocompromised patients	RDT	Bronchoalveolar lavage (BAL)	N/A	
			Immunoassay	Serum bronchoalveolar lavage (BAL)		

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer	Alpha-fetoprotein (AFP)	• To screen for hepatocellular carcinoma (HCC) in high-risk individuals with liver cirrhosis or with a family history, in conjunction with ultrasound • For staging and monitoring of germ cell tumours • To aid in the diagnosis and staging of hepatoblastoma	Immunoassay	Serum Plasma	N/A	Guidelines for the care and treatment of persons diagnosed with chronic hepatitis C virus infection (2018) https://apps.who.int/iris/handle/10665/273174 Guidelines for the prevention, care and treatment of persons with chronic hepatitis B infection https://apps.who.int/iris/handle/10665/154590 WHO classification of tumours of the urinary system and male genital organs. WHO classification of tumours, 4th edition, volume 8. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Urinary-System-And-Male-Genital-Organs-2016

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*						WHO classification of tumours of female reproductive organs. WHO classification of tumours, 4th edition, volume 6. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-Female-Reproductive-Organs-2014
	Basic panel for immunohisto-chemical (IHC) testing for diagnosis of lymphoma	To aid in the diagnosis, sub-classification, prognosis and treatment of lymphoma (including HIV-associated conditions)	IHC testing (ki-67, CD45, BCL6, IRF4/MUM1, MYC, CD20, CD5, CD10, BCL2, CD23, CD79a, cyclinD1, CD3, CD15, CD30, TdT, CD138/syndecan-1, kappa and lambda chains, PAX5)	Formalin-fixed paraffin-embedded tissue (FFPE)[31]	N/A	WHO classification of tumours of haematopoietic and lymphoid tissues. WHO classification of tumours, revised 4th edition, volume 2. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-Haematopoietic-And-Lymphoid-Tissues-2017 WHO Guide for establishing a pathology laboratory in the context of cancer control. https://www.who.int/publications/i/item/guide-for-establishing-a-pathology-laboratory-in-the-context-of-cancer-control

[31] Only for use in specialized anatomical pathology laboratories – see Anatomical Pathology under II.a General IVDs for use in clinical laboratories.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*	Basic panel of immunohisto-chemical (IHC) markers for diagnosis of solid tumours	To aid in diagnosis, prognosis and treatment of solid tumours, especially childhood cancer	IHC testing markers include desmin, cytokeratin, AE1/AE3, S100, synaptophysin, myogenin, hCG, PLAP, Oct3/4, NANOG, CD30, CD117/c-kit, WT1, SALL4	Formalin-fixed paraffin-embedded tissue (FFPE)[32]	N/A	WHO classification of tumours, 4th edition https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours WHO list of priority medical devices for cancer management https://apps.who.int/iris/bitstream/handle/10665/255262/9789241565462-eng.pdf WHO Guide for establishing a pathology laboratory in the context of cancer control. https://www.who.int/publications/i/item/guide-for-establishing-a-pathology-laboratory-in-the-context-of-cancer-control

[32] Only for use in specialized anatomical pathology laboratories – see Anatomical Pathology section under II.a General IVDs for use in clinical laboratories.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*	BCR-ABL1 and ABL1 transcripts	To diagnose and monitor therapy of chronic myelocytic leukaemia (CML) and CML variants (e.g., neutrophilic CML) and prognosis of acute lymphoblastic leukaemia (ALL)	NAT	Venous whole blood Bone Marrow	N/A	WHO classification of tumours of haematopoietic and lymphoid tissues. WHO classification of tumours, revised 4th edition, volume 2. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-Haematopoietic-And-Lymphoid-Tissues-2017 20th essential medicines list (2017) https://apps.who.int/iris/handle/10665/273826
	Epidermal growth factor receptor (EGFR) gene mutation	To aid in the diagnosis and treatment of non-squamous non-small cell lung carcinoma	NAT	Formalin-fixed paraffin-embedded tissue (FFPE) and buffered lung tumour specimen [33]	N/A	

[33] Only for use in specialized anatomical pathology laboratories – see Anatomical Pathology under II.a General IVDs for use in clinical laboratories.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*	Basic flow cytometry panel of antibodies for leukaemia[34]	To aid in the diagnosis of acute leukaemias	Flow cytometry (CD10, CD19, CD45, CD34, CD7, CD33, CD117, myeloperoxidase, CD79a, cytoplasmic CD3, HLA-DR, CD5, CD23, CD43)	Bone marrow Peripheral blood Body fluids Tissue Lymph node aspirate	N/A	WHO classification of tumours of haematopoietic and lymphoid tissues. WHO classification of tumours, revised 4th edition, volume 2. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-Haematopoietic-And-Lymphoid-Tissues-2017 WHO list of priority medical devices for cancer management https://apps.who.int/iris/handle/10665/255262 WHO Guide for establishing a pathology laboratory in the context of cancer control https://www.who.int/publications/i/item/guide-for-establishing-a-pathology-laboratory-in-the-context-of-cancer-control

[34] Detection of 10 unique items among 12 components for four-colour fluorochrome cytometry.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*	Faecal immuno-chemical test (FIT)	To screen for colorectal cancer	Latex agglutination immuno-turbidimetry	Stool	N/A	WHO priority medical devices for cancer management https://apps.who.int/iris/handle/10665/255262 Colorectal cancer screening. IARC handbooks of cancer prevention, volume 17 https://publications.iarc.fr/Book-And-Report-Series/Iarc-Handbooks-Of-Cancer-Prevention/Colorectal-Cancer-Screening-2019

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*	Human chorionic gonadotrophin (hCG) (measured as total beta-hCG[35])	To aid in the diagnosis of and monitoring of germ cell tumours and gestational trophoblastic disease	Immunoassay	Plasma	N/A	WHO classification of tumours of the urinary system and male genital organs. WHO classification of tumours, 4th edition, volume 8. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Urinary-System-And-Male-Genital-Organs-2016 WHO classification of tumours of female reproductive organs. WHO classification of tumours, 4th edition, volume 6. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-Female-Reproductive-Organs-2014

[35] To include both free and intact beta-hCG.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*	Lactate dehydrogenase (LDH)	To aid in the prognosis and monitoring of haematological malignancies (e.g. lymphoma) and germ cell tumours NOTE: Also used as a marker for heart disease, liver disease, *Pneumocystis* infection, thrombotic thrombocytopenic purpura (TTP), and other malignancies	Optical methods, automated chemistry analyser if available	Serum Plasma	N/A	WHO classification of tumours of haematopoietic and lymphoid tissues. WHO classification of tumours, revised 4th edition, volume 2. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-Haematopoietic-And-Lymphoid-Tissues-2017 WHO classification of tumours of the urinary system and male genital organs. WHO classification of tumours, 4th edition, volume 8. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Urinary-System-And-Male-Genital-Organs-2016

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*						WHO classification of tumours of female reproductive organs. WHO classification of tumours, 4th edition, volume 6. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-Female-Reproductive-Organs-2014
	Estrogen (ER) and progesterone (PgR) receptors	To aid in diagnosis, prognosis and treatment of breast cancer	IHC testing	Formalin-fixed paraffin embedded tissue (FFPE)[36]	N/A	WHO classification of tumours of the breast. WHO classification of tumours, 4th edition, volume 4. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Breast-2012 WHO list of priority medical devices for cancer management https://apps.who.int/iris/handle/10665/255262/

[36] Only for use in specialized anatomical pathology laboratories – see Anatomical Pathology section under II.a General IVDs for use in clinical laboratories.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*						WHO model list of essential medicines – 21st list, 2019 https://www.who.int/groups/expert-committee-on-selection-and-use-of-essential-medicines/essential-medicines-lists Guidelines for management of breast cancer. WHO Regional Office for the Eastern Mediterranean (2006) http://applications.emro.who.int/dsaf/dsa697.pdf WHO guide for establishing a pathology laboratory in the context of cancer control https://www.who.int/publications/i/item/guide-for-establishing-a-pathology-laboratory-in-the-context-of-cancer-control

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*	Papanicolaou (Pap) smear test	To screen for and to aid in early diagnosis of cervical cancer	Microscopic examination of cervical cells on slides	Cervical smear from liquid cytology specimen	N/A	Guidelines for screening and treatment of precancerous lesion for cervical cancer prevention. WHO guidelines (2013). https://apps.who.int/iris/handle/10665/94830 WHO guide for establishing a pathology laboratory in the context of cancer control. https://www.who.int/publications/i/item/guide-for-establishing-a-pathology-laboratory-in-the-context-of-cancer-control Comprehensive cervical cancer control. https://www.who.int/publications/i/item/978924154895 Global strategy to accelerate the elimination of cervical cancer as a public health problem. https://www.who.int/publications/i/item/9789240014107 Guide to cancer early diagnosis https://www.who.int/publications/i/item/guide-to-cancer-early-diagnosis

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*						WHO technical guidance and specifications of medical devices for screening and treatment of precancerous lesions in the prevention of cervical cancer https://apps.who.int/iris/handle/10665/331698
	Prostate-specific antigen (PSA)	To aid in diagnosis, prognosis and monitoring of prostate cancer	Immunoassay	Peripheral blood	N/A	WHO classification of tumours of the urinary system and male genital organs. WHO classification of tumours, 4th edition, volume 8. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Urinary-System-And-Male-Genital-Organs-2016

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Cancer *continued*	Tyrosine-protein kinase receptor (erbB-2) or human epidermal growth factor receptor 2 (HER-2) overexpression	To aid in treatment and prognosis of breast cancer	IHC testing	Formalin-fixed paraffin-embedded tissue (FFPE)[37] (Referred specimens must be fixed correctly before transport)	N/A	WHO classification of tumours of the breast. WHO classification of tumours, 4th edition, volume 4. https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-The-Breast-2012 WHO list of priority medical devices for cancer management https://apps.who.int/iris/handle/10665/255262 WHO 20th essential medicines list (2017) https://apps.who.int/iris/handle/10665/273826 WHO Guide for establishing a pathology laboratory in the context of cancer control. https://www.who.int/publications/i/item/guide-for-establishing-a-pathology-laboratory-in-the-context-of-cancer-control

[37] Only for use in specialized anatomical pathology laboratories – see Anatomical Pathology section under II.a General IVDs for use in clinical laboratories.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Coronavirus disease (COVID-19)	SARS-CoV-2 nucleic acid test	To diagnose infection by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in symptomatic and asymptomatic individuals suspected of having been exposed For surveillance and confirmation of outbreaks	NAT[38]	Upper respiratory specimens (e.g. nasopharyngeal and oropharyngeal) and lower respiratory specimens (e.g. BAL)	Emergency Use Listing (EUL) https://www.who.int/teams/regulation-prequalification/eul	Diagnostic testing for SARS-CoV-2. Interim guidance (11 September 2020). https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2 Guidance on SARS CoV-2 testing is reviewed regularly based on available evidence. For up to date guidance see: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance-publications?publicationtypes=f85a3610-b102-4287-a6df-f3bc0b2e9f7c

[38] Listing was based on evidence for RT-PCR tests. Other types of nucleic acid amplification require more evidence and will be subject to further review.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Diabetes mellitus	Glucose	• To diagnose and monitor[39] type 1 and type 2 diabetes mellitus • To diagnose impaired fasting glucose/impaired glucose tolerance • To screen for type 2 diabetes mellitus and impaired fasting glucose/ impaired glucose tolerance *Note: When used for emergency or critical care, results are time-sensitive.*	Optical methods, automated chemistry analyser if available	Serum Plasma	N/A	HEARTS-D: diagnosis and management of type 2 diabetes (2020) https://www.who.int/publications/i/item/who-ucn-ncd-20.1
	Haemoglobin A1c (HbA1c)	To diagnose and monitor diabetes mellitus	Immunoassay	Venous whole blood	N/A	HEART-D: diagnosis and management of type 2 diabetes (2020) https://www.who.int/publications/i/item/who-ucn-ncd-20.1

[39] If HbA1c is not available.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Endocrine disorders	Cortisol (total)[40]	• To diagnose central (pituitary) or primary (adrenal, Addison's disease) cortisol deficiency[41] • To diagnose central (pituitary) or primary (adrenal) hypercortisolism (Cushing's syndrome)[42]	Immunoassay	Serum Plasma	N/A	N/A
	Estradiol[40]	• To aid in the diagnosis of anovulation, gonadal dysfunction, precocious puberty, and primary and secondary amenorrhoea • To aid in the evaluation and management of infertility	Immunoassay	Serum Plasma	N/A	N/A

[40] For use in specialized health care settings.

[41] Often used with timed collection and stimulation with cosyntropin.

[42] Often used with timed collection and suppression with dexamethasone.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Endocrine disorders *continued*	Follicle-stimulating hormone (FSH)	• To aid in the diagnosis of anovulation, gonadal dysfunction, precocious puberty, and primary and secondary amenorrhea • To aid in the evaluation and management of infertility[43]	Immunoassay	Serum Plasma	N/A	N/A
	Luteinizing hormone (LH)	• To aid in the diagnosis of anovulation, gonadal dysfunction, precocious puberty, and primary and secondary amenorrhoea • To aid in the evaluation and management of infertility[43]	Immunoassay	Serum Plasma	N/A	N/A

[43] For use in specialized health care settings.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Endocrine disorders *continued*	Progesterone[44]	To confirm ovulation during infertility evaluation and treatment	Immunoassay	Serum Plasma	N/A	N/A
	Prolactin[44]	To diagnose and monitor hyperprolactinaemia (including prolactinoma)	Immunoassay	Serum Plasma	N/A	N/A
	Thyroid-stimulating hormone (TSH)	To screen[45] for and to diagnose hypothyroidism and hyperthyroidism	Immunoassay	Serum Plasma Capillary whole blood (newborns)	N/A	

[44] For use in specialized health care settings.

[45] Only in the context of neonatal screening for congenital hypothyroidism and of screening of patients with medical conditions such as type 1 diabetes mellitus where the incidence of hypothyroidism is higher than in the general population.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Hepatitis B	Hepatitis B virus (HBV) surface antigen (HBsAg)	To screen for HBV and to aid in the diagnosis of chronic hepatitis B virus (HBV) infection: infants > 12 months of age, children, adolescents and adults	RDT	Venous whole blood Capillary whole blood Plasma Serum	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=63	Guidelines on hepatitis B and C testing (February 2017) http://apps.who.int/iris/handle/10665/254621 https://www.who.int/news-room/fact-sheets/detail/hepatitis-b
			Immunoassay	Plasma Serum		
	Quantitative HBV nucleic acid test	To stage chronic HBV infection, to determine the need for treatment (including use of antivirals in mother to prevent mother-to-child transmission) and to monitor response to treatment.	NAT	Serum Plasma	N/A	
	Hepatitis B e antigen (HBeAg)	To stage chronic HBV infection and to determine the need for treatment	Immunoassay	Serum Plasma	N/A	

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Hepatitis B *continued*	IgM-specific antibodies to hepatitis B core antigen (IgM anti-HBc)	To aid in the diagnosis of acute HBV infection in the context of outbreak investigation	Immunoassay	Serum Plasma	N/A	
	Antibodies to hepatitis B surface antigen (anti-HBs)	To determine immune status due to HBV immunization.	Immunoassay	Serum Plasma	N/A	

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Hepatitis C	Antibodies to hepatitis C virus (HCV) (anti-HCV)	To screen for HCV infection, and to aid in the diagnosis of viraemic HCV infection: infants > 18 months of age, children, adolescents and adults	RDT	Capillary whole blood Venous whole blood Plasma Serum	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=59	Guidelines on hepatitis B and C testing (February 2017) http://apps.who.int/iris/handle/10665/254621 https://www.who.int/news-room/fact-sheets/detail/hepatitis-c
			Immunoassay	Serum Plasma		
	Combined antibodies to HCV (anti-HCV) and HCV core antigen (HCVcAg)	To screen for HCV infection, and to aid in the diagnosis of viraemic HCV infection: infants > 18 months of age, children, adolescents and adults	Immunoassay	Serum Plasma		
	HCV core antigen (HCVcAg)	To aid in the diagnosis of viraemic HCV infection	Immunoassay	Serum Plasma		
	Qualitative or quantitative HCV nucleic acid test	To diagnose viraemic HCV and to monitor response to treatment, and as a test of cure	NAT	Capillary whole blood Venous whole blood Serum Plasma Dried blood spots		

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
HIV infection	Antibodies to HIV-1/2 (anti-HIV Ab)	To screen for or to aid in the diagnosis of HIV infection: adults, adolescents, children and infants > 18 months of age	RDT	Venous whole blood Plasma Serum	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=58	Guidelines on HIV self-testing and partner notification (2016) http://apps.who.int/iris/handle/10665/251655
			Immunoassay	Serum Plasma		Consolidated guidelines on HIV testing services (July 2015) https://apps.who.int/iris/handle/10665/179870 WHO implementation tool for pre-exposure prophylaxis (PrEP) of HIV infection, module 10 for testing providers (2017) https://apps.who.int/iris/handle/10665/258516
	Combined HIV antibody/p24 antigen (anti-HIV/p24 Ag)	To screen for or to aid in the diagnosis of HIV infection: adults, adolescents, children and infants > 18 months of age	RDT	Venous whole blood Plasma Serum		Consolidated guidelines on HIV testing services (2015) https://apps.who.int/iris/handle/10665/179870
			Immunoassay	Serum Plasma		

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
HIV infection *continued*	Qualitative HIV nucleic acid test (NAT)	To diagnose HIV infection in infants < 18 months of age	NAT	Capillary whole blood Venous whole blood Dried blood spots Plasma	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=61	Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection (2016) https://apps.who.int/iris/handle/10665/208825 WHO HIV molecular diagnostics toolkit to improve access to viral load testing and infant diagnosis: Toolkit (July 2019) https://apps.who.int/iris/handle/10665/325961
	Quantitative HIV nucleic acid test (NAT)	• To monitor response to antiretroviral treatment • To diagnose HIV infection in infants < 18 months of age (only if validated by the manufacturer)	NAT	Dried blood spots (whole blood or plasma) Serum Plasma		
	CD4 cell enumeration	• To stage advanced HIV disease • To monitor response to antiretroviral therapy (in settings where quantifying viral load is not available)	Flow cytometry	Capillary whole blood Venous whole blood	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=66	Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection (2016) https://apps.who.int/iris/handle/10665/208825 Guidelines for managing advanced HIV disease and rapid initiation of antiretroviral therapy https://apps.who.int/iris/handle/10665/255884

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
HIV infection *continued*	Cryptococcal antigen	To screen for and diagnose cryptococcal meningitis in people with advanced HIV disease	RDT	Cerebrospinal fluid Capillary whole blood Venous whole blood Serum Plasma	N/A	Guidelines for the diagnosis, prevention, and management of cryptococcal disease in HIV-infected adults, adolescents and children (2018) http://apps.who.int/iris/handle/10665/260399 Guidelines for managing advanced HIV disease and rapid initiation of antiretroviral therapy (2017) https://apps.who.int/iris/handle/10665/255884
			Immunoassay	Cerebrospinal fluid Serum Plasma		
	Histoplasma capsulatum antigen	To aid in the diagnosis of disseminated histoplasmosis	Immunoassay	Urine	N/A	Guidelines for managing advanced HIV disease and rapid initiation of antiretroviral therapy (2017) https://apps.who.int/iris/handle/10665/255884

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Human papilloma-virus (HPV) Infection	HPV nucleic acid test (NAT)	For cervical cancer screening	NAT	Cervical cells collected in test-specific transport fluid/ vessel	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=65	Introducing and scaling up testing for human papillomavirus as part of a comprehensive programme for prevention and control of cervical cancer. A step-by-step guide. https://www.who.int/publications/i/item/9789240015166 WHO human papillomavirus laboratory manual, first edition (2009) http://apps.who.int/iris/handle/10665/70505 Comprehensive cervical cancer control. https://apps.who.int/iris/handle/10665/144785 Global strategy to accelerate the elimination of cervical cancer as a public health problem. https://www.who.int/publications/i/item/9789240014107 WHO technical guidance and specifications of medical devices for screening and treatment of precancerous lesions in the prevention of cervical cancer https://apps.who.int/iris/handle/10665/331698

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Influenza	Influenza A and B nucleic acid test (NAT)	To diagnose seasonal influenza	NAT	Nasal swab Nasopharyngeal swab Nasopharyngeal aspirate or wash	N/A	Manual for the laboratory diagnosis and virological surveillance of influenza (2011) https://apps.who.int/iris/handle/10665/44518 Global epidemiological surveillance standards for influenza https://www.who.int/influenza/resources/documents/WHO_Epidemiological_Influenza_Surveillance_Standards_2014.pdf

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Malaria	*Plasmodium* spp. antigens; species-specific (e.g. HRP2) and/or pan-species-specific (e.g. pan-pLDH)	To diagnose infection by one or more human malaria parasite species (*P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*)	RDT	Capillary whole blood Venous whole blood	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=64	WHO guidelines for the treatment of malaria, third edition (2015) http://apps.who.int/iris/10665/162441 Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 8 (2016–2018) https://apps.who.int/iris/handle/10665/276190/ Information note on recommended selection criteria for procurement of malaria rapid diagnostic tests https://apps.who.int/iris/handle/10665/259870 WHO good practices for selecting and procuring rapid diagnostic tests for malaria (2011) http://apps.who.int/iris/handle/10665/44530

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Malaria *continued*	*Plasmodium* spp.	To diagnose infection by one or more human malaria parasite species (*P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*) and monitoring response to treatment	Light microscopy	Capillary whole blood Venous whole blood	N/A	WHO guidelines for the treatment of malaria, third edition (2015) http://apps.who.int/iris/10665/162441 Basic malaria microscopy. Part I: learner's guide (2010) http://apps.who.int/iris/handle/10665/44208 Malaria microscopy standard operating procedures (2015) https://apps.who.int/iris/handle/10665/274382

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Neglected tropical diseases	Qualitative dengue virus nucleic acid test	For surveillance (serotype differentiation) and confirmation of outbreaks	NAT	Serum Plasma Dried blood spots (DBS)	N/A	*Dengue: guidelines for diagnosis, treatment, prevention and control (2009)* https://apps.who.int/iris/handle/10665/44188
	Dengue virus IgM antibody	To aid in the diagnosis of dengue fever (always in combination with NS1) and for population surveys	RDT	Serum Venous whole blood	N/A	
			Immunoassay	Venous whole blood Dried blood spots (DBS) Saliva		
	Dengue virus antigen (NS1)	To aid in the diagnosis of dengue fever (always in combination with IgM) and for population surveys	RDT	Serum Venous whole blood	N/A	
			Immunoassay	Serum Plasma		

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Neglected tropical diseases *continued*	Kato-Katz faecal smear	For surveillance and diagnosis of soil-transmitted helminthiases and schistosomiasis caused by *Schistosoma mansoni*, *S. intercalatum*, *S. japonicum*, *S. mekongi*	Microscopic slide examination	Fresh stool	N/A	
	Trypanosoma cruzi IgG antibody	• For surveillance of *T. cruzi* infection • To screen girls, women of childbearing age and pregnant women without previous treatment for *T. cruzi* infection • To screen children and other at-risk populations • To diagnose chronic *T. cruzi* infection (Chagas disease) • To monitor treatment of *T. cruzi* infection	Immunoassay	Serum Plasma	N/A	Guidelines for the diagnosis and treatment of Chagas disease (2019) https://iris.paho.org/handle/10665.2/49653

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Neglected tropical diseases *continued*	Visceral leishmaniasis direct agglutination test	To aid in the diagnosis of clinically suspected visceral leishmaniasis (kala-azar)	Agglutination assay	Serum Dried blood spots (check validation)	N/A	Control of the leishmaniases: report of a meeting of the WHO Expert Committee on the Control of Leishmaniases. Geneva: World Health Organization; 2010 (WHO Technical Report Series, No 949) https://apps.who.int/iris/handle/10665/44412
Pneumocystis pneumonia	*Pneumocystis jirovecii* nucleic acid test	To aid in the diagnosis of *Pneumocystis* pneumonia[46]	NAT	Respiratory specimens (sputum, bronchoalveolar lavage fluid)	N/A	

[46] Particularly relevant in immunocompromised patients.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Primary Immuno-deficiencies	HIV 1/2 antibody (anti-HIV Ab)	For differential diagnosis of primary immunodeficiencies	RDT	Oral fluid Capillary whole blood Venous whole blood	N/A	N/A
	Plasma immunoglo-bulin levels (IgG, IgA, IgM)	To identify patients with low plasma immunoglobulin levels and to monitor replacement	Radial immuno-diffusion (RID)	Serum	N/A	N/A
			Immunoassay	Serum Plasma		
	Lymphocyte subtype enumeration: CD3, CD4, CD8, B cells CD19 and/or CD20, CD16/56 T cells and NK cells (Refer to HIV infection for enumeration of CD4 cells only)	To aid in the diagnosis of primary and secondary immunodeficiencies	Flow cytometry	Venous whole blood	N/A	

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Sexually transmitted infections	Qualitative nucleic acid test (NAT) for *Chlamydia trachomatis* (CT) and *Neisseria gonorrhoeae* (NG) infections	To diagnose chlamydial and/ or gonorrhoeal urogenital disease and extragenital infection	NAT	Urine, urethral swabs endocervical swabs, vaginal swabs, rectal swabs, oropharyngeal swabs, Liquid cytology	N/A	WHO sexually transmitted infection laboratory manual https://apps.who.int/iris/handle/10665/85343 Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations https://apps.who.int/iris/handle/10665/246200
	Antibodies to *Treponema pallidum*	To diagnose or to aid in the diagnosis of syphilis	RDT	Venous whole blood Plasma Serum	N/A	WHO laboratory diagnosis of sexually transmitted infections, including human immunodeficiency virus (2013) http://apps.who.int/iris/handle/10665/85343
			Immunoassay	Serum Plasma		
	Antibodies to *T. pallidum* and to HIV-1/2 (anti-HIV Ab)	To diagnose or to aid in the diagnosis of HIV-1/2 infection and/or syphilis	RDT	Venous whole blood Plasma Serum	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=57	WHO Information note on the use of dual HIV/syphilis rapid diagnostic tests (RDT) (2017) http://apps.who.int/iris/handle/10665/252849/ Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations https://apps.who.int/iris/handle/10665/246200

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Sexually transmitted infections *continued*	Non-treponemal rapid plasma reagin (RPR)	To screen for syphilis and monitor treatment effectiveness	Particle/ charcoal agglutination assay	Serum Plasma	N/A	WHO sexually transmitted infection laboratory manual https://apps.who.int/iris/handle/10665/85343
	Non-treponemal venereal disease research laboratory (VDRL) test	To screen for, diagnose and confirm neurosyphilis	Flocculation test	Serum Plasma Cerebrospinal fluid	N/A	
	T. pallidum haemagglutination (TPHA) test	To confirm syphilis infection and diagnose early and late syphilis infection	Haemagglutination assay	Serum (preferred) Plasma	N/A	Laboratory diagnosis of sexually transmitted infections, including human immunodeficiency virus https://apps.who.int/iris/handle/10665/85343
	T. pallidum article agglutination (TPPA) test		Particle agglutination assay		N/A	
Streptococcal pharyngitis	Group A *Streptococcus* antigen	To aid in the diagnosis of Group A streptococcal pharyngitis	RDT Immunoassay	Throat swab	N/A	N/A

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products[47]	WHO supporting documents
Tuberculosis (TB)	*Mycobacterium tuberculosis* bacteria	To diagnose and monitor treatment of active TB	Microscopy	Sputum or other specimen types	Implementing tuberculosis diagnostics: policy framework (2015) https://apps.who.int/iris/handle/10665/162712	Compendium of WHO guidelines and associated standards: ensuring optimum delivery of the cascade of care for patients with tuberculosis, 2nd edition (2018) https://apps.who.int/iris/handle/10665/272644
		To diagnose and monitor treatment of active TB, including drug-resistant TB	Bacterial culture	Sputum or other specimen types		
	M. tuberculosis DNA	To diagnose active TB and simultaneously or sequentially detect rifampicin resistance	NAT	Sputum Broncho-alveolar lavage (BAL) or extra-pulmonary TB specimen types	WHO meeting report of a technical expert consultation: non-inferiority analysis of Xpert MTB/RIF Ultra compared to Xpert MTB/RIF (2017) http://apps.who.int/iris/handle/10665/254792 Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: policy update (2013) https://apps.who.int/iris/handle/10665/112472	Implementing tuberculosis diagnostics: policy framework (2015) https://apps.who.int/iris/handle/10665/162712

[47] All TB tests are evaluated and guidelines developed by the WHO Global TB Programme.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products[48]	WHO supporting documents
Tuberculosis (TB) *continued*	*M. tuberculosis* DNA	To diagnose active TB	Loop-mediated isothermal amplification (LAMP)	Sputum	The use of loop-mediated isothermal amplification (TB-LAMP) for the diagnosis of pulmonary tuberculosis: policy guidance (2016) http://apps.who.int/iris/10665/249154	
	M. tuberculosis DNA mutations associated with resistance	To detect resistance to first-line anti-TB medicines	Molecular line probe assay (LPA)	Sputum	The use of molecular line probe assays for the detection of resistance to isoniazid and rifampicin: policy update (2016) https://apps.who.int/iris/handle/10665/250586	Compendium of WHO guidelines and associated standards: ensuring optimum delivery of the cascade of care for patients with tuberculosis (2017) https://www.who.int/tb/publications/Compendium_WHO_guidelines_TB_2017/en/ Implementing tuberculosis diagnostics: policy framework (2015) https://apps.who.int/iris/handle/10665/162712
	M. tuberculosis DNA mutations associated with resistance	To detect resistance for second-line anti-TB medicines	Molecular line probe assay (LPA)	Sputum	The use of molecular line probe assays for the detection of resistance to second-line anti-tuberculosis drugs: policy update (2016) http://apps.who.int/iris/handle/10665/246131	

[48] All TB tests are evaluated and guidelines developed by the WHO Global TB Programme.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products[49]	WHO supporting documents
Tuberculosis (TB) *continued*	Drug susceptibility testing of *M. tuberculosis* culture	To detect resistance to first-line and/or second-line anti-TB medicines	Drug susceptibility testing	Bacterial culture of *M. tuberculosis*	Technical report on critical concentrations for drug susceptibility testing of medicines used in the treatment of drug-resistant tuberculosis (2018) http://apps.who.int/iris/bitstream/10665/260470/1/WHO-CDS-TB-2018.5-eng.pdf?ua=1	
	Immune response by interferon-gamma release assay (IGRA)	To diagnose latent TB infection	Immunoassay or enzyme-linked immune absorbent spot (ELISpot) assay	Venous whole blood		Latent TB Infection: updated and consolidated guidelines for programmatic management (2018) http://apps.who.int/iris/handle/10665/260233

[49] All TB tests are evaluated and guidelines developed by the WHO Global TB Programme.

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Vaccine-preventable diseases	Measles nucleic acid test	To diagnose clinically suspected measles infection	NAT	Oral fluid Throat swab Nasopharyngeal aspirates or swabs Urine	N/A	Manual for the laboratory diagnosis of measles and rubella virus infection https://www.who.int/immunization/monitoring_surveillance/burden/laboratory/Manual_lab_diagnosis_of_measles_rubella_virus_infection_ENG.pdf?ua=1 Manual for the laboratory-based surveillance of measles, rubella, and congenital rubella syndrome https://www.who.int/immunization/monitoring_surveillance/burden/laboratory/manual/en/ Surveillance standards for vaccine-preventable diseases, 2nd edition https://apps.who.int/iris/handle/10665/275754 The immunological basis for immunization series. Module 11: rubella. Geneva: World Health Organization; 2008. https://apps.who.int/iris/handle/10665/43922
	Measles IgM antibody	To diagnose clinically suspected measles infection	Immunoassay	Serum Plasma Dried Blood Spots Oral fluid	N/A	
	Measles IgG antibody	To aid in the diagnosis of clinically suspected measles infection	Immunoassay	Serum Plasma Dried blood spots Oral fluid	N/A	
	Rubella IgM antibody	To diagnose active rubella infection or recent exposure	Immunoassay	Serum Plasma Dried blood spots Oral fluid	N/A	

II.b Disease-specific IVDs for use in clinical laboratories *continued*

Disease	Diagnostic test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Vaccine-preventable diseases *continued*	Rubella IgG antibody	To screen for prior exposure to rubella infection or vaccination, particularly in pregnant women	Immunoassay	Serum Plasma Dried blood spots Oral fluid	N/A	WHO. Rubella vaccines: WHO position paper. Weekly epidemiol record. 2011;29(86):301–316. https://www.who.int/wer/2011/wer8629.pdf?ua=1
Zika virus infection	IgM antibodies to Zika virus	To aid in the diagnosis of suspected Zika virus infection[50]	Immunoassay	Serum (Not to be used with CSF)	N/A	Laboratory testing for Zika virus infection: interim guidance https://apps.who.int/iris/handle/10665/204671
	Zika virus nucleic acid test (NAT)	To diagnose acute Zika virus infection[51,52]	NAT	Venous whole blood Serum Plasma Urine CSF	WHO Emergency Use Listing (EUL) https://extranet.who.int/pqweb/vitro-diagnostics/zika-virus-disease	

[50] Because of potential cross-reactivity with dengue and other flaviviruses and persistence of Zika IgM antibody that may reflect infection prior to pregnancy, currently available Zika virus IgM test results should not be used alone for clinical decision-making in pregnancy.

[51] Zika virus RNA is typically detectable in serum by NAT assays only within the first week of infection. A negative result does not rule out infection.

[52] To reduce risk of false-positive results in pregnant women, a positive NAT test should be confirmed by re-extraction and repeat NAT testing of the same specimen.

II.c Disease-specific IVDs for blood screening laboratories

Disease	Screening test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Hepatitis B virus (HBV)	Hepatitis B surface antigen (HBsAg)	To screen blood donations for HBV	RDT[53,54]	Capillary whole blood Venous whole blood Plasma Serum	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=63	Screening donated blood for transfusion-transmissible infections: recommendations (2009) http://apps.who.int/iris/handle/10665/44202
			Particle agglutination assay[53,54]	Plasma Serum		
			Immuno-assay[53]	Plasma Serum		

[53] The only assays recommended for blood screening are those that have been validated for this purpose by the manufacturer.

[54] May be performed in laboratories with small throughput, in remote areas or emergency situations.

NOTE: Please refer to the Haematology section for information on General IVDs for blood transfusion.

II.c Disease-specific IVDs for blood screening laboratories *continued*

Disease	Screening test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Hepatitis C virus (HCV)	Antibodies to HCV (anti-HCV)	To screen blood donations for HCV	RDT[55,56]	Capillary whole blood Venous whole blood Plasma Serum	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=59	
			Immuno-assay[55]	Serum Plasma		
	Combined antibodies to HCV (anti-HCV) and HCV core antigen (HCV cAg)	To screen blood donations for HCV	Immuno-assay[55]	Serum Plasma		

[55] The only assays recommended for blood screening are those that have been validated for this purpose by the manufacturer.

[56] May be performed in laboratories with small throughput, in remote areas or emergency situations.

II.c Disease-specific IVDs for blood screening laboratories *continued*

Organism	Screening test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
HIV	Antibodies to HIV-1/2 (anti-HIV Ab)	To screen blood donations for HIV	RDT[57]	Capillary whole blood Venous whole blood Serum Plasma	Public reports of WHO prequalified IVDs https://extranet.who.int/pqweb/vitro-diagnostics/prequalification-reports/whopr?field_whopr_category=58	Screening donated blood for transfusion-transmissible infections: recommendations (2009) http://apps.who.int/iris/handle/10665/44202
			Particle agglutination assay[57]	Serum Plasma		
			Immuno-assay[57,58]	Serum Plasma		
	Combined HIV antibody/p24 antigen (anti-HIV/p24 Ag) test	To screen blood donations for HIV	Immuno-assay[57,58]	Serum Plasma		

[57] The only assays recommended for blood screening are those that have been validated for this purpose by the manufacturer.

[58] May be performed in laboratories with small throughput, in remote areas or emergency situations.

II.c Disease-specific IVDs for blood screening laboratories *continued*

Organism	Screening test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Malaria[59]	Antibodies to *Plasmodium* spp.	To screen blood donations for one or more human malaria species (*P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*) **in non-endemic areas**	Immunoassay	Serum Plasma		
	Plasmodium spp. antigens	To screen blood donations for one or more human malaria species (*P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*) **in endemic areas**	Immunoassay	Serum Plasma		
	Plasmodium spp.	To screen blood donations for one or more human malaria species (*P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*) **in endemic areas**[60]	Light microscopy	Capillary whole blood Venous whole blood		

[59] Testing should be implemented in combination with the donor selection and deferral strategies.

[60] To selectively screen blood donations from individuals identified to be at risk of transmitting malaria, travellers or previous residents of endemic countries to detect antibodies to the *Plasmodium* species prevalent in their donor population.

II.c Disease-specific IVDs for blood screening laboratories *continued*

Organism	Screening test	Test purpose	Assay format	Specimen type	WHO prequalified or recommended products	WHO supporting documents
Treponema pallidum	Antibodies to *T. pallidum*	To screen blood donations for syphilis	Immuno-assay[61,62,63]	Serum Plasma	N/A	
Other transfusion-transmitted organisms	To screen for other blood-transmitted microorganisms (e.g. *T. cruzi*, human T-lymphotropic virus (HTLV I/II), Zika virus, *Babesia* species and West Nile virus in blood donations, depending on local risks of contamination.		Immuno-assay[61,62]	Serum Plasma	N/A	

[61] The only assays recommended for blood screening are those that have been validated for this purpose by the manufacturer.

[62] May be performed in laboratories with small throughput, in remote areas or emergency situations.

[63] In populations with a high incidence of syphilis, screening should be performed with a non-treponemal assay: venereal disease research laboratory (VDRL) or rapid plasma reagin (RPR).

IV. Do Not Do recommendations

The following test categories have been listed for discontinuation. These recommendations are based on either evidence of harm or a lack of benefit. Listings are supported by current WHO policies.

Disease	Diagnostic test	Recommendation	Assay format	WHO supporting documents
HIV	HIV western blot	Western blotting and line immunoassays should not be used in national HIV testing strategies and algorithms (strong recommendation, low-quality evidence).	Western blot assay Line immunoassay	Consolidated guidelines on HIV testing services (December 2019) https://www.who.int/publications/i/item/978-92-4-155058-1 WHO recommends countries move away from western blot and line immunoassays in HIV testing strategies and algorithms (27 Nov 2019) https://www.who.int/publications/i/item/WHO-CDS-HIV-19.30
Tuberculosis	Mycobacterium tuberculosis serology	Serodiagnostic tests for diagnosis of tuberculosis should not be used in individuals suspected of active pulmonary or extra-pulmonary TB, irrespective of their HIV status.	RDT Immunoassay	Commercial serodiagnostic tests for diagnosis of tuberculosis (2011) https://apps.who.int/iris/handle/10665/44652

SELECTED WHO PUBLICATIONS OF RELATED INTEREST

The selection and use of essential in vitro diagnostics
Report of the second meeting of the WHO Strategic Advisory Group of Experts on In Vitro Diagnostics, 2019 (including the second WHO model list of essential in vitro diagnostics)
WHO technical report series, No 1022, 2020
Website: https://www.who.int/publications/i/item/978-92-4-121031-7

Second WHO model list of essential in vitro diagnostics, 2019
Website: https://www.who.int/publications/i/item/WHO-MVP-EMP-2019.05

First WHO Model List of Essential In Vitro Diagnostics
WHO Technical Report Series, No. 1017, 2019
Website: https://apps.who.int/iris/handle/10665/311567

Further information on these and other WHO publications can be obtained from
WHO Press, World Health Organization, 1211 Geneva 27, Switzerland
tel.: +41 22 791 3264; fax: +41 22 791 4857; email: bookorders@who.int
order online: www.who.int/bookorders

This report presents the deliberations and findings of the Strategic Advisory Group of Experts on In Vitro Diagnostics (SAGE IVD) sessions held in 2020, which were convened to make recommendations on the test categories to be included in the Third WHO Model List of Essential In Vitro Diagnostics (EDL). SAGE IVD is tasked with acting as an advisory body on matters of global policies and strategies related to in vitro diagnostics (IVDs). The report describes the methods used to develop the third EDL, provides progress updates on the different associated products (eEDL, country guidance and technical specifications) and harmonization initiatives. The report contains the recommendations from the SAGE IVD on each test category together with a full description of the evidence considered for each test submission and requested edits. In addition, the report describes the evidence considered and the recommendations made for reversing conditional listings and Do Not Do recommendations. Finally, it presents general recommendations from SAGE IVD for the future direction of the list.

ISBN 9789240019102