THE PERFORMANCE OF STAFFORDSHIRE AMBULANCE SERVICE – A REVIEW

An independent study carried out by the Medical Care Research Unit of the University of Sheffield on behalf of the Department of Health. The views expressed are those of the authors and not necessarily those of the Department of Health.

Janette Turner, Jon Nicholl

with the assistance of

Tom Judge, Mark Cooke

June 2002

Medical Care Research Unit
University of Sheffield
Regent Court
Regent Street
Sheffield
S1 4DA
Email: j.turner@sheffield.ac.uk
Telephone: 0114 222 0761
Fax: 0114 222 0749

Acknowledgement

We would like to thank the trust Board, management and line staff of Staffordshire Ambulance Service NHS Trust for their unreserved co-operation during this project.

Contents

List of tables

List of abbreviations

A&E	accident & emergency
AMPDS	Advanced Medical Priority Dispatch System
AVL	automatic vehicle location
CAD	computer aided dispatch
EMD	electromechanical dissociation
GIS	geographical information system
HPAS	high performance ambulance service
NVS	no vital signs
PALS	paediatric advanced life support
PEA	pulseless electrical activity
PHTLS	pre-hospital trauma life support
PI	performance indicator
PTS	patient transport services
PUM	public utility model
ROSC	return of spontaneous circulation
RTI	response time interval
SAS	Staffordshire Ambulance Service
SSM	System Status Management
UHU	unit hour utilisation
VF	ventricular fibrillation

Summary

Prior to the year 200/2001 Staffordshire was the only ambulance service which consistently reported achieving the response time standard of 75% of potentially life-threatening A calls responded to in 8 minutes. We have investigated the accuracy of this reported response time performance, the organisational changes that have been introduced to produce this improvement, the process of implementing the changes and the consequences for the service and staff. We have also assessed the sustainability of these achievements and the extent to which changes could be transferred to other services.

The main findings are:

- Staffordshire Ambulance Service has introduced a fundamental change in the practice of emergency ambulance deployment and response using the approach of “High Performance” services in the USA.

- The principal changes have been the introduction of a system status plan to match supply and demand; a dynamic deployment system to implement this plan which involves ambulance crews being based outside stations; the use of first responders; changes in the practice of control, dispatch and communications; continuous measurement and management of response performance; and a general and important change in the culture of the service towards a universal patient and performance focus.

- These changes are real, fundamental, and have been achieved in a large part because of the vision and leadership of the Chief Executive and Trust board.

- A detailed analysis of the reported response times found that real improvements had been achieved, although there was some evidence that a small proportion of response times were being under-recorded. This under-recording may have led to some exaggeration of the proportion of responses achieved within 8 minutes, but we estimate that Staffordshire Ambulance Service has nevertheless been achieving at least 75% of emergency responses within 8 minutes.

- The principal technical changes, which we believe have brought about the improved performance, are the introduction of rapid ambulance dispatch, System Status Management and the use of a number of first responder schemes.

- There is evidence of dissatisfaction amongst staff at all levels because of the increased pressures and lack of reward both in terms of pay and better training, facilities and accommodation despite improved performance. Nevertheless, the staff recognise the

significant achievements and few would go back to the old system of working. There has indeed been a dramatic shift in the culture of the organisation.

- The performance achievements are probably sustainable. This would be certain if resources to the Trust and rewards to the staff could be linked to their performance. If this were to happen further improvements might be possible. However, without additional resources we judge that improvements have now reached a plateau.

- The changes which have been introduced, and the improvements in performance which have resulted, have arisen from a sustained, 5 year process of change, adaptation and learning.

- Although the ideas are transferable we do not believe that they can all be implemented nationally in the short term. In our judgement, although rapid dispatch and first responder schemes might be introduced relatively quickly, implementing a system status plan is a longer term task. Furthermore, the improvements achieved in Staffordshire are, to some extent, dependent on characteristics of the Staffordshire Ambulance service area; the level of resources available to Staffordshire; the commitment of the Chief Executive and the Trust management to change; and the qualities and efforts of Staffordshire Ambulance service front-line emergency ambulance crews and other staff.

In summary, a significant part of the improvements in performance achieved by Staffordshire are real and sustainable with the right level of resources because of the change in the culture and a service wide commitment to a patient focused response delivery. We believe that the changes and improvements are transferable, though perhaps not as a whole systems approach, to other services in the longer term. This scale of change requires long term strategic planning. However Staffordshire has demonstrated that the High Performance approach provides a means of delivering national performance targets.

1. Introduction

The 1996 review of ambulance performance standards recognised a need to develop more clinically relevant performance measures[1]. A particular focus was the potential to save additional lives if shorter response times could be achieved for life-threatening conditions. The review recommended a long term performance target of 90% of life – threatening category A calls responded to within 8 minutes. Subsequently an interim target of 75% of potentially life-threatening A calls responded to within 8 minutes was introduced and services were required to achieve this target by April 2001.

Progress towards meeting the new standard has been slow. In 2000-2001 28/31 services had not achieved the revised target of 75% of A calls within 8 minutes[2] , although this number has since increased. Staffordshire Ambulance Service was the first service to report response times that have reached, and surpassed, this target and have shown a substantial improvement response times over the last 5 years. In 1994/5 55.8% of all emergency calls resulted in a response within 8 minutes[3]. By 200/01 the DoH performance table showed that 87.4% of category A calls and 81.9% of category B/C calls received a response in 8 minutes or less[2]. It is also reported that much of this improvement has been achieved without additional resources.

This exceptional performance was examined so that any transferable lessons could be identified and appropriate recommendations made to enable other services to move towards the new standards. The Department of Health therefore commissioned the Medical Care Research Unit to conduct a study of Staffordshire Ambulance service performance. The aim was to evaluate the impact of introducing an ambulance service High Performance System model and make an assessment of the extent to which this model is both transferable and sustainable in the UK setting.

1.2 Objectives

The objectives of this evaluation were;

1. To measure the performance of Staffordshire Ambulance Service and validate its performance data.
2. To catalogue and assess the key changes in the operation and management of Staffordshire Ambulance Service that may have led to improvements in performance.
3. To measure clinical performance and outcomes following out of hospital cardiac arrest.
4. To identify the transferable lessons which need to be disseminated throughout the UK ambulance services.
5. To assess the sustainability of this ambulance service model in the longer term.

1.3 High Performance Ambulance Services – The US model

The framework for the changes that have taken place in Staffordshire has been the adoption and adaptation of the High Performance model for Emergency Medical Service (EMS) systems.

The term “high performance” was originally developed in the USA by a small group of health economists, municipal officials, and ambulance managers involved in the redesign of EMS systems during the late 1970’s. A high performance system is one designed to deliver the attributes required by patients with out of hospital cardiac arrest namely; access, response, defibrillation, advanced cardiac life support and transport to hospital on a timely, reliable and reproducible basis[4]. Development of a high performance system requires enormous change and this has invariably arisen as a result of some high profile event leading to the decision or threat to remove current management and market test the service. Despite nearly twenty years of experience, there are still very few providers in the USA meeting the general definitions of high performance[4]. Those that do are now defined as public utility models (PUMS) the key features of which are[5]:

1. A competitive bidding process to contract for the cost-effective provision of services for the entire market.
2. The award of performance based contract with exclusive market rights for the given community to a single provider. Continual performance monitoring.
3. Contractual penalties for failure to perform including removal of the contract if necessary.

These PUMS broadly resemble the NHS Trust model for the provision of ambulance services in that there is a purchaser provider relationship and a single provider delivers services to a defined area. Unlike the NHS system, the model is based on exclusivity for the entire market area on a reasonably long-term contract – usually 5 years with renewals to allow the provider to recapture initial investment. Contracts for services are defined within a given envelope and additional services or activity exceeding an agreed cap are reimbursed by purchasers. High performance is rewarded by the ability to retain productivity gains and providers are allowed to “make” money on non-emergency services to offset costs of providing the emergency service. The PUM concept requires the entire jurisdiction to be contestable rather than any individual component and this allows the provider to control and manage all services.

The evolution of the term “high performance” has been often been interchanged with another, “system status management” or SSM, a descriptor of the dynamic ambulance deployment necessary to achieve response reliability. SSM is a system of sophisticated resource management targeting deployment of mobile resources within economic and clinical goals. High performance is the outcome if the system operates well. However SSM is one strategy within a wholesale change process. Development of clinical quality of care, maximising productivity of available resources (including provision of support services, for example fleet

management) and changes in organisational culture are also required if the high performance objectives of highest quality of service at least cost are to be achieved.

As described by Jack Stout, a health economist and the individual most often associated with the development of the concept:

> A "high performance" EMS organisation earns that title by producing extraordinary service from modest financial resources, often under difficult market conditions. Patients of high performance organisations receive the best service possible from the dollars available, or very nearly so. ……..High performance is rarely and perhaps never achieved without sincere opposition, risk to career, false starts, misunderstandings, and outright mistakes. In every case it is harder than it looks"[6].

1.4 UK Environment and drivers for change

Demand for emergency ambulance services has increased by 50% over the last ten years whilst funding has increased in real terms by 17.5% during the same period[7]. This has created an impetus for change in the provision of ambulance services within the NHS. Improvements have to be made as some services are still failing to meet the current 8 minute emergency response standard [2].

Concurrent with the requirement for improved response time performance the need to improve value for money has and continues to lead to substantial reorganisation of the ambulance services through mergers and replacement of senior management. Although there is good practice in many Trusts[8], and despite changes that have already taken place, there remains a gap between expected and obtained performance. The developments at Staffordshire ambulance service may provide useful strategies for other services to adopt in their efforts to close this gap.

2. Organisational objectives and change process at Staffordshire Ambulance Service

The move to a High Performance Ambulance Service (HPAS) requires major organisational change both in terms of the structure of the organisation and the working conditions of staff. In order to describe this process and identify transferable lessons the key developments during the transition period were elicited and assessed.

2.1. Methods

Semi-structured interviews were held with both current and former Staffordshire ambulance service staff. In order to assess the impact of change across the whole Trust opinions were sought from staff at each level in the organisation. This included executive and non-executive directors, senior and middle managers from different parts of the organisation, union officials and field operations including front line ambulance crews.

A formal questionnaire was not used for these interviews as the aim was to gain as broad an overview of events as possible rather than restrict the conversations to a narrow range of topics. However a small number of key questions were used to guide the interviews and provide some structure. These were

- Brief personal details – length of service, role, specific responsibilities in implementing HPAS
- The perceived advantages and disadvantages of HPAS
- The features of HPAS considered to be critical to improving response times and the relative ease or difficulty of implementing these features
- Impact on staff
- Which features are transferable and which are not
- Whether this level of performance is sustainable over the long term. If yes, why; if no, what is needed to sustain it.

These interviews were conducted by the researcher (JT) and an independent EMS consultant from the USA (TJ).

Much of the achievement of improved performance is dependent on the road crews. This is the biggest single group of employees and it is in this group that working practices have substantially changed. Therefore, additional interviews were held with ambulance crews to assess the impact of HPAS on frontline staff. Over a 2 day period randomly selected paramedics and technicians in North & South divisions were selected and asked if they were willing to be interviewed. Eight crew members were interviewed and a further 11 completed a paper questionnaire left at the posting points visited during the course of these two days. No one refused to be interviewed. These interviews were conducted by a graduate paramedic from another ambulance service (MC), who also went out on calls with crews. The results of

these interviews were combined with the crew interviews conducted by the other two researchers. A total of 22 crewmembers were interviewed. Their ages ranged from 24-58 years and length of service ranged from 7 – 19 years.

Interviews focused on their views of HPAS and the impact it had had in terms of job satisfaction, staff morale, welfare and stress and working conditions. SAS has conducted its own independent staff satisfaction survey[9] and where similar questions were asked in this survey we have included the results as a comparison.

In addition to the interviews, key trust documents and papers describing the implementation process and strategic development of Staffordshire ambulance service were also examined.

2.2 Baseline service and barriers to service development

A description of the baseline service has been developed from interviews primarily with the Chief Executive and the Trust Chairman and Directors. The CE identified a number of features that describe the organisational culture at the start of the change process.

i) Background

As with EMS systems in the USA, the impetus for change at the Staffordshire Ambulance Service NHS Trust (SAS) was the need to overcome a number of threats. The service was faced with a widening gap between resources and performance and was under pressure for a potential merger. The move to Trust status in 1992 saw the start of a strategy with high performance objectives. Later, a number of patient transport service contracts were lost and competition from another service instigated the possibility of market testing for the contract for emergency ambulance services in Staffordshire. This potential threat was used to persuade staff that change was necessary.

ii) Baseline service

A number of features identified by the CE describe the characteristics of the organisation when he arrived:

- There was a lack of information on virtually every component of service delivery.
- There was a view that little strategic planning was taking place.
- No rigorous demand analysis was available. There was virtually no information technology support and consequently there was little information about hourly, daily, and seasonal variation in demand.
- There was no standardisation of processes. ("15 stations = 15 methods of doing the same thing.")
- There was a hierarchical structure in the ambulance service. Station officers were always last out on calls and consequently the most available under-utilised resources resided in protected management positions.

- Promotion was on length of service not merit.
- There were conflicts between control room staff and field crews.

In late 1994 the reorganisation of the Trust was complete, demand analysed and a rudimentary plan produced. The concept of SSM was introduced by senior staff with personal experience of the system and this was adapted and adopted as the way forward.

As described in a 1995 Trust document[10], the scope of change included the following goals:

a) Highest quality for least cost

b) The maximum potential from each member of staff

c) Response time standards of:

- 8 minutes in 90% of all life threatening cases
- 12 minutes for all other emergency calls, within 5 years

d) Provision of medical care to strict protocols, where compliance and outcomes are closely audited

e) Provision, through a variety of agencies, of rapid basic life support including early defibrillation, to patients facing life threatening medical emergencies

f) Capability to react effectively to major incidents

g) Radical improvement to training and development

These goals and objectives provided the framework for the subsequent changes and developments in organisational and operational practice.

2.3 Key features of the change process

The change process has been described during interviews with the CE, Trust Directors and Trust Chairman. Change was implemented at several levels using a top down approach. These were principally:

<u>i) Leadership:</u>

Having developed an agenda for change the CE had to persuade existing senior staff and the non-executive Board to accept fairly radical and immediately unpopular ideas for change. The key lessons and issues identified were:

- The CE was not from an ambulance background and not tied to the traditional ambulance service culture. However experience of military medical operational planning helped effect change. The business background of non-executives was also considered essential.
- Quality Management and Quality Control industrial process techniques to measure and implement change were introduced

- CE and the Board were committed to the NHS as a public service. It was considered necessary to create a service which “accommodates the needs of patients not employees”.
- CE and the Board were clear that change could be accomplished without new resources.
- CE and Board were certain that the medical model for change found in the USA could be replicated it in the UK . Senior staff visited “high performance” sites in the USA.
- CE and the Board believed the entire plan must be accomplished as all components are interrelated and accepted there was a consequent risk of failure.

ii) Staff:

The second step was to develop support at staff level.

- Once senior staff understood the system, the Trust sent staff, including control, line, and union stewards for hands-on training at HPAS sites in the USA. Open management and support of staff during the change was seen as imperative if the plans were to succeed.
- The Trust agreed to rely wherever possible on current staff as the workforce needed to feel secure in their positions during major change.
- The Trust recruited several key members of staff including a senior production manager, a doctor and a nurse specialist to lead clinical changes and paramedics with knowledge and experience of HPAS in the USA
- The use of the ‘unit hour’ as the operational currency was developed and unit hour utilisation as the measure of efficiency. A unit hour is one fully crewed emergency ambulance available or tasked for one hour. Unit hour utilisation (UHU) reflects the use of these unit hours. SSM measures unit hours including surplus and wasted hours. Measurement and understanding of how hours are produced and then utilised and where and why differences between demand and production exist provides information for planning and the efficient use of resources. As a result of SSM Staffordshire uses UHU as the budgeting tool. Each year they set a trust UHU efficiency target (currently 0.51 - 0.50 is considered high performance in US urban based PUMs) and plan to match available resources and demand as close to budget as possible. This then forms the basis of resource requests.
- Operations were divided into two parts. Production is responsible for fleet operations, field operations, estates, supply, and scheduling. Distribution manages contracting, demand analysis and system status plans, control and dispatch operations and IT.
- The team concept was introduced to replace the traditional station model. Team leaders were used to brief the team about developments and changes.

iii) The public

Reorganisation of an ambulance service on the scale undertaken by Staffordshire clearly has implications in terms of public perception of change and the need to maintain public confidence in the service. Of particular importance is the impact of moving from a relatively small number of fixed ambulance stations to a larger number of response posts and standby points. Closure of ambulance stations is a sensitive issue as these provide a more permanent reminder of the presence of the ambulance service in a community than vehicles at standby points for short periods.

A number of strategies were employed by SAS to gain public support and acceptance of their plans including:

- Adopting the same open style of access to information that was used to gain staff support
- Conducting a series of public meetings where service options were presented before decisions were taken on the closure of ambulance stations. The public was allowed to choose whether or not an ambulance station remained in their community
- Proactive press briefings and public relations
- Establishing paramedic first responders in some communities before any station closure took place.
- Proactive use of other ambulance crew members to inform and reassure the public in their own local communities.

We have not attempted to measure public satisfaction but the Trust has conducted a patient satisfaction survey[11]. Although there was a very low response rate (7.5%), the responses they did receive indicate a very high level of satisfaction with the speed of response and the service provided by crews. There has also been the successful implementation of a number of community first responder schemes operated by volunteers from the public in partnership with the ambulance service in rural and remote areas.

Having set Trust goals and made changes in the structure of the organisation a number of key operational strategies were introduced to facilitate the transition to a high performance service and the implementation of system status management as a framework for actual delivery of improved performance. Some of these were directly concerned with implementation of SSM. Others were adjunct developments designed to improve response time and clinical performance and improve efficiency by increasing productivity of available resources in order to meet the high performance objectives of best quality care for least cost. The principal developments have taken place in unit hour distribution and production which are discussed in more detail in the following sections.

3. Key Developments in Unit Hour Distribution

The principal function of unit hour distribution is to determine how many ambulances are required at a given time, where they will be needed and to then deploy vehicles and crews in response to demand. Operationally this means responsibility for IT and communications and control room operations. A number of planning and deployment processes aimed at improving response time performance have been introduced, namely:

1. Development and implementation of the system status plan
2. Introduction of paramedic staff into the call taker and dispatcher roles
3. Rapid deployment in response to a call
4. Improved communications with crews
5. Introduction of priority dispatch

3.1. Development of a response plan.

Staffordshire Ambulance Service has adopted and adapted to the UK setting, the US model of System Status Management (SSM) to maximise performance and efficiency. This is achieved through a variety of processes:

- Better positioning of vehicles in locations determined by geographical and clinical urgency demand patterns to reduce the distance between incident and responder
- Improved temporal allocation of resources by matching supply of vehicles with expected demand
- Dynamic deployment with available resources being moved around to maximise efficiency.

Stout describes SSM as “the formal or informal systems, protocols and procedures that determine where the remaining ambulances will be when the next call comes in”[12].

The starting point for this process is detailed analysis of demand patterns and the development of system status plans (SSP’s). The original plan in Staffordshire was developed using demand analysis and a purpose designed Geographical Information System (GIS). This plan was introduced in November 1994. Subsequently, activity was examined in 10 hour blocks over a 10 week period. From this information 168 plans (1 for each hour of the day for 7 days) were developed. The plans reflect both demand and clinical urgency (and hence the resources required to meet that demand) and where the demand occurs (and hence where available ambulances need to be placed). In the first instance planning the positioning of ambulances was a time consuming task. Groups of staff from all levels were involved and were asked, if only one ambulance was available where should it be placed to get to most calls. This was repeated for 2 ambulances, 3 ambulances etc and all combinations were explored. The most favourable positions were then incorporated into the

system status plans. This process resulted in the second plan, which became operational in 1996 when a new SSM computer aided dispatch (CAD) system was installed.

To be effective the SSM plan must be understood and adhered to by control staff. There are many processes that facilitate the system status plan working in practice but the dispatcher role is crucial to the overall success of the plan. Dispatching in a HPAS environment is a complex task and is not confined to tasking crews. It also involves the continual review of available resources and moving these resources around in accordance with the SSP in order to use crews most efficiently and effectively. The success or failure of the SSM plan is therefore critically linked to dispatcher performance and dispatchers are subject to continuous performance measurement. In order to support this role SAS has made considerable investment in its IT systems. This includes priority dispatch, automatic vehicle location and geographical mapping, automatic paging of information to road crews, computerised information systems for managing the system status plan. The Staffordshire CAD system has a reporting system designed to manage the SSM plan and is capable of prompting control staff to shift resources on a continual basis to stay within plan—matching resources with demand both temporally and geographically.

A more detailed analysis of the impact of SSM on response time performance is given in chapter 5.

3.2 Introduction of paramedic staff into the call taker and dispatcher roles

Recruiting operational staff to control is a common strategy in high performance systems in the US. Controllers are seen to need to understand and "feel" street response—routing, road timing, whole event timing. With the exception of 5 protected control staff posts, call taking and dispatching are now performed by paramedics who rotate, on a voluntary basis, through control from the road staff. Paramedic staff also manage the control room.

The perceived advantages of this change are that:

- Paramedic staff are more familiar with the environment the field crews operate in
- Field crews respond more positively to paramedic's requests, particularly when they are asked to move around between posting/standby points.
- Rotating paramedics through control gives field crews a better understanding of the dispatcher and call taker role and the need to continually move around resources
- Paramedics in control can also act as first responders at times of peak demand
- Management and decision making training is provided for paramedics and staff learn in more detail about SSM and how the system operates.

Interviews with key staff have revealed conflicting opinions about the extent to which the anticipated benefits have been realised. It is generally accepted that these advantages can

work in practice. However, there are also disadvantages in adopting this approach. The main problems identified are:

- Although relationships between some field crews and some control staff are better, difficult relationships still exist. In particular, the continual pressure on crews and control staff to achieve high targets for 8 minute responses creates tensions and conflicts. This can result in a blame culture existing between some field staff and distribution staff.
- Employing primarily paramedic staff in the control room is expensive. Some consider this a worthwhile investment both in terms of control/field operations relationships and the provision of additional resources for responding to calls, both of which can impact on performance. Others consider the use of paramedics in control should be confined to the dispatcher role and not call taking. It is estimated, by SAS, that employing paramedics as call takers costs an additional £75,000pa.
- Using control staff to respond to calls creates pressures on remaining staff whilst they are absent.

The impact of this change on response times is most likely to be indirect in that the effect is concerned with the relationship between control and field staff. Where this relationship is positive there will be high compliance of road crews with control instructions. Where it is poor opportunities arise for both crews and dispatchers to negatively impact on each others performance. The extent to which the benefits of employing only paramedics in the control room outweigh the disadvantages is difficult to quantify but the necessity for this change as an absolute requirement for HPAS operations remains open to question.

3.3 Rapid deployment in response to a call

Immediate dispatch from the control room can reduce activation time and hence contribute to improved response time intervals. In Staffordshire call takers and dispatchers sit together and both watch the call entry screen. The dispatcher alerts the crew nearest to the incident as soon as a postcode is obtained. The clock starts when the chief complaint has been obtained or after 45 seconds, whichever is the sooner. In addition, crews now spend more time in their vehicles with consequent faster mobilisation. Analysis of activation times before and after the introduction of HPAS changes shows mean activation time has decreased from 3.2 minutes in 1995 to 1.7 minutes in 1999. However, this is the result of eliminating some very long activation times. The median activation time has changed very little, from 60 seconds in 1995 to 62 seconds in 1999. This is confirmed by SAS who are confident that the real improvements in activation times occurred in 1994.

3.4. Improved communications with crews

In Staffordshire all crews and other operational staff are provided with pagers and incident information data is rapidly transmitted from the control room CAD system via these pagers. There is therefore less reliance on the more conventional telephone or radio links that can

produce delays, for example, when crews are away from their vehicle. This contributes to reducing long activation times The system also allows some flexibility for crews as they do not need to rely solely on radio contact. They may, for example, choose to use home rather than a posting point in the same location.

3.5. Introduction of priority dispatch

Concurrent with other substantial developments in control room IT, the Advanced Medical Priority Dispatch System (AMPDS) of call prioritisation has been introduced. The principal driver in Staffordshire is to respond to as large a proportion of calls as possible within 8 minutes. There is therefore, a conflict with the principles of priority dispatch. Immediate dispatch when an address is available, and

a hot response to every call irrespective of urgency means that use of prioritisation at the time of dispatch is limited to providing information on where resources can be diverted if the system is under pressure. As such prioritisation makes little impact on response times in this operational environment. A key role prioritisation does play is in the provision of information on call urgency, which is then incorporated into the system status plan.

3.6 Impact on road crews

Interviews with crews produced a generally positive response to the principles of a High Performance Ambulance service. Staff who stated that they did not like HPAS did acknowledge that the system worked to provide a better response and more efficient ambulance service. In particular they felt that HPAS was better for patients because they received a quicker response. However there were some concerns that patients in outlying areas often received the slowest response or that of a ‘first responder whilst calls from a town centre will get the fastest response as ambulance provision is concentrated around high demand areas. Whilst the staff understand the logic behind this there is a perceived feeling of ‘inequality of service’ as the nature of calls in these high demand areas is often minor.

Crews were asked if the 82% performance figure for 8-minute responses was a true reflection of actual performance. The majority were in agreement that it probably was a true reflection although there were a few comments from individuals who thought performance was not at this level and that Control were ‘editing’ calls. One paramedic commented that recently Control had provided a response time of less than eight minutes for a job that was over 10 miles away. There were several comments that the details of job times passed from the control room, via their pagers, do not always match the times recorded by crews on the patient report forms and that this happened often.

Although staff agreed with the principles of HPAS the practice has led to some dissatisfaction. There was a common perception that HPAS means there is often not enough vehicles to provide adequate cover, particularly at night, which results in greater pressure being placed upon crews. However, there was consensus that under the HPAS system workload is much

more evenly distributed between crews when compared to the old way of working where urban based crews felt they responded to far more calls than rurally based crews. Crews also liked the flexibility that data transmission via a pager gave them as they are not tied to their vehicles.

Staff were asked their opinions about responding 'hot' to emergency calls when they did not know the nature of the call. Whilst the majority of staff interviewed recognised the need for a fast response to achieve the eight minute response targets, some felt uneasy about this type of response. A typical response was 'I don't drive fast until I know what I am going to'

When asked about how other members of staff feel about 'hot responding', one paramedic commented that probably half of the front line emergency staff feel the same way. The relentless pursuit of fast response times, regardless of the nature of the call, the need to drive fast and pressure to be 'green and available' were all cited as reasons for increased stress with the HPAS system. A common response from interviewees was that morale was higher with the old system of working. One lead Paramedic commented; "They need to back off on the push for response. We know we can get to calls we need to. There is too much pressure on crews for non-life threatening problems."

3.7 Summary

A number of strategies undertaken by the distribution function in a HPAS can potentially impact on response time performance. Principally these are:

- System status planning and operation of the SSP which places resources in the right place at the right time
- The combination of rapid deployment, improved communications and crews being more likely to be in their vehicles which results in a consistent activation time performance of 1 minute or less and
- the availability of a range of alternative responders which provide the dispatcher with additional resources at times of peak demand.

However, rapid transfer of incident information, and other developments such as pre-alerting crews as soon as a postcode is obtained do not in themselves produce shorter activation or response times. It is whether the crews themselves respond quickly, or not, to this information that is the real determinant. Similarly moving vehicles around in accordance with the SSP can only produce benefits if dispatchers adhere to the plan and crews comply with the instructions given by dispatchers. Planning and sophisticated technology provide the practical framework for implementing SSM but ultimately the success of this strategy is dependant on staff buying in to and understanding the system and subsequently carrying this through in terms of the way they actually work.

4. Key developments in Unit Hour Production

Unit Hour Production provides distribution with the unit hours determined by the SSP at the times these unit hours are required. Unit hours are purchased by distribution from production at an agreed cost and quality. The key areas of production responsibility are:

- Field Operations
- Scheduling
- Estates, fleet and supply

4.1 Field Operations

Field operations is essentially concerned with organisation and management of field staff and has responsibility for personnel performance, welfare and discipline, health and safety, field training and team management. Organisationally field operations is divided into 2 divisions – North & South – and each division is supported by a divisional officer. Paramedic and technician staff are organised into 21 teams of 10 people supported by 10 team leaders.

Teams

The team concept was introduced to support staff and replace the traditional station culture following the restructuring of the fixed station model. The team approach is seen as essential for gaining co-operation, staff development and training and the distribution of information. It was also used to provide some internal promotions with the post of team leader. Each team provides 168 unit hours per week and organises their own rotas. Some teams provide ambulance cover over 24 hours and others during peak times only.
The development of some sort of team strategy is considered by SAS to be essential to the process of implementing HPAS as the disappearance of stations also means the loss of an important support network. Teams also provide the information that is critical in enabling crews to understand and accept why changes are being made. The experiences in Staffordshire indicate that without this structure other organisational changes are difficult as staff either fail to fully comply or deliberately flout operational protocols. The opinion of senior managers is that, although a long lead-time is required to set up and gain acceptance of the team concept, it is a relatively easy task. A more difficult task is the process of moving crews out of stations and into an environment where they spend most of their time in their vehicles.

Crewed vehicles are also provided by “pool” staff who are not members of teams but provide cover at short notice. These staff fill in shortfalls in unit hours required by the SSP and often only get 24 hours notice of when they will be required. Pool staff are not attached to teams although they can be assigned temporarily to a team to cover, for example, long term absence. However, this does not provide the stability of permanent membership of a team. For some staff the flexibility of pool working and not being tied to rotas is advantageous, however, for others, who would prefer to work in a team, it is viewed as a disadvantage.

There has been a trend in SAS to increasing the number of pool staff and decreasing the number of team staff.

Other Staff

Not all responses to emergency calls are made by fully crewed ambulances provided by the teams. Operation of the SSP requires sufficient unit hours to be available to match predicted demand. This cannot always be achieved within the constraints of crew rotas and there will sometimes be shortfalls. Similarly the plan is based on average expected activity and there will sometimes be unexpected peaks in demand or, conversely, expected low demand (for example at night) with consequent low utilisation of crewed ambulances. In keeping with the HPAS principle of high performance within available resources SAS has developed alternative strategies to resolve these problems rather than simply providing more fully crewed ambulances. Principally this means the use of a variety of first responder options, namely:

- 7 community paramedic officers (CPO's) who are provided with a vehicle and equipment and work from home. CPO's work shifts although some volunteer to be called out at any time. They provide a first response service and clinical support to lay first responder schemes. They also develop clinical links and networks with the wider health community, particularly primary care, and as a consequence have reduced the number patients requiring transport to hospital by 20%. The CE considers that CPO's have played an important part in gaining public acceptance of the changes and maintaining public confidence in the service. Explanation of the role of first responders to the public, and putting them in place before ambulance stations are closed, are considered by the Trust to be key factors in implementing change.
- Management and control staff are expected to staff ambulances or provide first response at peak demand periods in order to maximise resource allocation. The underlying principle is that any qualified employee can be tasked to a call and all have access to equipped vehicles. There are currently 33 of these support vehicles. These staff, together with the CPO's and community lay responders also provide a voluntary "on call" service allowing additional response flexibility that can be utilised outside peak demand hours, for example, at night in rural areas.
- In some rural areas community first responder schemes using volunteers from the community have been established. Basic training, including the use of automatic external defibrillators (AED's) and driver training, is provided by the ambulance service. Otherwise the schemes are financed by the local communities and are managed in partnership with the ambulance service.
- SAS have also trained their Patient Transport Service (PTS) staff in the use of AED's. The PTS fleet now has AVL and PTS crews provide a first response service.

The use of a variety of first responders is now an integral part of the day to day management of emergency calls and account for the initial response to 10-12% of emergency calls. The effects are twofold. Firstly, use of first responders who are nearer to the incident has reduced some of the very long distances that previously resulted from tasking vehicles from distant stations with a consequent reduction in response times. Secondly, the availability of operational staff for response provides the dispatcher with additional resources at times of peak demand. This results in shorter response times as dispatchers do not have to wait for a vehicle to become available. SAS figures show that use of first responders produces a 5.5% improvement in response time performance.

Impact on staff

In general, the movement to teams has been widely accepted and supported by the staff we interviewed. Dissatisfaction was around the issue of pool staff being unable to gain team posts. This is supported by Staffordshire's own survey where staff reported they were very satisfied with the team philosophy and felt their skills and experience were valued within their teams. A high proportion of responders also considered their job provided them with the opportunity to use their skills and initiative.

The use of first responders produced mixed views. For some staff this is their primary role and, for example, the CPO's enjoy the flexibility that working from home allows them and the opportunity to develop their role. However, for others there are disadvantages. Responding to calls, on a regular basis, means that substantial amounts of time can be spent away from an individual's real job. This makes their task more difficult, and regularly disrupts scheduled work so staff are continually trying to catch up. It also makes planning difficult. Furthermore, the first responder role extends outside normal working hours and many members of the service respond to calls voluntarily in their own time. Whilst this in itself is not a problem if properly agreed, there is an implicit expectation that staff can be called at any time. This leaves some staff feeling that they are "on call" 24 hours a day. One suggestion, made by staff, was that deployment of, for example headquarters staff, could be rotated with some protected time to enable other tasks to be completed without interruption. There is also a perception amongst some that to refuse to undertake this work will seriously impede career progress within the organisation.

Clinical operations and training

Improved response times do not necessarily equate to improved clinical performance. Impact on patient outcomes is dependent not only on rapid response but also appropriate and timely action once the ambulance crew arrive. In the early transitional period there was a focus on the provision of training and an acknowledgement that clinical training is as important as response times. Staffordshire employed a director level physician in operations and training and a senior RN as clinical specialist for operations audit in order to maximise clinical

performance. There was detailed review and reinforcement of clinical protocols by a strong team of clinically focused managers. Senior mangers see this as a key area in facilitating progress towards a patient focused service as it clarifies the patient benefits and increases the confidence of crews in carrying out clinical protocols. Strategies adopted include:

- The production of a medical procedures manual
- Improvements in record keeping with the emphasis on their contribution to clinical audit
- In the early phase additional training was provided. All staff received 5 days clinical skills development to enhance and improve what were perceived as poor clinical skills.
- Additional courses in Pre-hospital trauma life support and paediatric advanced life support.

A scheme of clinical supervisors was introduced to help maintain clinical skills but did not prove to be successful. Similarly, early in the process training was done within teams but now this is not the case. This is seen by some as a retrograde step as it has removed a valuable opportunity to maintain and support the team concept.

Impact on staff

Interviews with crews have revealed that the combined effects of faster response and improved clinical skills are key factors in terms of both gaining acceptance of the changes implemented and sustaining motivation in a HPAS environment as they genuinely consider that they have made a difference to patient care. However, as HPAS has developed there is a perception amongst some staff that some of the focus on improving clinical performance has been lost. Some additional training is provided, for example, PHTLS and PALS courses. However, the pressure to maintain the high level of response time performance means that sometimes, during statutory training, crews can be called away to respond to calls or that training is cancelled and has to be re-scheduled. It is felt by some to be similarly difficult to provide adequate training when new protocols are introduced. The SAS staff survey showed that, although a high proportion of respondents felt that training and development had helped them do their job more effectively, more than half felt that they did not have adequate access to training and development opportunities.

The push for efficiency also means that staff numbers are strictly controlled and maintained at levels necessary to implement the system status plans. There are some concerns that the long term effect of this strategy will make it increasingly difficult to maintain the training cycle and provide a suitably qualified pool of staff so that, for example, people leaving the service can be immediately replaced. Funding for training is ring fenced but at what SAS consider to be inadequate levels. If the service were allowed to keep some efficiency savings this could be invested in the training programme. This problem is exacerbated by the requirement for ambulance service trusts to finance the training of their staff.

4.2 Fleet and estates

A critical component of the SSM plan is the placement of vehicles according to predicted demand and available resources. This necessitates a radical re-structuring of the traditional fixed station model. The key features in Staffordshire to support SSM are;

- A network of 42 response posts across the county. Some of these are leased rooms in, for example, hospital or industrial premises that provide crews with access to food and drink, somewhere to sit and toilet facilities. Others are standby points on roads, car parks etc.
- 3 Depots - These provide a make - ready service, that is, centralised operations for fleet management which includes cleaning, maintenance and stocking of vehicles. A cleaned and stocked vehicle is provided for each crew at the beginning of each shift so they are immediately available for tasking. Previously, crews performed cleaning and stocking of vehicles.
- 4 central reporting posts – Crews report to one of these posts or a depot at the beginning of their shift rather than at a station. This is a necessary change with the implementation of SSM but does result in some disadvantages for staff in that some now have to travel much further distances at the beginning and end of shifts.
- 7 community paramedic officers who work from home in areas where stations have been closed. These are located in more rural or less densely populated areas outside the main conurbations where demand is lower.

This process included a substantial review of assets and the elimination of those that did not support core activities, for example stations in geographical areas that did not fit the SSM plan.

The fleet was also reviewed in terms of design, preparations, maintenance, equipment, supplies and replacement resulting ion a number of developments including:

- . The introduction of the make ready programme which has been a key factor in producing efficiency gains
- Ambulances are purchased with staff input and designed for comfort as crews now spend much more time in their vehicles
- Ambulances are supplied with self loading stretchers to reduce fatigue and injury.

<u>Impact on staff</u>

Few of the people interviewed said that they would return to the old system of working which provided an inadequate and ineffective service. All the staff interviewed stated that they did enjoy and value their jobs and commented that the interaction between themselves and the patients was important and that patient care was the priority. However, they felt more stress exists with HPAS than with the old system. Other sources of dissatisfaction were:

- Travelling distance to work with the introduction of central reporting

- Loss of station camaraderie and support. Staff are concerned that they rarely see their work colleagues and consider ‘chatting with colleagues’ as an important aspect of debriefing and coping with stress. The loss of stations has reduced the opportunities for staff to support each other and often the only person they see is their crew mate who has been involved in the same calls.
- Staff acknowledged that the use of posting points and standby, being mobile quickly and not being sat on station were the factors responsible for the high response performance. They felt that a valuable lesson for other services would be to get the posting points in place before closing stations. In Staffordshire this was a gradual process over time and was a considerable source of dissatisfaction during the implementation period. There were also several comments that the posting points are of poor quality and could be better maintained.
- Many of the staff interviewed were unhappy about their salary and vehicles, and said that the CE had not kept his promises of ‘high pay for high performance’ and ‘better vehicles’. Staff felt aggrieved that the Staffordshire Ambulance Service were recognised nationally as having the best response performance and that this credit was not being handed down to crews in the form of enhanced pay. As a result staff are demoralised by their perception that there is little differential in their pay compared to paramedics and technicians in other services with poorer performance.

The SAS staff survey supports some of these statements but conflicts with others. In general, the survey reported a high degree of job satisfaction and motivation. There was also a high degree of satisfaction with uniforms and equipment and a reasonable degree of satisfaction with vehicles and facilities. However staff reported dissatisfaction with pay, particularly in relation to the pay received by other emergency services. There was also dissatisfaction with arrangements for meal breaks and specifically the opportunity to take breaks away from the working environment.

4.3 Scheduling

Scheduling is a centralised operation that provides the field crews to meet daily unit hour requirements and therefore has responsibility for rota management, monitoring of sickness and leave and authorisation of pay. As such the scheduling function of production is accountable for unit hour consumption.

Demand analysis and historical patient volumes provide the basis for determining unit hours required. Daily cover plans estimating hourly UH requirements and budget available are agreed with distribution. Rotas from teams provide the bulk of cover available and these are matched to the plan. Any shortfalls in unit hours needed are covered using a variety of sources including pool and returned part time staff, pay back hours owed by staff and overtime. There are no relief shifts and scheduling decides which shifts to cover. If shortfall is small then available shifts, first responders or officers may provide the necessary cover.

Similarly officers can be used to cover expected peaks in demand for short periods which would not warrant a whole crewed shift. The process is dynamic in that it can be reviewed and amended in real time to take account of sickness or other unexpected absence. Demand analysis also provides a long-term view of staffing requirements so that, for example, maximum annual leave and training can be provided when demand is lowest.

The advantages for staff are that they can take extra leave in the summer months when demand is low, the system allows flexibility so that leave can be taken at short notice , for example when domestic problems arise, and the hours paid back later, and they can move into different job areas, for example first response. Use of officers and first responders also provides additional support when there is unexpected peaks in demand.
The disadvantages are that, particularly for pool staff, notice of when they are required can be short. There is also a perception amongst crews that the push for efficiency means there is, at times, inadequate numbers of vehicles available and that there is an over reliance on other types of response rather than fully crewed ambulances.

4.4 Summary

The fundamental principle of HPAS is the provision of best performance within current financial resources. Consequently the extraction of efficiencies across all parts of the organisation are critical to the achievement of high performance. The re-engineering of production processes has been responsible for many of the efficiency gains in SAS. The cost per A&E ambulance activated has been reduced from £140.24 in 1994/95 to £96.09 in 1998/99. The fleet has reduced from 72 to 53 vehicles. To maintain this efficiency fleet planning, servicing and supply acquisition and the make ready programme provides 85% availability of vehicles.
Within the organisation there has been a steady increase in unit hour utilisation despite annual increases in demand and each year better performance has been achieved with fewer unit hours as efficiency strategies have become more refined.
The strategies they have undertaken provide valuable lessons for other services. However, there are limitations to the efficiencies that can be gained from any organisation and the continued pursuit of further efficiencies causes tensions. The sophisticated strategies for matching supply and demand means that the service is always operating at the margin and there is less “slack” in the system which increases pressures on crews. The introduction of HPAS has had a substantial impact on the working conditions of road crews. Numbers have been reduced, staff are required to be much more flexible and they feel that they work harder as they now respond to a larger number of calls per shift than previously. The push for high unit hour utilisation means that if possible the service will try to operate using, for example, uniformed management staff to man ambulances or first respond, when the available hours are less than the planned ones. This makes the system efficient but causes its own pressures on crews and officers. The difficulty here is finding the balance between maximising efficiency and maintaining staff morale and satisfaction.

In general, crews recognise the need for changes and have responded positively to them. They appreciate the tangible benefits that have resulted in terms of improved response times, and the SAS's current position as the best performing ambulance service is a key motivator in sustaining the HPAS model. However, after almost five years of continued improvement some of the consequences of the changes are beginning to become apparent. During the process several key personnel who were instrumental in achieving these changes have left the service. For road crews, the most frequently stated areas of dissatisfaction during our interviews are:

- The continued pressure to respond to ever more calls within 8 minutes even though many of these calls are not life-threatening
- Insufficient vehicles and crews to cope with peaks in demand
- Poor standards and maintenance of response posts, old vehicles and continued attempts to cut costs on equipment
- Failure to translate promises of reward for improved performance into better pay

It is generally accepted within the organisation that crews have not been rewarded for improved performance. There is concern amongst some members of staff that unless these issues are addressed potential problems such as increased sickness and refusal to undertake overtime, more complaints and more staff leaving the service could emerge. However, it is also true that many of the comments about HPAS have been made in the light of comparisons between old and new methods of working. It is possible that over time, and as a bigger proportion of staff only experience the high performance environment, some of the perceived effects on staff may diminish. Also, after our interviews were conducted, some additional funding was made available to the service from the Coronary Heart Disease national service framework. This was used to make an additional payment to road crews. Changes in future funding arrangements as the NHS modernisation agenda progresses may provide the service with the means to resolve the issues around pay and performance.

5. Location and response analysis

5.1 Methods

We have examined changes in the SAS response time distribution from 1995-1999 and compared the latter distribution with those seen in other ambulance services.

In order to estimate the extent to which the current dynamic deployment system for vehicle location in SAS has improved response times, we have compared distances to incidents, calculated from the actual locations of vehicles when they responded, with simulated distances from the nearest fixed ambulance station which existed in 1995. On occasions there would have been no vehicle available at the nearest ambulance station to respond to the incident and instead a vehicle from further afield, with a longer response time, would have been used. We have simulated this situation by imagining that i) 10% or ii) 20% of all responses have been made from the second nearest station instead of from the nearest.

We have also estimated the response times for these actual and simulated distances by using the empirical relationships between straight-line distance and response time observed in other ambulance service data.

5.2 Sample

A random sample of 999 calls covering all times of the day was studied. Data was taken from the 6 month period April – September 1999. Eight 3 hour time periods were used; 0000 – 0259, 0300 – 0559, etc. For each time period two samples were chosen using a randomly identified month and a day of the month. All emergency calls where an ambulance or ambulance service first responder arrived on scene during each sampled time period were retrieved from the ambulance service computer aided dispatch system. All calls were included, rather than only category A calls, to reflect the practice in Staffordshire of responding immediately to every call, regardless of priority. Calls where no ambulance arrived on scene, for example stopped calls or duplicate calls, were excluded. A total of 271 calls were used in the final analysis. 14 cases were excluded as they contained missing information.

5.3 Data Collected

a) Location analysis

For each call the following data were extracted from the call log:

- Call time
- Clock start
- Time at scene
- Response time

- Crew Type (Fully manned ambulance or single responder)
- Address of the incident
- Address of the responding vehicle

Eight digit ordnance survey grid references were then calculated for:

- Location of the incident
- Location of the responding vehicle
- Location of the nearest historical station
- Location of the 2nd nearest historical station

b) Response time distributions

Data for the reported response time distributions were provided by Staffordshire Ambulance Service from their CAD system. All calls in the 3 month period April – June for the years 1995, 1997, 1998 and 1999 were included. For each call the following information was provided:

- Date
- Time clock started
- Time en-route to scene
- Time at scene
- Time depart scene
- Time at destination
- Time complete
- Priority (life-threatening, emergency or urgent)

Equivalent data for comparison with other services was already available from another MCRU project.

5.4 Results

5.4.1. Response time distributions

Shape of distributions

The distribution of response times in SAS for four years from 1995 – 1999 are shown in figures 1 – 4. The figures show three key characteristics:

1. The proportion of responses recorded as achieved within 0-3 minutes increased from 5% in 1995 to 10%, 11% and 12% in 1997, 98 and 99 respectively. This accords with the change in dispatch procedures introduced by SAS at around this time (see section 4.1). The procedures were designed to ensure that vehicles would be dispatched within a minute of a call coming in rather than the usual 2-3 minutes. Although this will improve the response times of all calls the effect would be most visible around 0-3 minutes.

2. The proportion of responses taking 19 minutes or longer was over 5% in 1995. Since then the proportion has varied from 1-2%. This suggests that either response to the more remote rural communities has improved or that the frequency with which incidents occur which cannot be covered by locally available units has decreased. The latter case arises if more incidents occur in quick succession in high demand areas than there are units available to respond. Again, two changes which have been introduced – a number of first responder schemes, and the implementation of the system status plan which moves vehicles to cover areas when local units are deployed - might have been expected to have just these effects.

3. The third distinct feature of the graphs is the curious discontinuity, which was at 9 minutes in 1997, and then when the new performance standards were introduced clarifying that 75% of responses were to be achieved within 8 minutes, shifted to 8 minutes. It is unlikely, but not impossible that the discontinuity could arise naturally. It implies either that

a) responding units can be stationed in such a way that response times are 7 – 8 minutes rather than 8 – 9 minutes, or

b) that by some means responses which would otherwise take 8 – 9 minutes can be speeded up so that they are achieved in 7 – 8 minutes.

The first possibility, resulting from the location of vehicles, could only be the explanation if the distance to incidents also showed the same discontinuity. This could be the case, for example, in an urban area with no suburbs so that the population, and hence the distribution of incidents, was either within the urban area and within 8 minutes distance, or in a large rural surrounding area 9 or more minutes away. However, this is unlikely because

i) it is not a pattern seen in other ambulance services we have studied (see figures 5 - 7).and

ii) the distribution of straight-line distances from responding vehicle locations to incidents in SAS data does not show the same discontinuity (see figure 8).

The second possibility could be the explanation if the responding crew (or the control room) could foresee that the response time was likely to be over 8 minutes, and then the responding vehicle speeded up or an alternative response unit (e.g. a first responder) was called in. However, if this is the case it is important to ask why a response is not <u>always</u> dispatched in the minimum possible time. If it is, then this explanation cannot be the case.

If the recorded response times are genuine we might expect the <u>total time</u> from call to hospital to show the same discontinuity, but this is not apparent (Fig 9). In fact, examining the component times (response, on-scene, travel time to hospital) we can see

that as response times decreased dramatically between 1995 and 1997 (mean response 17 mins to 8 mins; median 9 mins to 7 mins) so the 'on-scene time' increased (mean 13 mins to 17 mins, median 11 to 15) (figs 10 and 11). The impression is that some part of the response time reduction is artificial and hence is simply added to the scene time, but that some part is genuine.

It seems likely therefore that part of the discontinuity around 7-9 minutes is artificial rather than natural. One explanation for this artificiality comes from the fact that SAS control staff selectively edit the response time data after it is initially recorded. Since crews sometimes fail to notify the control room when they arrive on scene, the response time initially recorded is sometimes wrong by over recording the true response time. In SAS all response times just over the 8 minute threshold are reviewed to see whether the crews were in fact already on scene at the time arrival at scene was recorded. Some of these response times are thus edited downwards. Because performance targets are set in terms of the 8 minute standard (rather than the mean time) SAS do not of course need to review times initially recorded as under 8 minutes even though these may also be over-recorded.

Similarly, in real time, dispatchers are alerted by the CAD system as 8 minutes approaches and no arrival time has been logged. This prompts the dispatcher to contact the crew to see if they have in fact arrived at the scene. If they have, this time (between 7-8minutes) will be recorded, even if they actually arrived a few minutes earlier. Thus, the combined effects of this and selective editing creates an artificial discontinuity at 8 minutes. It is important to appreciate that this itself does not necessarily mean that the reported proportion of responses completed within 8 minutes is wrong. Thus it seems likely that rapid dispatch, first responders and System Status Management together with selective editing of wrongly recorded response times have led to the improvement in the response time distributions shown in figures 1-4.

The only question that remains is whether there is also any under recording of response times which is also contributing to the very high level of performance recorded by SAS. This can be explored by comparing the recorded (edited) response times with the actual journey distances.

5.4.2 Distances

The mean, median, 75th and 90th centile straight-line distances from the incidents to the actual vehicle location, the nearest (old) ambulance station, and the second nearest station are shown in table 1. Surprisingly, there was no difference between the mean distance from the incidents to the actual vehicle location and the nearest station. However, dynamic deployment did result in a shorter median and 75th centile distance, although the 90th centile

recorded distance was longer than the distance to the nearest station. The distributions of the actual and simulated distances to the nearest station are shown in Figures 12 and 13. It can be seen that the actual distances are generally shorter than the distances to the stations, but there are also longer distances. This presumably occurs in the real situations represented by the actual data in Figure 12 because sometimes no nearby vehicle is available to be deployed, and one has to be brought in from further afield.

Table 1. Straight-line distance (kms) from incidents to actual and theoretical vehicle locations

Distance to	Mean	Median	75th centile	90th centile
Actual vehicle location	3.5	2.4	4.2	8.6
Nearest station	3.6	3.1	4.9	7.1
Second nearest station	10.6	10.0	13.0	14.8

We have attempted to simulate this situation by imagining that in either 10% or 20% of these incidents no vehicle would have been available at the nearest station and instead one would have been deployed from the second nearest station. The effect of this is shown in Table 2 and Figs 14 and 15.

Table 2. Straight-line distance (kms) from incidents to actual and possible vehicle locations

Distance to	Mean	Median	75th centile	90th centile
Actual vehicle location	3.5	2.4	4.2	8.6
Nearest station + 10% from second nearest station	4.3	3.5	5.6	8.5
Nearest station + 20% from second nearest station	5.0	3.9	7.0	10.9

We do not know which of these models most closely represents what would occur without dynamic deployment. But the models clearly show that most distances are shorter using dynamic deployment. The estimated reduction from using dynamic deployment in the average response distance is about 1.0 - 1.5 kms.

5.4.3 Speeds

The relationship between the actual straight-line distance between the incident and vehicle locations and the recorded response times is shown in Figure 16. The figure also shows the estimated relationship between response time and distance calculated by the straight-line regression of time on distance. The regression gave the relationship

time (minutes) = 3.7 + 0.7 x straight line distance (kms).

The regression coefficient of 0.7 implies a travel speed of 86 straight-line kms/hr. This seems unrealistically fast since it is equivalent to a speed in excess of 70mph (road miles/hr).

We have compared this relationship with that from the automatically recorded incident and vehicle location and recorded response times in another Ambulance Service. The relationship in their data is shown in Figure 17. The regression of the recorded response time on the straight-line distance in this comparator data gave the relationship

time (minutes) = 2.8 + 1.2 x straight line distance (kms)

implying a more realistic travel speed of approximately 45 mph. This comparison suggests that some response times are being under-recorded in SAS leading to the impression of unrealistically fast travel speeds.

In fact if Figure 16 is re-examined it will be seen that of 29 response times recorded by first responders rather than conventionally crewed emergency ambulances, only two exceeded 8 minutes. The response time distributions of normal crews and first responders are shown separately in Figures 18 and 19 and the 8 minute 'barrier' for first responders is plain.

Excluding the first responders therefore, the estimated relationship between response times and travel distances in SAS was

Time (minutes) = 3.5 + 0.8 x straight-line distance (kms).

This still implies a travel speed of 79 straight-line kph with a lower 95% confidence limit of 70 straight-line kph, a speed in excess of 60 road mph.

Thus it seems that in a small minority of journeys emergency ambulance response times are being under-recorded to the extent that the implied speeds are unrealistic.

5.5 Summary and conclusions of the analysis

1. Analysis of the component times suggests that some of the reduction in mean response times was just added into the scene times, but some of the improvement was real.

2. Real improvements have been achieved by rapid dispatch and by first responder schemes.

3. Selective editing of wrongly recorded response times has led to some artificiality in the response time distribution but this does not compromise the accuracy of the reported 8 minute performance level.

4. The analysis of distances suggests that SAS has also achieved a real reduction in response distances compared to dispatch from fixed ambulance stations.

5. Nevertheless, the analysis of SAS distance and response data indicates that there is under recording of some response times implying impossibly fast speeds

Figures 1-4: **Recorded response times in Staffordshire 1995-1999**

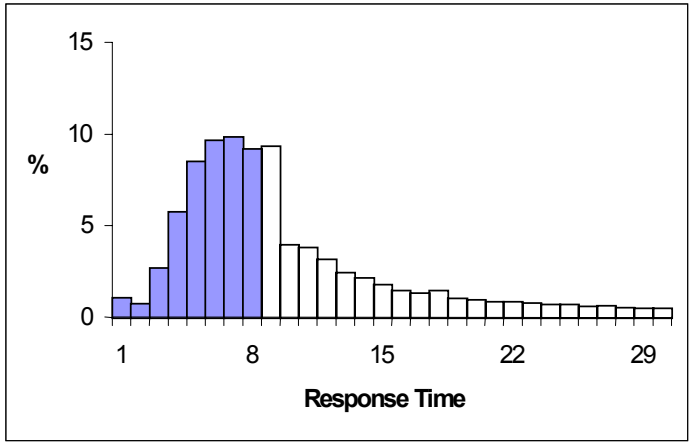


Figure 1: April - June 1995

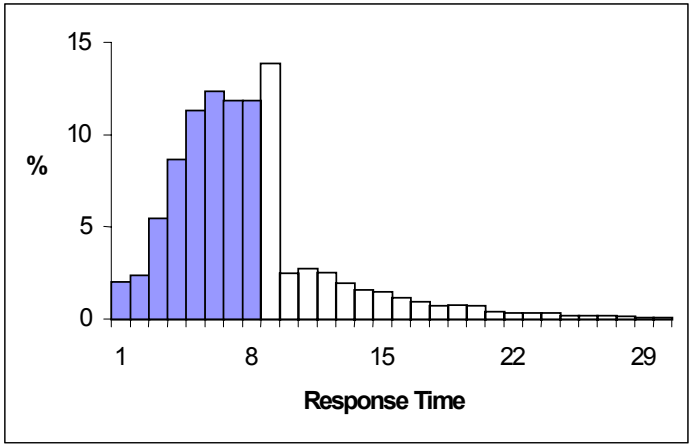


Figure 2: April - June 1997

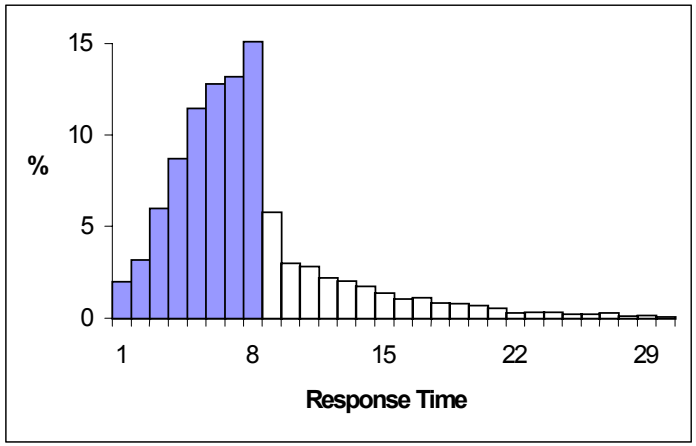


Figure 3: April - June 1998

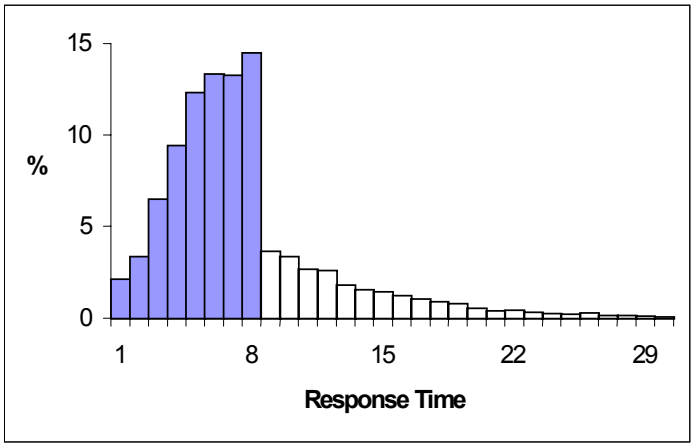


Figure 4: April - June 1999

Figure 5: Recorded response times, Staffordshire: April - June 1999

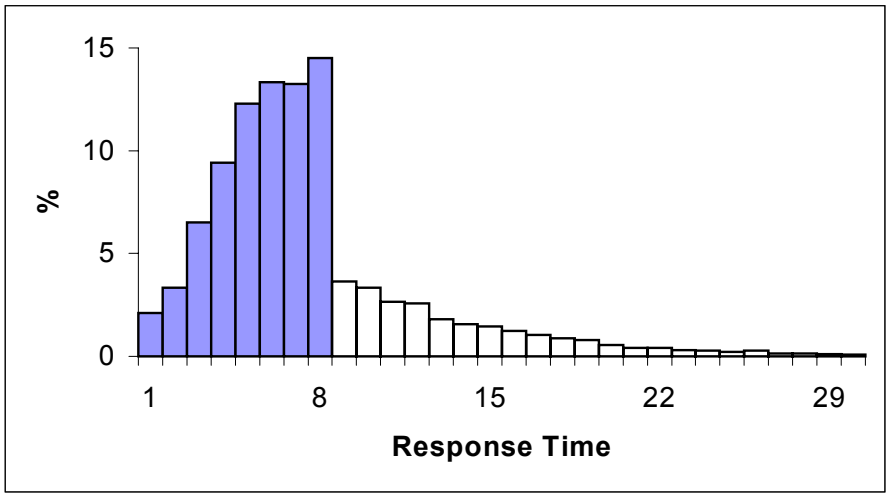


Figure 6: Recorded response times, Rural comparator service: November 1997

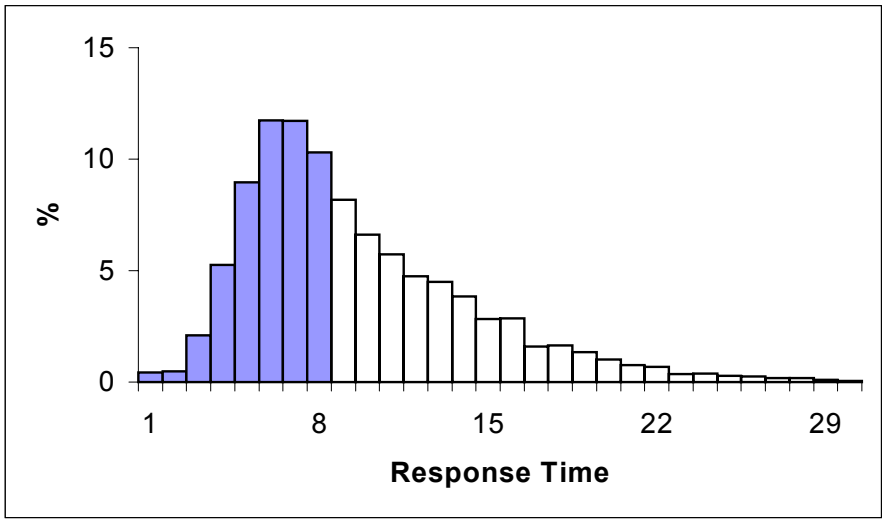
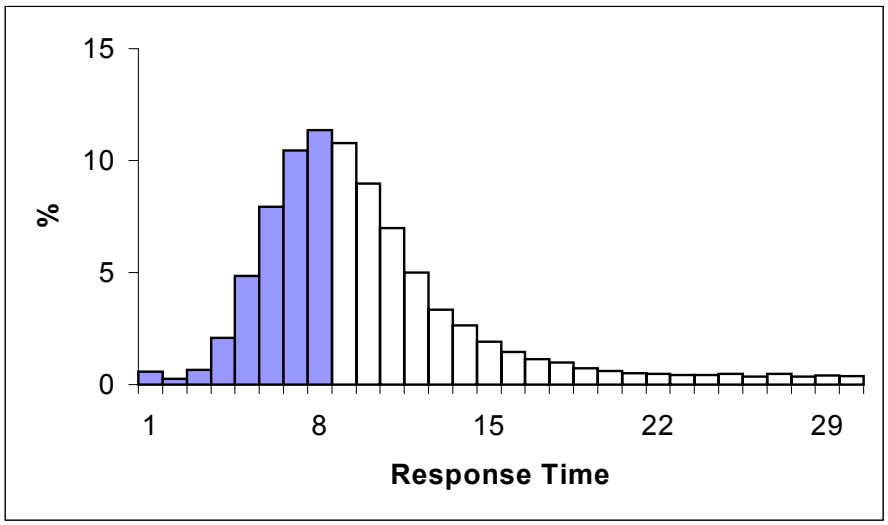


Figure 7: Recorded response times, Urban comparator service: April 1998



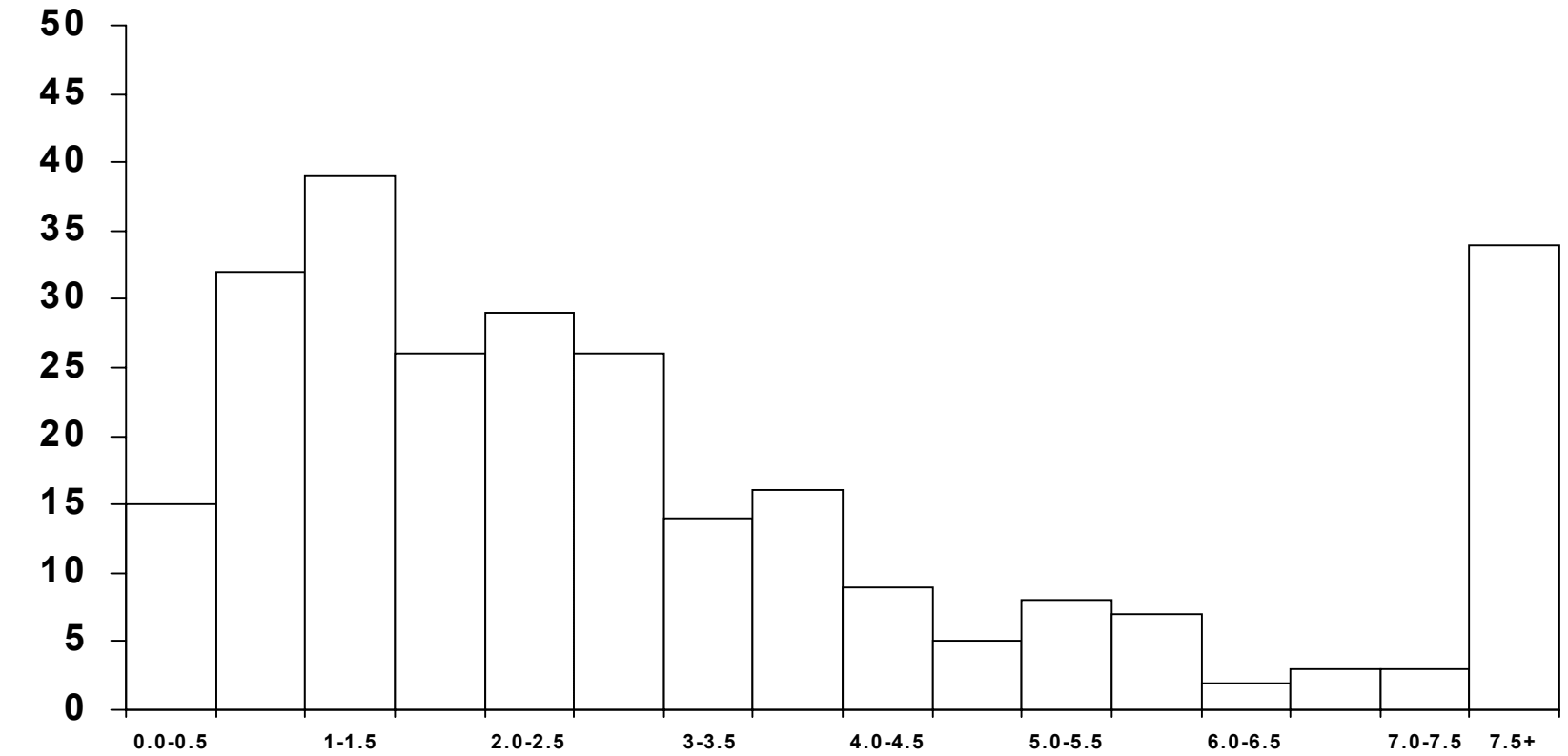


Figure 8 Straight line distance from incident to actual responding vehicle location in Staffordshire

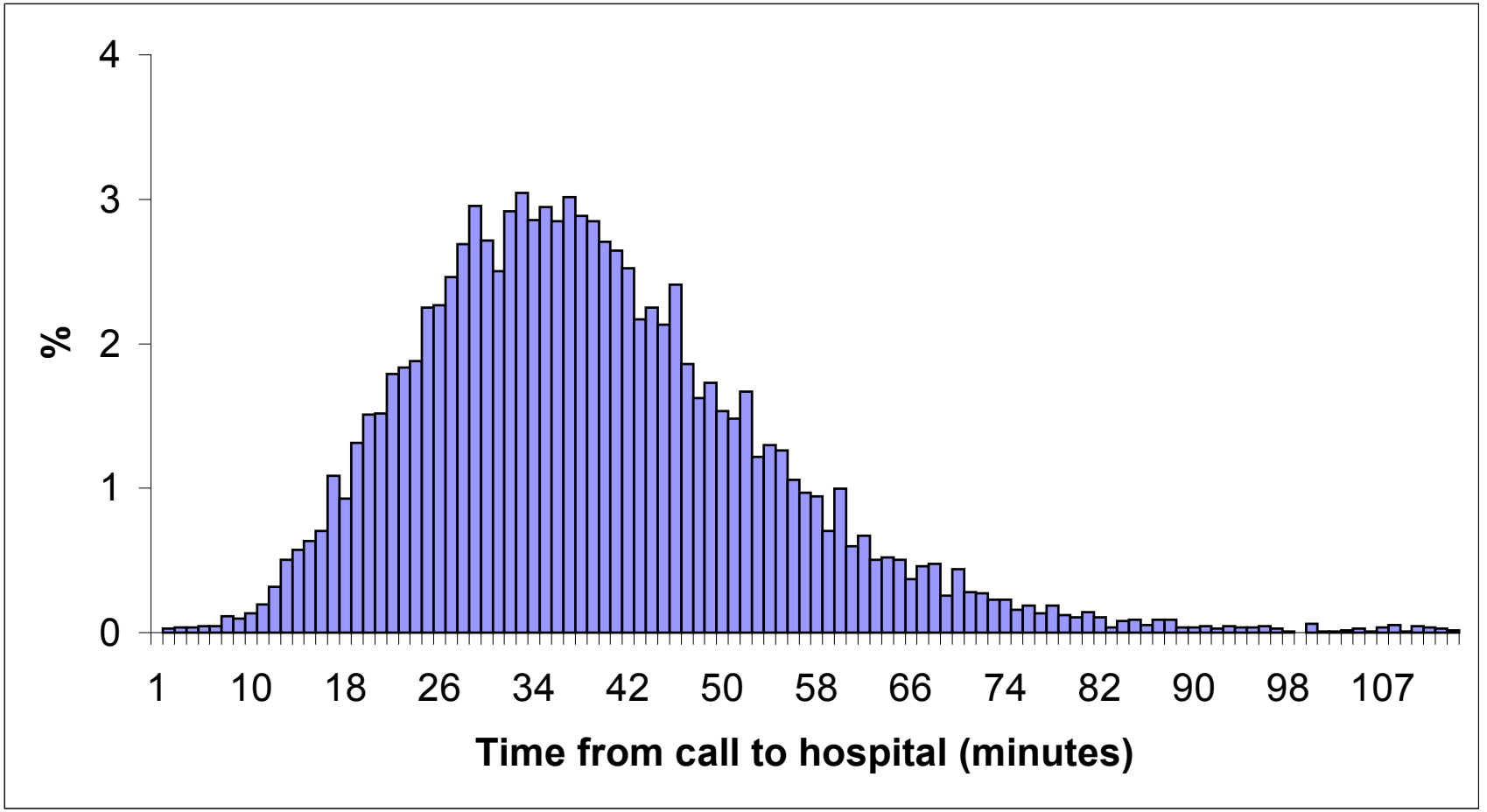


Figure 9: Total time to hospital, Staffordshire

Figure 10: Median component times - Staffordshire

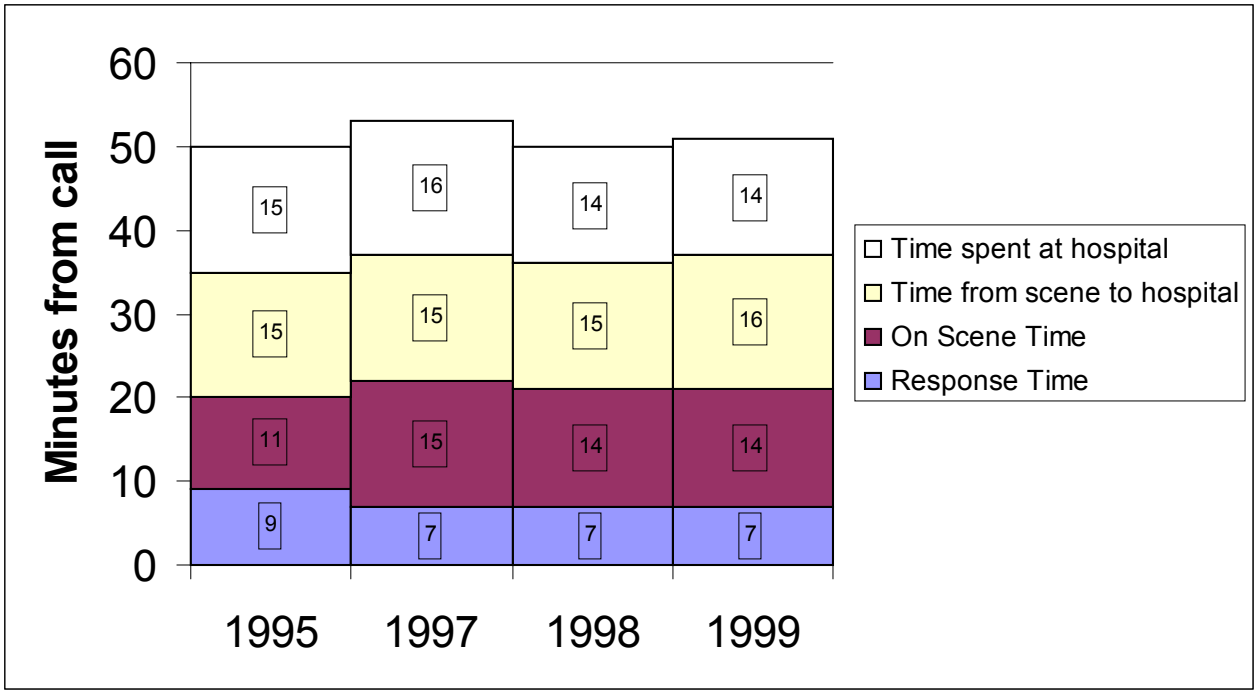


Figure 11: Mean component times - Staffordshire

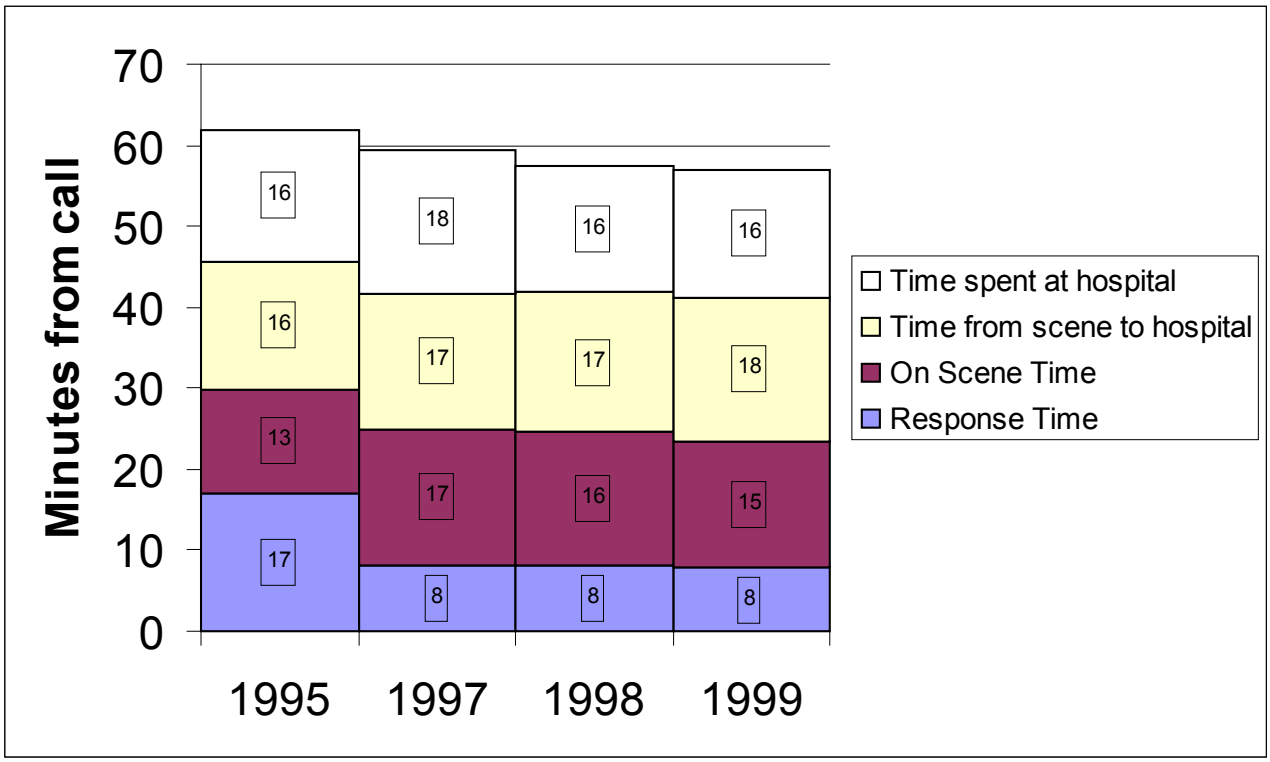


Fig 12. Straight line distance from incident to actual responding vehicle location in Staffordshire

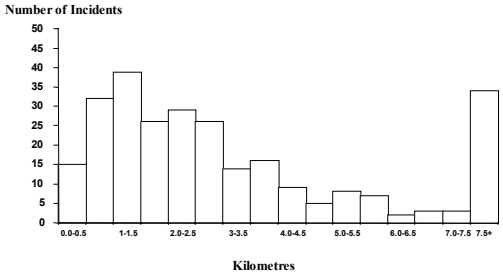


Fig 13. Straight -line distance from incident to nearest ambulance station

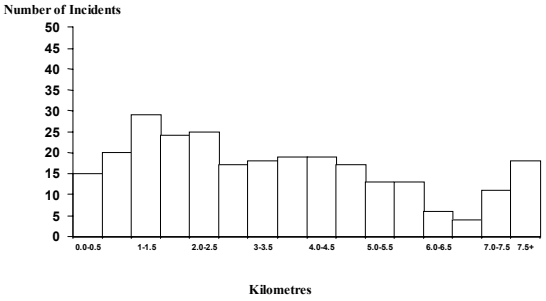


Fig 14. Simulated straight line distance with 90% of journeys from nearest station and 10% from the second nearest

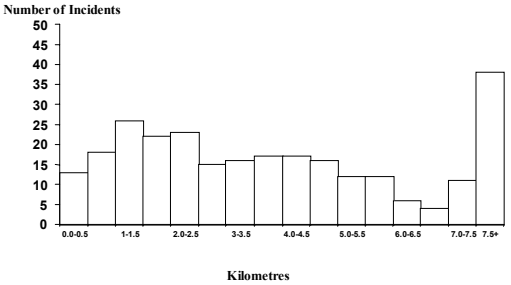


Fig 15. Simulated straight line distancel with 80% of journeys from nearest station and 20% from the second nearest

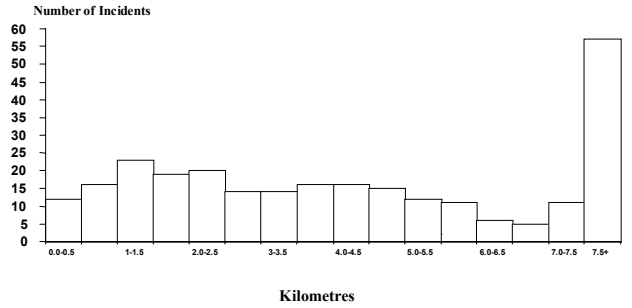


Figure 16 - Relationship between recorded response time and calculated straight line distance of responder

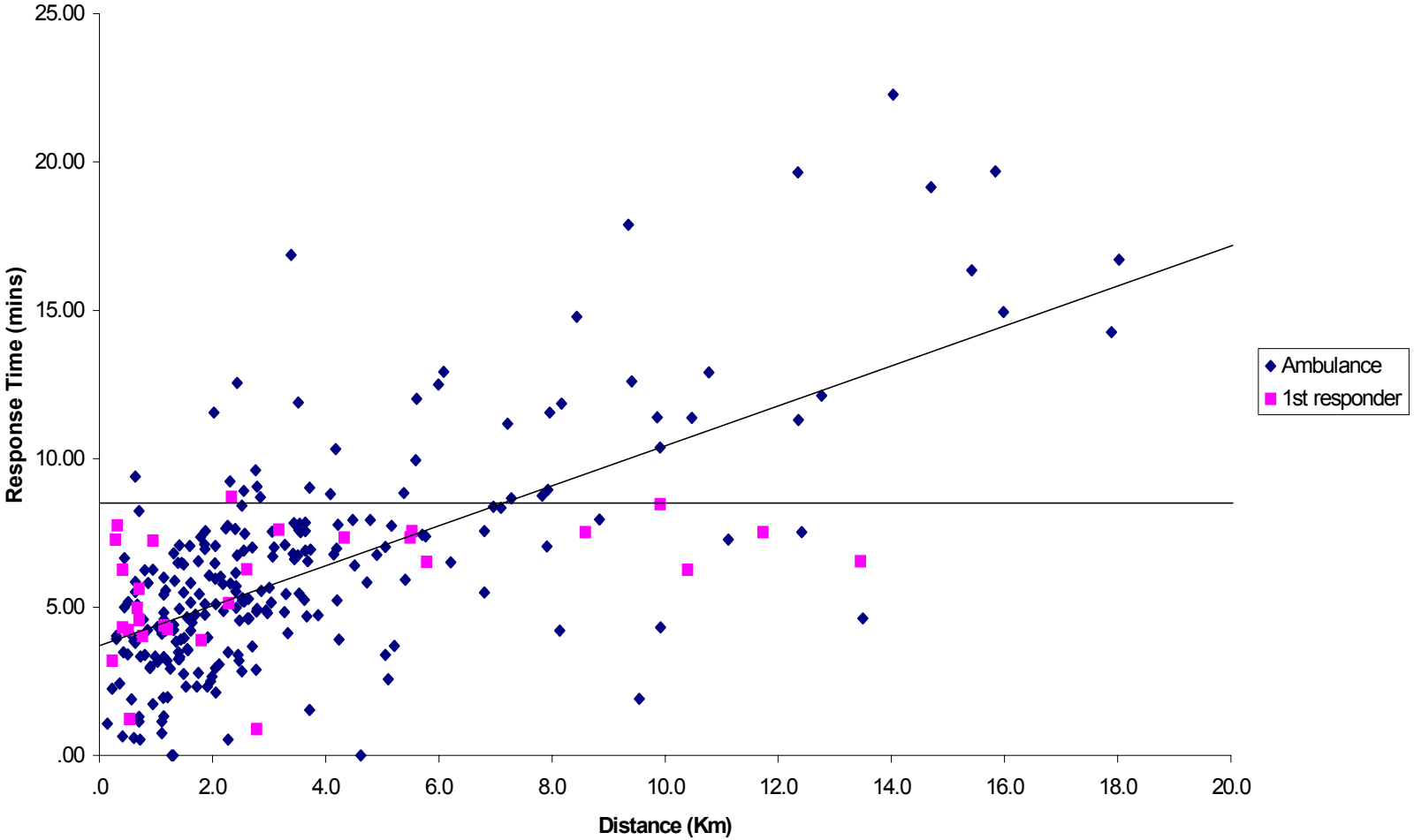
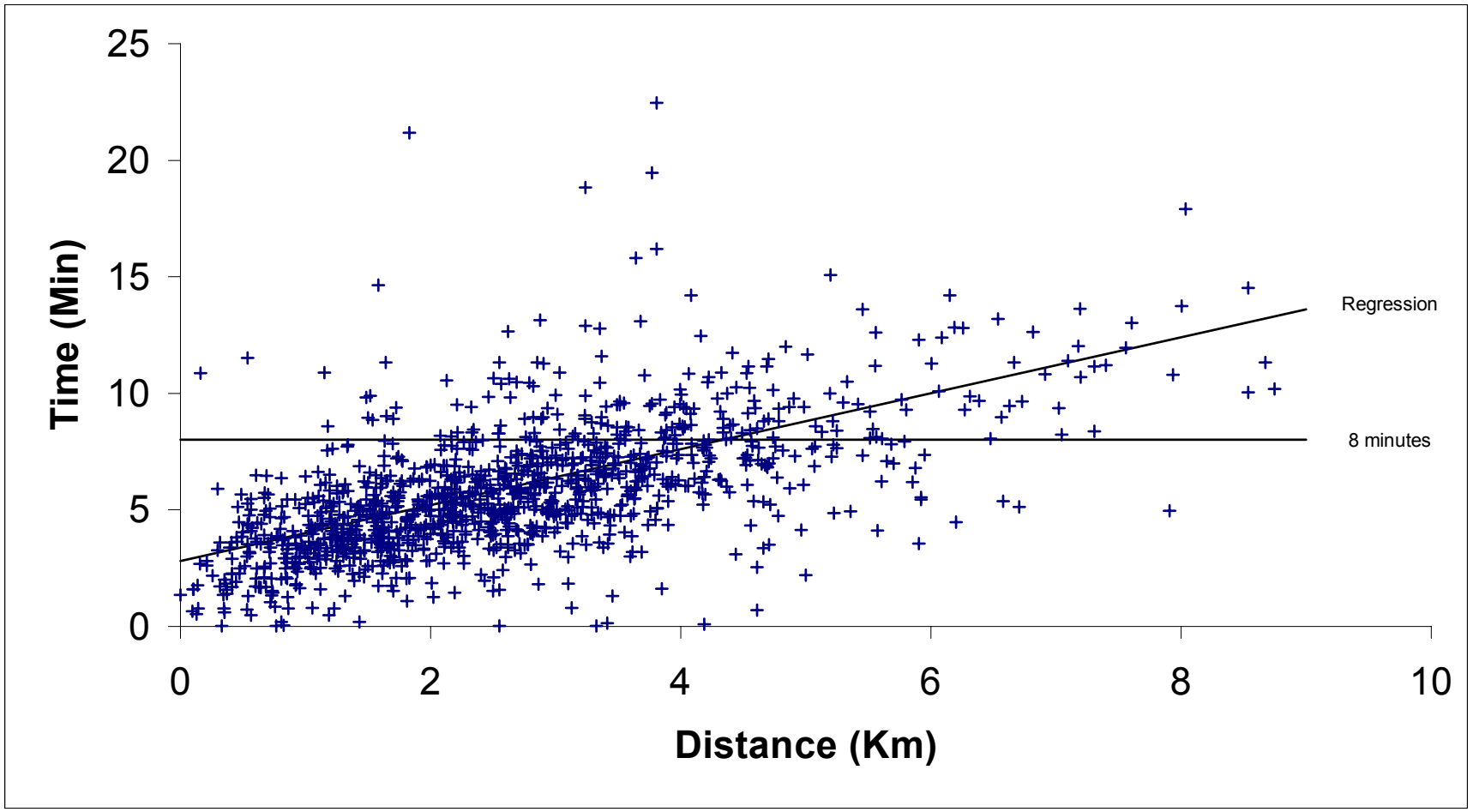


Figure 17: Relationship between response time and straight line distance from incident to recorded vehicle location in a comparator service



6. Audit of out of hospital cardiac arrest

6.1 Introduction

The rationale for faster response time performance is that early intervention in time critical events will lead to better patient outcomes. For some years Staffordshire Ambulance Service have been conducting an annual audit of cardiac arrests attended as a measure of clinical performance both in terms of process (protocol compliance) and outcome. The principal patient outcome measure used is Return of Spontaneous Circulation (ROSC) during the pre-hospital phase of care and Staffordshire have reported an improvement in ROSC from 9.4% in 1996/7 to 11.9% in 1999/2000. ROSC does not, of course, necessarily result in an increase in survival[13] but in the absence of any information which allows the Trust to measure outcome following hospital care ROSC does provide a proxy measure which allows year on year changes to be detected.

We have conducted our own audit of out of hospital cardiac arrests attended by SAS crews in order to validate the figures reported and provide an example of the potential improvements in clinical care that can result as a consequence of introducing HPAS principles.

6.2 Methods

A random sample of cardiac arrests attended by SAS crews during 1999 was obtained and outcomes in terms of ROSC on arrival at hospital and survival to discharge from hospital were measured.

6.2.1 Sample

There were two sampling options available:

i) Identify cases from the CAD system with problem codes recorded as cardiac/respiratory arrest. The advantage of this method is that a sample independent of that analysed by the Trust audit is obtained. The disadvantage is that it assumes all cases of cardiac arrest can be identified from this single problem code. The audit carried out by SAS has shown that about 30% of cardiac arrest cases have problem codes other than cardiac/respiratory arrest. This is because insufficient or incorrect information is given to the call taker at the time the call, or because the arrest takes place after the call.

ii) Take a random sample of cases from those already identified by SAS for their own audit. The advantage of this strategy is that cases with problem codes other than cardiac or respiratory arrest will be included in the sample. The disadvantage is that this cohort may already have been subject to some degree of selection because the cases are identified by paper records rather than by the CAD system.

We have therefore taken two samples for our analysis. The first was derived from the CAD

system using the problem code of cardiac/respiratory arrest. All cases with this code for alternate months during 1999 (starting with January) were identified. In each of these 6 months the CAD details of every 3rd case were retrieved. Where there was more than one ambulance service response to an incident details of each response were obtained. One hundred and eighty four cases were identified. For each of these cases a patient report form (PRF) and /or SAS cardiac arrest form was retrieved. SAS has a separate paper form for cardiac arrest cases which the attending crew complete. Fifteen cases were not related to a cardiac arrest, 9 cases had missing PRF's/cardiac arrest forms and 3 cases were excluded as insufficient information was available. This provided a sample of 157 cases for analysis.

The second sample was derived from the cohort identified by SAS for their own audit. The paper records (CAD details, PRF and or/cardiac arrest form) for all cases are kept for each year by the clinical audit department. Cases are filed in date order. For each of the same 6 months used in the first sample, every 3rd case was retrieved. This provided a sample of 189 cases of which 3 were duplicates identified for the first sample giving 186 cases for analysis.

In keeping with the SAS audit, all cases where there was no detectable cardiac output were included. This includes cases where the patient was obviously dead on arrival of the crew and also includes cardiac arrest from all causes (e.g. trauma, poisoning, asphyxia etc.), not just those of known cardiac origin.

6.2.2 Information collected

For every case in each sample the following information was recorded:

i) CAD system

- Patient details (name, age address for subsequent follow up to hospital)
- Incident date
- Problem code
- All relevant timings (call, mobile, at scene, left scene, at hospital)

ii) PRF/cardiac arrest form

- Presenting rhythm on arrival of crew (or at time of arrest where this was crew witnessed)
- Estimated time of collapse
- Whether the arrest was witnessed and who by (bystander, relative, crew)
- Whether resuscitation was attempted
- Disposal (left at scene or transported to hospital)
- Presence or absence of any vital signs (respiration, pulse) on arrival at hospital

iii) At hospital outcome classified by

- Died in A&E
- Admitted to hospital
- Died in hospital

- Discharged alive

6.3 Results

As the sample has been derived by different methods and it is known the casemix within each sample is likely to be different the cases were divided into three groups for the initial analysis. These were:

Sample 1 - All cases derived from the CAD system using the problem code cardiac/respiratory arrest

Sample 2 - Cases derived form the SAS audit cohort with the problem code cardiac/respiratory arrest

Sample 3 - Cases derived form the SAS audit cohort with problem codes other than cardiac/respiratory arrest

A summary of all cases for each sample by presenting rhythm and whether they were certified as dead at the scene (DAS) or were transferred to hospital is given in Table 5. Presenting rhythm is classified as asystole; Electromechanical Dissociation (EMD) or Pulseless Electrical Activity (PEA); or Ventricular Fibrillation (VF).

Table 3 - Summary table of all cardiac arrest cases by presenting rhythm and initial disposal.

	Sample 1 (n=157)	**Sample 2 (n=82)**	**Sample 3 (n=104)**
Dead at scene			
All cases – number (%)	96 (61%)	57 (69.5%)	56 (53.8%)
No resuscitation - (all asystole)	60 (38.2%)	41 (50%)	29 (27.8%)
Resuscitation attempted all cases	34 (21.6%)	16 (19.5%)	25 (24%)
asystole/EMD	28 (17.8%)	12 (14.6%)	22 (21.1%)
VF	6 (3.8%)	3 (3.6%)	3 (2.9%)
Transferred to hospital			
all cases	62 (39%)	24 (29.3%)	49(47.1%)
asystole	22 (14%)	9 (11%)	10 (9.6%)
EMD/PEA	6 (3.8%)	5 (6.1%)	22 (21.1%)
VF	33 (21%)	10 (12.2%)	17 (16.3%)
unknown	1(0.6%)	1 (1.2%)	

Not surprisingly, there are some clear differences in the characteristics of each sample. In particular, a much larger proportion of patients not identified at the time of the call as an arrest (sample 3) had resuscitation attempted and were transferred to hospital than in the 2 groups where the problem category of cardiac arrest was recorded. This may reflect the greater likelihood in sample 3 that the arrest is more recent and therefore more amenable to treatment. In contrast, a high proportion of patients identified as cardiac or respiratory arrest at the time of the call were clearly deceased at the scene and no resuscitation was attempted.

The ROSC rate and survival to discharge form hospital rates for both samples is given in table 4. Because of variations between the two samples, particularly the inclusion of cardiac arrests not identified at the time of the call we have calculated an adjusted ROSC rate for the combined samples. This is based on the proportion of cases achieving ROSC in each sample and the proportion of cases identified as AMPDS cardiac arrest code or other code. Twenty seven cases (10 in sample 1 and 17 in sample 2) could not be traced at hospital and we were unable to obtain outcome data at one hospital (5 cases). The survival to discharge rate has been calculated from all cases where outcome is known.

Table 4 – ROSC and survival rates for all cases of cardiac arrest

	Sample 1	Sample 2 + 3	Combined adjusted sample
ROSC			
All cases	157	186	343
Number of ROSC	18	22	40
ROSC rate	11.5%	11.8%	11.7%
ROSC rate of patients transported to hospital	29%	30.1%	29.6%
Survival to discharge rate			
All cases traced	145	166	311
Number of survivors	3	3	6
Survival to discharge rate	2.1%	1.8%	1.9%
Survival to discharge rate of patients admitted to A&E	7.5%	6.6%	7.0%

The ROSC rate is comparable with those reported in two other UK studies[14,15]. However, in one study all cases of non-cardiac origin were excluded[14] and in the other the selection process for cases was not described[15]. As all cases, regardless of aetiology, were included in this study the ROSC rate would be higher if cases of non-cardiac origin were excluded. Similarly the survival to discharge rate is similar to other studies[15,16] for patients admitted to

an A&E department (i.e. excluding all cases certified dead on arrival). The overall survival to discharge rate of 2% is more difficult to compare, as the majority of studies of out of hospital cardiac arrest do not include patients not transported to hospital. However this rate is comparable to one other UK study which included all cases of cardiac arrest[17].

6.4 Summary

The ROSC rate is very similar in both samples and, when adjusted for the differences between the two samples with regard to whether or not the call was identified as a cardiac arrest at the time of the call, gives an overall ROSC rate of 11.7%. This reflects the high proportion of cases identified with non-cardiac arrest codes in sample 2. In this analysis 66% of cases had other AMPDS codes. This is much higher than the proportion reported in the SAS cardiac arrest audit report where 38% of arrests had non cardiac arrest codes. If the analysis is adjusted using this proportion of non-arrest codes the ROSC rate remains as 11.7%. The ROSC rate found in this analysis concurs with those reported by Staffordshire and confirms that the currently reported rate is accurate. Survival to discharge from hospital rates are comparable to those reported in other UK studies.

7. Transferable lessons, reproducibility and sustainability

7.1 - Summary of current position at Staffordshire Ambulance Service

After more than 5 years of change it would appear that Staffordshire has probably now reached a plateau and that, for current levels of activity, maximum production and distribution efficiencies have been achieved. The CE believes that, as a traditional ambulance service SAS has probably got as far as it can go and further improvements can only be made by exploring alternative processes of patient management, for example by extending the community paramedic initiative, so that fewer patients require transport to hospital.

The ability to produce further improvements in Staffordshire is limited by both the organisation and factors present in the wider NHS. The organisational limitations are:

- The “easy” and most of the difficult improvements have been accomplished. Line and control staff are struggling with increased demand. Senior management believes more efficiencies can possibly be extracted but line staff do not, which is creating an increasing organisational gap.
- Tension exists between the desire to maintain and improve clinical performance and the production goals of maintaining efficiency. Economic constraints are forcing the Trust to delay clinical development and progress.
- Staff compliance with SSM and the vision of Trust: - Crews appear increasingly under pressure and are subjected to continuous demands to maintain performance. The constant pressure is a source of conflict. Some, particularly the CE, perceive this as a necessity since if pressure is not maintained performance will fall off. Others believe that staff have taken on and accepted the cultural change, that this sustained pressure is unreasonable and that to persist will produce serious problems.
- The lack of ability to reward staff for improved performance exacerbates this problem. This is not confined to pay although this is a major factor and the ability to properly financially reward staff would undoubtedly improve morale and motivation.

Future developments and further service improvements are also constrained by limitations posed by the organisation of the NHS. The most pertinent factors identified by SAS are:

- Contracts should be based on quality standards, not just lowest cost and should include all services within a defined area. There needs to be a much better understanding of what funding is required to meet demand.
- Efficiency gains have been made and it would be helpful to retain savings as pump priming for new developments.

- The year on year increases in demand pose a particular problem. Because the Trust is achieving performance targets they feel there is little incentive for purchasers to provide funds to meet extra demand.

7.2 Transferability

The emerging lessons from Staffordshire with respect to changing to a high performance system are:

1. Fast activation, rigorous demand analysis, development and refinement of system status plans to match supply and demand and use of dynamic deployment and first responders have produced the biggest impact in terms of improved response time performance.
2. Implementation of SSM results in changes to working conditions. For road crews this means more (or often most) of their time is spent in vehicles responding to calls or moving around between response posts. Many of the previous 'natural' breaks (e.g. returning to stations) have disappeared. For control staff, dispatching carries much greater responsibility and involves not only deploying vehicles but also maintaining the system status plan. Lessons from Staffordshire have demonstrated that this is one of the most difficult areas of change and one which requires significant forward planning. The process could be achieved more easily if response posts are put in place before stations are closed. This strategy also helps in gaining public acceptance of changes.
3. Adoption of the high performance approach requires significant change in organisational culture with an emphasis on the needs of patients rather than the organisation. This change can only be achieved if the whole organisation 'buys in' to the idea. This is particularly true of the road crews and control staff as, ultimately, it is these people who deliver the service at the front end and will make the system succeed or fail. The achievements in Staffordshire have only been made because these key groups of staff have accepted the principles of HPAS and adapted to change. There is real risk attached to the HPAS strategy and the organisation, including the board and senior management, has to accept the possibility of failure.
4. The process of implementing these changes is lengthy and difficult. Significant forward and long term planning and strong leadership is required for any change to be successful. This can produce a financial clash as considerable investment needs to made in some areas before savings are released. Staffordshire have achieved improvements by incremental change over a long period of time. However they feel that it would have been helpful to have pump priming funds to facilitate some of their developments. Any attempt to try to implement this level of change over a much shorter time period would require an appropriate injection of resources at the outset.
5. The use of first responders has also produced measurable benefits. Some of these are paramedics who are employed for this purpose. However, much use is also made of

other staff who have paramedic, technician or basic life support and AED training which has contributed to operational efficiency. The use of operational staff to respond to calls is not questioned, rather it is the extent of this practice which requires consideration. A balance is needed between the needs of the organisation and response time performance and the need for staff to be able to perform their roles properly and efficiently. The use of first responders could have an important role to play, particularly in rural areas where there is low demand and distances are long and efficiency gains are more difficult to realise. Transferring high performance system working to rural areas will require additional planning and possibly greater emphasis on developing robust and properly funded first responder networks.

6. Investment in information technology is critical to success both in terms of the transition to HPAS working and the continuous assessment and refining of plans and drive for efficiency. Equally important is investment in the training and development of people to use information and develop and implement HPAS methods. Staffordshire sent staff on a number of visits to the USA to learn the principles of SSM and HPAS. At present there is probably not sufficient detailed knowledge and expertise of high performance systems in the UK to adopt this approach nationally in the short term. If there were widespread interest in adopting the HPAS approach then some investment in the development of appropriate UK training courses would be required.

7. The introduction of SSM and other efficiency measures has resulted in improved response performance in Staffordshire with no additional funding. That there remains scope within other services to use resources more efficiently and effectively is not disputed. However, the magnitude of achievable improvements in performance may vary depending on the current level of funding. The Audit Commission report[8] demonstrated that the costs of ambulance services vary enormously. In their analysis Staffordshire were considered to be averagely funded. So, whilst it may be possible that some services could make the same improvements, or better, than Staffordshire it is also possible that for others the capacity will be limited. The level of improvement will be dependent on the baseline funding.

8. It is acknowledged that one of the reasons for the successful implementation of a high performance system in Staffordshire is the presence of a relatively stable workforce. There are serious questions surrounding the transferability of the model to areas where there is a more mobile or less compliant workforce.

9. Improving performance leads to greater public expectation, particularly when the changes are given a high profile in, for example, the local press. Early and sustained involvement of the public is essential if changes are to be accepted.

A summary of areas of good practice identified in Staffordshire, the relative ease of implementation and the organisational implications is given in table 9

Table 5 - Initiatives for improving response times

Initiative	Difficulty of implementation	Organisational implications
Fast Activation		
1. Call takers and dispatchers working adjacently	Easy to Medium	May require control room re-organisation if call taking & dispatch are separated.
2. Pre-assigning vehicle on receipt of postcode	Easy	Dependant on communications system. May require some updating of CAD system. AVL essential.
3. Rapid transfer of information to crews	Easy to Medium	Requires reliable communications system – radio, paging system, mobile phones
4. Crews in vehicles rather than stations	Medium to Difficult	Culture change for crews. Identification of suitable standby points. Other issues of dynamic deployment (see below)
System Status Management		
1. Detailed demand analysis	Easy to medium	Depends on quality of information systems available and ability of staff to use them. Where IT system is old investment required as information is a key issue.
2. Development of system status plans	Difficult	Labour intensive. Time consuming .Requires good historical temporal and geographical demand data.
3. Development of scheduling systems to		

match supply and demand	Medium to difficult	Follows from step 2. Dependant on good demand analysis. All 3 steps require significant organisational change – adoption of UHU as currency, development of distribution and production functions. Significant staff information and training implications
Implementation of SSM – direct measures		
1. Control room – dispatcher function	Medium	Training implications for dispatchers to implement of SSM plan. Manpower and financial considerations if control room roles transferred to paramedic staff.
2 Field crews – dynamic deployment	Difficult	Substantial change in working conditions. Some vehicle adaptations. Development of team approach and central reporting. Cultural change towards patient focussed service.
3. Other operational staff	Medium to difficult	Cultural change and acceptance of requirement for all operational staff (including managers) to respond to calls. Development of voluntary out of hours first response function.

Implementation of SSM – indirect measures		Long term planning required on management of estates – station closures, identification and leasing or purchase of suitable response posts. Implement public information programme to gain public support.
1. Posting points	Difficult	
2. Make ready function	Medium to difficult	Re-organisation of maintenance functions and development of central reporting and make ready facilities. Streamlining of equipment and supplies purchase. Strategies for delivery and collection of vehicles.
First responders		
1. Community paramedic first responders	Easy	Good communications system (paging. May require some investment in vehicles and equipment.
2.Other operational staff	Easy to medium	Cultural change required to replace traditional manager/officer role. Vehicles and equipment.
3. PTS crews	Medium	Training and equipment investment. Possible review of PTS contracts. Good development opportunity for staff.
4. Community lay first responders	Medium to difficult	Recruitment and information programme. Training, monitoring & supervision requirements. Equipment.

7.3 Reproducibility

It is possible other Trusts, despite many examples of innovative practice, will not take on the level of change necessary to reproduce Staffordshire performance. Increasing numbers of Staffordshire managers are moving on to other Trusts but lack of belief at senior management level may constrain full utilisation of ideas and experience gained in Staffordshire. Further, merely recruiting middle managers from Staffordshire is unlikely to produce good results elsewhere unless Trust boards are committed to HPAS principles and the consequent operational re-organisation takes place. A key lesson from Staffordshire is that to implement the HPAS approach requires commitment across the whole organisation to the changes this entails.

7.4 Sustainability

A number of issues have been identified relating to the likely long term success of the Staffordshire service and any other service that adopts this approach.

- A fundamental principle of high performance systems is that efficiency is maximised by matching resources to demand to achieve the best unit hour utilisation (UHU). Staffordshire has undoubtedly become a very efficient organisation and most of the developments undertaken have been made possible through efficiency savings. Staffordshire believe they are now operating at about maximum performance level and without the input of additional resources it would be difficult to sustain performance. Similarly, if SAS were to lose staff now, it would probably be detrimental to performance. While the morale and vision of the Trust by line and middle management remains high, frustration levels are increasing.

- There are concerns that it will be difficult to maintain the current level of performance and continue to improve the service without additional resources. The focus on improving response times has been to the detriment of other parts of the service particularly in the development of staff training programmes with potential consequences for the quality of clinical care. Current resources allow the day to day running of the A&E service and high performance in terms of response times at current levels of activity. However, there is little left in the system to finance other parts of the operation – training, fleet replacement, long term planning and development. Whilst response time performance is important others may not consider the trade-off in terms of reductions in other parts of the service as acceptable and may make different choices. There also continues to be annual increases in demand for emergency ambulance services and these cannot be absorbed indefinitely within current budgets. The Trust feels strongly that if funding were linked to performance with financial incentives rather than disincentives for high performance then

there would be more scope to properly invest in other parts of the organisation.

- The service acknowledges that the improved performance has been achieved through significant changes in working conditions for staff but with limited change in their pay. The system has worked so far in Staffordshire because staff, particularly road crews and control room staff, have seen the benefits in terms of response time improvement. Having achieved that, there are concerns that it will be difficult to maintain motivation without suitable reward, particularly when staff are aware that they are paid the same as their counterparts in other services whose performance is much lower. Pay and training are crucial issues which need to be resolved if performance and clinical standards are to be maintained or improved further. Experience in mature HPAS systems in the USA has shown that the sustained pressure of working in this type of environment is difficult to sustain. Staff turnover in US services is high with many people staying for less than a year. Some of these services are now looking at alternative modern workload management techniques to try and resolve some of these difficulties and achieve a better balance between the needs of patients and the requirement to provide fast response times and staff welfare. These include more creative performance measures to reward staff for achieving performance targets[18]. Similarly, consideration needs to be given to long term career structures as there may be a limited life span to working at the front line in a HPAS environment. Staffordshire have already begun to develop some strategies in this direction with early retirement, part time and first responder options. In the future credible and creative schemes will need to be in place to support and retain staff over a working lifetime.

References

1. Chapman R. Review of ambulance performance standards. Final report of steering group. 1996; NHS Executive, England.

2. Government statistical service. Statistical bulletin. - Ambulance services, England: 2000-2001 Department of Health, London.

3. Government statistical service. Statistical bulletin. - Ambulance services, England: 1994-95 Department of Health, London.

4. Judge T. A mosaic in transition: contemporary EMS in the United States. Pre-hospital Immediate Care 1997; **1:** 204-212

5. Ryan J. Quality Management. In: Khuel A et al. Pre-hospital systems and medical oversight 1994;. Mosby year book, St Louis, Missouri.

6. Stout J. There's a train coming

7. Nicholl J, Turner J, Martin D. The future of ambulance services in the United Kingdom. The Medical Care Research Unit, University of Sheffield 2001. Ambulance Service Association, London.

8. Audit Commission. A life in the fast lane. Value for money in emergency ambulance services. 1998; The Audit Commission, London.

9. Staffordshire Ambulance service NHS Trust. Staff survey 1999-2000. Pay and Workforce Research. Harrogate.

10. Thayne R, Introduction to High Performance Ambulance Systems. Staffordshire Ambulance Service Trust Document July 1995.

11. Devlin P. Patient satisfaction survey: 1st August 1999 – 31st May 2000. Staffordshire Ambulance service NHS Trust, June 2000.

12. Stout J. System Status Management: The strategy of ambulance placement. In: Pre-hospital care administration, J.J. Fitch (Ed) 1995; Mosby year book, St Louis, Missouri.

13. Devlin P. Treating cardiac arrest patients in the community. Staffordshire Ambulance service NHS Trust, June 2000.

14. Mitchell R.G, Guly U.M, Rainer T.H, Robertson C.E. Can the full range of paramedic skills improve survival from out of hospital cardiac arrests?. J Accid Emerg Med 1997; **14**:(5) 274-277.

15. Mann J.C, Guly U.M. Paramedic interventions increase the rate of return of spontaneous circulation in out of hospital cardiac arrests. J Accid Emerg Med 1997; **14:**(3) 149-150.

16. Rainer T, Marshall R, Cusack S. Paramedics, technicians and out of hospital cardiac arrest. J Accid Emerg Med 1997; **14**:(5) 278-282

17. Wright D, Bannister J, Ryder M, Mackintosh A. Comparison of two methods of transporting paramedics to cardiac arrests outside hospital. Resuscitation 1992; **23:** 193-197.

18. .Fitch J.J. Personal communication.