

# The WHO Adverse Reaction Database

## Basic Facts

The WHO Adverse Reaction Database basic facts.....	1
Background and Contributors .....	1
The WHO database system.....	3
The WHO database .....	3
Format and data exchange .....	3
Controlled vocabulary.....	4
Medical terminology .....	4
Drug classification .....	4
Lexicon tables.....	5
Quality management .....	5
Documentation grading.....	6
Signal detection and analysis .....	6
Conclusions.....	7
Advantages of the WHO database system .....	7
Disadvantages .....	7
Further reading.....	8

## Background and Contributors

The WHO database holds more than 3,000,000 individual case safety reports (ICSRs) contributed by National Centres participating in the WHO International Drug Monitoring Programme.

The WHO Programme started in 1968, with 10 countries providing data from national spontaneous reporting systems. In 1978 the operational responsibility for the Programme was transferred from WHO, Geneva, to a WHO Collaborating Centre in Uppsala, Sweden. The centre is now known under its field name *the* Uppsala Monitoring Centre (UMC).

*the* UMC was set up as a foundation, responsible for the development of the international system on the basis of a two-way flow of information on suspected adverse reactions to drugs, in collaboration with National Centres. The main tasks of *the* UMC are to:

- Collect and analyze reports of adverse drug reactions from worldwide members of the Programme and to issue international *signals* of drug safety concerns arising from the data
- Actively support and provide training for both aspirant countries in establishing their own national drug safety surveillance systems, and for current members in maintaining and developing their systems
- Develop the science of *pharmacovigilance* (drug safety surveillance), in theory, methodology, practice and research
- Communicate the drug safety message throughout the world to all audiences

Currently, 72 countries participate in the WHO Programme, with another 14 countries being associate members (not yet actively contributing data).

	Full members	Associate members
Africa	7	6
Asia/Pacific	23	3
Europe	30	3
North America	2	-
Latin America/Caribbean	10	2

The table below shows the top ten contributors to the WHO database, by *number of reports* received in total from joining up to March 2002.

Country	# reports	Start year
United States	1,314,525	1968
United Kingdom	391,868	1968
Germany	160,648	1968
Australia	146,116	1968
Canada	136,192	1968
France	113,713	1986
Sweden	77,058	1968
Spain	71,993	1984
Netherlands	48,472	1968
Denmark	44,196	1968

The table below shows the top ten contributors to the WHO database, by *number of reports/million inhabitants/per year* (average 1996-2000)

Country	# reports/ mill. inhabitants
New Zealand	740.7
Australia	479.7
United States	416.1
Sweden	312.0
United Kingdom	310.8
Netherlands	305.7
Ireland	274.1
Denmark	220.8
Switzerland	170.4
France	163.8

The next table shows type of reporting and sources of reports, by country.

Type of reporting	Doctors	Other health professionals	Consumers	Marketing authorisation
Mandatory	9	5	1	28
Mandatory/Voluntary	6	1		4
Voluntary	41	38	22	8
Not accepted	0	12	33	16

# The WHO database system

## The WHO database

The data provided by the National Centres participating in the WHO International Drug Monitoring Programme is stored in a relational database management system (RDMS), Mimer 8.2. This system is ODBC (Open Database Connectivity) compatible, and uses SQL for the database communication. Mimer SQL adheres 100% to the internationally agreed SQL standard. The RDMS resides on a UNIX server, and is accessible through client-server applications, ODBC, and Internet applications (using standard web browsers as the interface).

The database is updated with incoming ICSRs on a continuous basis. National Centres are recommended to send reports at least quarterly; most National Centres adhere to these guidelines, but several report more frequently.

All National Centres have access to a web based search tool which runs against the current database. In addition, quarterly outputs based on data mining of the whole database are produced by *the* UMC, and distributed to the National Centres (see also Signal Detection & Analysis below).

## Format and data exchange

The current data model of the WHO database was designed in the mid 90s based on the data elements proposed in the CIOMS 1a document, which formed the basis for the ICH E2B format for ICSR exchange. Being E2B compatible, the WHO database allows processing of incoming data in the E2b format. Currently, E2b format ICSRs are received from the US, the Netherlands, and Norway, with more countries following in the near future.

Although the WHO database allows for the transmission and storage of a large amount of data for each individual case, there are few reports that have all the possible data elements actually filled in (*detailed figures on data elements filled in, overall, and by country, are available from the UMC; see also Documentation Grading below*). The issue of missing data, together with under-reporting, are well recognized problems for pharmacovigilance in general, but the fact that busy health professionals do not fill in all the expected data fields does not invalidate their concerns.

*the* UMC has, in collaboration with the Swiss authority, developed a web-based reporting tool. This tool allows a streamlined flow of information, both ways, from the original notifier, via regional centres and the national pharmacovigilance centre, to the WHO database. The data is stored directly in the WHO database, removing the need for extraction and transfer between different database systems. This reporting tool is now being made available to other National Centres.

A third option, for countries which do not yet produce an E2b format output or use the web-based reporting tool, is to send ICSRs to the UMC as ASCII text files, in the old WHO format, which is a sub-set of the E2b format.

## Controlled vocabulary

### ***Medical terminology***

At the start of the WHO Programme in 1968, hierarchical terminologies for coding adverse reactions (WHO-ART) and drugs (WHO Drug Dictionary, WHO-DD) were created, with the purpose of aiding data input and analysis. By linking the recorded case safety data to the corresponding classification, the data can be aggregated and analysed at different levels of precision.

National Centres may use either WHO-ART terms or MedDRA terms when reporting to the WHO database. Both terminologies allow for groupings and aggregation of data on different levels, from broad system-organ classes to individual signs and symptoms. The main difference between WHO-ART and MedDRA is the number of terms included (WHO-ART has around 2000 preferred terms (PT), and 3000 lower level terms (IT); MedDRA also has one more grouping level: high level grouping terms (HLGT)). On the PT and IT level, WHO-ART is a sub-set of MedDRA.

Until now, MedDRA terms in incoming reports have been mapped to the corresponding WHO-ART term by UMC staff; all ICSRs are coded in WHO-ART, and database output has been provided according to the WHO-ART hierarchy. However, since the database system is not restricted to the use of one medical terminology only, *the* UMC has decided to run WHO-ART and MedDRA in parallel. An obvious advantage in doing this is that there is flexibility both on the input and output side; those who so wish can continue using WHO-ART, whereas MedDRA reports will not have to be recoded.

### ***Drug classification***

The WHO Drug Dictionary is an integral part of the WHO database system. All medicinal products mentioned in ICSRs sent to the WHO database are coded by UMC staff according to the WHO Drug Dictionary classification. This applies to both drugs reported as 'suspected' of having caused the adverse reaction, and those reported as 'concomitant' or 'interacting'. In addition, newly registered drugs posted on the FDA and EMEA web sites are included, as are drug products coded on request by WHO-DD users.

Currently, the WHO-DD contains 63,000 drug records, of which 19,000 are generic (single ingredient/unique combination of ingredients). In the near future, *the* UMC will be able to extend the number of products considerably through an agreed collaboration with IMS Health which will lead to the loading of all product data in their system into the WHO-DD.

Since 1968, the following main data elements have been recorded for *each drug product* : product name, name source and source version, company, country, active ingredient(s), CAS numbers, therapeutic indication according to the ATC (Anatomical – Therapeutic – Chemical) classification.

Each drug entry is assigned a three-level hierarchical code, which groups products based on their active ingredient(s), base/salt of ingredient, and product name (the WHO Drug Record Number System). In addition, all drugs with the same ingredients are allocated the same 'preferred name'.

In connection with the implementation of the new, extended version of the WHO database system, the drug database was also extended, so that much more detailed information on medicinal products could be captured. The database model and the

nomenclature used for the data elements are based on the European Committee for Standardization (CEN) PreStandard prEN12610 Health Informatics - Identification of Medicinal Products.

For the purpose of international pharmacovigilance, it is unrealistic, and certainly not cost-effective, to populate the 'ideal' data set provided for by the new drug database format. Therefore, the *additional* level of detail currently used consists of a limited number of data elements such as form and strength.

The WHO-DD allows for the following levels of precision:

- ATC level, denoting the main indication for which a medicinal product is used; the ATC is in itself a hierarchy, with 5 levels;
- generic (ingredient/combination of ingredients) level;
- pharmaceutical product level (combination of ingredients/form/strength); and
- medicinal product level (referring to the named product marketed and sold in a particular country).

The WHO-DD also incorporates herbal medicinal products with a unique new classification system based on the ATC, and which links to internationally accepted botanical names and synonyms (in collaboration with the Royal Botanical Gardens, Kew, UK). This is an important development, considering the increased use of herbals and traditional medicines all over the world, and, therefore, the increased need for safety monitoring of these products.

### ***Lexicon tables***

The WHO database includes free text fields for e.g. background information and descriptions of the adverse reactions. However, most fields are linked to so-called *lexicon tables* which contain predefined, allowed values, expressed as formatted text or codes. When linked to a field in a database table, the lexicon table ensures that a value entered in that field matches an existing value in the lexicon table. Lexicon tables also allow for translations of values into different languages, as well as short and long text versions for each value stored.

## **Quality management**

As part of the processing into the WHO database, each incoming report is checked according to predefined quality criteria. Syntactic accuracy is obtained using controlled vocabularies (see above): entered values are compared and checked against reference dictionaries (lists of permissible data values).

Reports that contain rejected values are flagged, and subsequently examined by UMC staff. After correction of the problem, the report is re-processed. Missing data does not lead to a rejection, unless it involves one of the mandatory fields. Currently, these are: reporting country, case-ID, (at least one) medical term, and (at least one) drug.

Adverse reaction terms are checked against WHO-ART/MedDRA; drug names are checked against the WHO Drug Dictionary; the values in those fields that are linked to a lexicon table are checked against these. In addition to these checks, reports are also matched against a knowledge database with contains correct values for previously identified errors/synonyms to accepted values.

The main reason for rejection is due to the reporting of drug names which are not included in the WHO-DD. Many of these are mis-spellings, or drug names recorded using a different nomenclature as compared with that of the WHO-DD. These are corrected, and, when applicable, entered into the knowledge database. In case of new, valid products, UMC staff make the necessary updates of the WHO-DD.

## Documentation grading

In the 1990s a 'documentation grading' field was added to the WHO database. This grading is used for statistical purposes, and to identify problems related to missing data in the reports received. It is also used in signal detection, to facilitate the identification of well documented cases. Thus, in the output from regular screenings of the database, reported drug-adverse reaction combinations with a high documentation grading score are marked.

The documentation grading is based on the following core data fields (apart from the mandatory fields: reporting country, case-ID, adverse reaction, and drug):

- Age and gender
- Onset of reaction and treatment dates
- Patient outcome
- Drug dosage
- Route of administration

## Signal detection and analysis

Since 1998, routine data mining of the WHO database has been carried out using a Bayesian Confidence Propagation Neural Network (BCPNN), on a quarterly basis. The output is available to all 72 National Centres. The method has been tested, and over a seven year period nearly 50% of the drug-ADR associations found have been cited in the literature.

In addition to the availability of the data mining results, a triage procedure and expert clinical review are also undertaken by *the* UMC. The results of this further analysis are also available to all National Centres, and to pharmaceutical companies when a single branded product is in question.

Within the last two years, unsupervised pattern recognition using artificial intelligence, has been added to identify new complex relationships.

Another facility, which is used in further analysis, is to link the information in the WHO database with IMS Health sales and prescription data.

*the* UMC has made surveys of the usefulness and use of signals from the WHO Database, which have confirmed that National Centres find them both timely and valuable.

# Conclusions

## Advantages of the WHO database system

- 72 countries
- 3,000,000 individual case safety reports
- Whole populations covered
- Continuous data collection
- Low cost
- Possibility to identify differences between countries and make country comparisons
- E2b compatible database
- Legacy data from 1968 to 2002 converted to new format
- Controlled vocabularies
  - WHO Drug Dictionary
  - WHO-ART/MedDRA
  - Lexicon tables
- Quality management
- Quality grading of all reports
- Data mining in routine use (since 1998)
- Pattern recognition
- International panel of experts for clinical review

## Disadvantages

- Dependent on National Centres for timeliness, completeness and quality of reports
- Under-reporting (same as for any spontaneous reporting system)
- Missing data (made transparent by the 'documentation grading', and can be handled by the BCPNN)
- Various biases, but some analyzable by comparison of national differences

### Marie Lindquist

Uppsala Monitoring Centre

WHO Collaborating Centre for International Drug Monitoring

Stora Torget 3

753 20 Uppsala

Sweden

Tel. +46 18 656060

Fax. +46 18 656080

e-mail: [marie.lindquist@who-umc.org](mailto:marie.lindquist@who-umc.org)

## Further reading

The UMC homepage: <http://www.who-umc.org>

- Bate, A. The Use of a Bayesian Confidence Propagation Neural Network in Pharmacovigilance: Umeå University, 2003.
- Bate, A., Lindquist M., Edwards I.R., et al. A Bayesian neural network method for adverse drug reaction signal generation. *Eur J Clin Pharmacol* 1998;**54**(4):315-21.
- Bate, A., Lindquist M., Edwards I.R., Orre R. A data mining approach for signal detection and analysis. *Drug Saf* 2002a;**25**(6):393-7.
- Bate, A., Lindquist M., Orre R., Edwards I.R., Meyboom R.H. Data-mining analyses of pharmacovigilance signals in relation to relevant comparison drugs. *Eur J Clin Pharmacol* 2002b;**58**(7):483-90.
- Bate, A., Orre R., Lindquist M., Edwards I.R. Explanation of data mining methods. *BMJ website* 2001a:<http://www.bmj.com/cgi/content/full/322/7296/1207/DC1>.
- Bate, A., Orre R., Lindquist M., Edwards I.R. Pattern recognition using a recurrent neural network and its application to the WHO database. *Pharmacoepidemiol Drug Saf* 2001b;**10**(S1):S163 (abstract).
- Bate, A.J., Lindquist, M., Edwards, I. R., Orre, R. Understanding Quantitative Signal Detection Methods in Spontaneously Reported Data. *Pharmacoepidemiol Drug Saf* 2002;**11**, **Supplement 1**:S214-215.
- CIOMS. Harmonization of data fields for electronic transmission of case-report information internationally. CIOMS: Geneva; 1995: 48.
- Edwards, I.R. The accelerating need for pharmacovigilance. *J R Coll Physicians Lond* 2000a;**34**(1):48-51.
- Edwards, I.R. Pharmacovigilance - beyond 2000. *Reactions* 2000b(783):3-5.
- Edwards, I.R., Fucik H. Impact and Credibility of the WHO Adverse Reaction Signals. *Drug Inf J* 1996;**30**(2):461-464.
- Edwards, I.R., Lindquist M., Wiholm B.E., Napke E. Quality criteria for early signals of possible adverse drug reactions. *Lancet* 1990;**336**(8708):156-8.
- Edwards, I.R., Olsson, S., The WHO International Drug Monitoring Programme. In: Aronson, J.K., ed. Side Effects of Drugs, Annual 25: Elsevier Science B.V., 2002: 589-598.
- Farah, M.H., Edwards, I.R., Lindquist, M. Key issues in herbal pharmacovigilance. *Adverse Drug Reactions Journal* 2000;**2**(2):105-109.
- Finney, D.J. Monitoring adverse reactions to drugs - its logic and its weaknesses. Proceedings of the European Society for the study of Drug Toxicity 1966: 198-207.
- Finney, D.J. Systematic signalling of adverse reactions to drugs. *Methods Inf Med* 1974;**13**(1):1-10.
- Fucik, H., Backlund, A., Farah, M. Building a Computerized Herbal Substance Register for Implementation and Use in the World Health Organisation International Drug Monitoring Programme. *Drug Inf J* 2002;**36**:839-854.

- Lindquist, M. The WHO Programme for International Drug Monitoring: The Present and Future. In: Mitchard, M., ed. *Electronic Communication Technologies*. Buffalo: Interpharm Press Inc., 1998: 527-549.
- Lindquist, M. Seeing and Observing in International Pharmacovigilance - Achievements and Prospects in Worldwide Drug Safety: University of Nijmegen, 2003.
- Lindquist, M., Edwards I.R., Bate A., Fucik H., Nunes A.M., Ståhl M. From Association to Alert - a revised approach to International Signal Analysis. *Pharmacoepidemiol Drug Saf* 1999;**8**:S15-S25.
- Lindquist, M., Farah, M.H. , Edwards, R. Monitoring of Herbal Medicines. *Clin Pharmacol Ther* 2000b, Blackwell Science Ltd: 18.
- Lindquist, M., Pettersson M., Edwards I.R., et al. Omeprazole and Visual Disorders: Seeing alternatives. *Pharmacoepidemiol Drug Saf* 1996;**5**:27-32.
- Lindquist, M., Pettersson M., Edwards I.R., et al. How does cystitis affect a comparative risk profile of tiaprofenic acid with other non-steroidal anti-inflammatory drugs? An international study based on spontaneous reports and drug usage data. ADR Signals Analysis Project (ASAP) Team. *Pharmacol Toxicol* 1997;**80**(5):211-7.
- Lindquist, M., Sanderson J., Claesson C., Imbs J.L., Rohan A., Edwards I.R. New Pharmacovigilance Information on an Old Drug - an International Study of Spontaneous Reports on Digoxin. *Drug Investigation* 1994;**8**:73-80.
- Lindquist, M., Stahl M., Bate A., Edwards I.R., Meyboom R.H. A retrospective evaluation of a data mining approach to aid finding new adverse drug reaction signals in the WHO international database. *Drug Saf* 2000;**23**(6):533-42.
- Lindquist, M. Data quality management in pharmacovigilance. *Drug Saf* 2004 (in press).
- Orre, R., Lansner A., Bate A., Lindquist M. Bayesian neural networks with confidence estimations applied to data mining. *Computational Statistics and Data Analysis* 2000;**34**(8):473-493.
- Ståhl, M., Edwards I.R., Bowring G., Kiuru A., Lindquist M. Assessing the Impact of Drug Safety Signals from the WHO Database Presented in 'SIGNAL': Results from a Questionnaire of National Pharmacovigilance Centres. *Drug Saf* 2003;**26**(10):721-7.
- Ståhl, M., Lindquist, M., Edwards, I.R., Brown, E.G. Introducing Triage Logic as a New Strategy for the Detection of Signals in the WHO Drug Monitoring Database. *Drug Saf* 2003;**26**(10):721-7.