Why the Closest Ambulance Cannot be Dispatched in an Urban Emergency Medical Services System

Stephen F. Dean, PhD

Department of Emergency Health Services, University of Maryland, Baltimore County, Baltimore, Maryland USA

Correspondence:
Stephen F. Dean, PhD
Department of Emergency Health Services
University of Maryland, Baltimore County
1000 Hilltop Circle
Baltimore, Maryland 21250 USA
E-Mail: sdean@umbc.edu

Funding/Support
The author has no financial arrangements that may represent a conflict of interest.

Keywords: ambulances; ambulance response time; Emergency Medical Services communication systems; paramedic

Abbreviations:
ALS = advanced life support
CAD = computer-aided dispatch
EMS = emergency medical services
NREMT-I = Nationally Registered Emergency Medical Technician-Intermediate

Received: 26 March 2007
Accepted: 01 August 2007
Revised: 09 October 2007

Web publication: 18 April 2008

Abstract
Introduction: Response time performance is related to increased survival for a relatively small group of patients with critical emergencies. Effectively utilizing current resources is a challenge for all emergency medical services (EMS) systems for reasons of cost-effectiveness and safety.
Problem: The objective of this study was to identify opportunities for improving ambulance response-time performance in an urban EMS system using fixed deployment.
Methods: This was a qualitative and quantitative case study which consisted of structured interviews with policy makers, managers, and workers in a fire department EMS division, as well as analysis of dispatch data and observation of dispatch operations.
Results: The current computer-aided dispatch (CAD) system does not identify the closest ambulance to the emergency, and therefore, dispatchers must guess which unit is closer when units are not within their stations or "first due" areas. There is no means to track how often dispatchers guess correctly or how often the closest ambulance actually is dispatched to the emergency. Temporal and geographic patterns were identified. Opportunities also were identified to improve response time performance through the use of dynamic deployment and peak-load staffing.
Conclusions: The results suggest that there were opportunities for improving ambulance response times by implementing strategies such as peak-load staffing and dynamic deployment. However, the most important improvement would be the implementation of a policy to send the closest ambulance to the emergency. More research is needed to identify how prevalent the failure to send the closest ambulance is within EMS systems that use fixed-deployment response strategies and computer-aided dispatch systems that are incapable of tracking unit locations outside of their stations.

Dean S: Why closest the ambulance cannot be dispatched in an urban emergency medical services system. ***Prehospital Disast Med*** **2008;23(2):161–165.**

Introduction
Response time performance is related to increased survival for a relatively small group of patients with critical emergencies.[1,2] The American Heart Association Guidelines for Cardiopulmonary Resuscitation (CPR) and Emergency Cardiovascular Care recommend that emergency medical services (EMS) systems "try to shorten response times" when feasible, for victims of cardiac arrest.[3,4] Evidence demonstrates the effectiveness of two strategies for improving response times when temporal and geographical demand patterns are identified: peak-load staffing and geographic deployment.[5–8]

Problem
The objective of this study was to identify opportunities for improving response times in an EMS system in a major metropolitan area that utilizes fixed deployment. In a system using fixed deployment, units are located at specific stations, respond from those stations, and then attempt to return to their station after the completion of a call. In a system using dynamic deploy-

ment, units are moved to different locations within the service area based upon the temporal and geographical patterns of demand.[9]

Early research by Savas examined the relationships between reduced response times and lower costs.[5] Reducing total "out of service" time resulted in increased unit availability and a reduction in the amount of units needed to service total demand. The use of satellite stations reduced response times as well. Arreola-Risa *et al* also documented improvements in response times by using satellite locations for ambulance stations.[10] Increasing the number of locations from which ambulances were dispatched from two to four resulted in a six-minute reduction in the mean value of the response times. Results of a study by Peleg and Pliskin[11] suggested that improvements in response time and cost-effectiveness also could be achieved through dynamic deployment of ambulances based on temporal demand patterns.

More recently, Maguire *et al* discussed the risks experienced by EMS personnel in the workplace.[12] Motor vehicle crashes account for the majority of on-the-job deaths in EMS. As Clawson notes in the justification for priority dispatching, a great part of that risk occurs as the result of driving to emergency scenes.[3,13]

Therefore, in addition to patient care, there are operational reasons to send the closest ambulance and improve response time performance. These include reducing overall service time and therefore, reducing the number of needed units, and decreasing the risk of motor vehicle accidents by decreasing the amount of time units are responding in emergency mode.

Methods

Both qualitative and quantitative methods were employed to analyze the response time performance of an urban ambulance service including case study methodology, an interview tool, and the analysis of dispatch data. Providers, managers, and administrators were interviewed to identify opportunities for response time improvement and to identify what obstacles might prevent implementation of these improvements.

An exemption was granted by the Institutional Review Board at the University of Maryland-Baltimore County.

Study Design

Paramedics, managers, administrators, and physician medical directors were interviewed. Operations also were observed at the dispatch center. The temporal distribution of EMS responses was analyzed using 40 weeks of dispatch data. One month's data were used to analyze geographic demand patterns and to calculate response time performance.

Site or Population Selection

A fire department serving an urban area with a population of approximately 600,000 persons and a population density of 8,000 per square mile was studied. The EMS Division agreed to cooperate with the study and was able to provide dispatch data and access to personnel. Ambulances transported more than 80,000 patients during the study period.

Data Gathering Methods

Interviews were conducted among members of the EMS Division including administrators, supervisors, medical directors, and paramedics. The interviews were tape recorded and transcribed. In the Communications Center, call-takers and dispatchers were observed processing requests for emergency medical assistance.

Dispatch Data

Dispatch data for the 12-month period from January to December 2003 were obtained from the Dispatch Center for analysis of temporal demand patterns. The 40 weeks with the highest weekly volumes were used to calculate the temporal analysis. This included 95,960 responses. Geographic demand patterns and actual response time performance were calculated using data from August 2003, which were judged to be a typical month in which there was no unusual weather and included 7,233 transports. Department statistics also were used to examine individual unit call volume, but these data were only available for analyzing call volume by unit and not by geographic factors such as response zone, zip code, or census tract.

Ambulance Service

The system uses advanced life support (ALS) ambulances staffed with at least one [US] Nationally Registered Emergency Medical Technician Intermediate (NREMT-I). In the US, a NREMT-I has participated in approximately 400 hours of training, including advanced airway management and the treatment of lethal cardiac arrhythmias with pharmacological agents.

The ambulance crews work the same shift as the firefighters in the city. They work two 10-hour day shifts followed by two 14-hour night shifts, followed by four days off. Their average work week is 42 hours. Some EMS crews are busy the entire night shift responding to high call volumes.

9-1-1 and Call Receipt

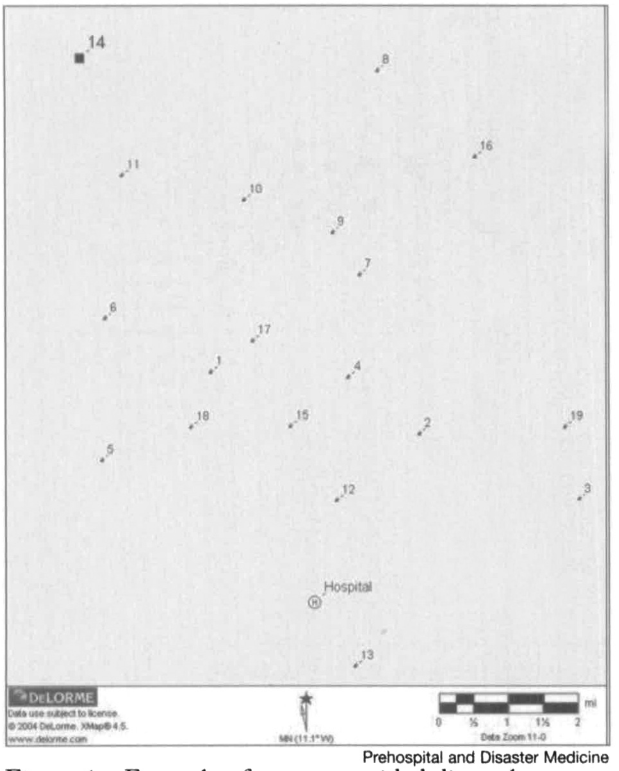
Requests for emergency service made via 9-1-1 (emergency telephone service) are received in the police dispatch center. Requests for ambulance service are transferred to the dispatch center located at the backup public safety answering point (PSAP). Both the caller and electronic dispatch records are received at one of the two EMS call-taking positions. The call-takers use the Clawson protocols to determine the severity of the emergency and to provide pre-arrival first aid instructions.

Ambulance and First-Responder Dispatch

The system's ambulances are staffed 24 hours per day, seven days per week. However, all units may not be available to the system at any given time because of mechanical failures, training, or administrative details. The system does not track these lost unit hours.

The Dispatch Process

The dispatcher uses a computer-aided dispatch (CAD) system to assist in identifying the nearest ambulance to the location of the emergency. The CAD displays the address



Prehospital and Disaster Medicine

Figure 1—Example of computer-aided dispatch (CAD) recommendation

of the emergency and a box number. Every address in the city is assigned a box number that is based on the location of the nearest fire station. The first part of the box number identifies the number of the nearest fire station, and the second part identifies a district within the box area. The dispatcher is shown the numbers of the nearest fire stations whose ambulances are available for a call.

The CAD also shows a list of available ambulances identified as "Medic" units. If the ambulance is available at a location other than its own station it is marked with a "@" sign before the ambulance number.

Most of the time, fire suppression units either are at their stations or within their "first due" districts, the area of the city in which the unit is listed as the closest unit. Therefore, whether the unit is available at the station or in its district makes no difference in terms of whether the suppression unit is closest to a possible fire.

However, in the case of ambulances that transport patients to hospitals within the city, and in some cases outside of the city, there is a difference between being available at the station or at another location. During periods of peak activity, many, if not all of the available ambulances will be far from their home stations. Therefore, the CAD recommendation for the closest unit often will be incorrect.

If Ambulance 14 had finished a patient transport and had just become available for service at a hospital in the southern portion of the city (circle labeled "Hospital"), and all other medic units were available at their station locations as shown on the map, Ambulance 14 would be recommended as the closest available ambulance to a call at its own station in the northwest corner of the city even though 15 of the other 18 stations in the city actually are closer (Figure 1).

Results

The demand for ambulance service varies by the hour of the day and by the day of the week (Figure 2). The busiest times in the system are from 09:00 to 22:00 hours (h). There are times when the system has excess capacity and also times when the system is overtaxed. When ambulances complete a call, they attempt to return to their stations. During busy periods, the units seldom make it back to the station before they receive another call.

Analysis of geographical demand patterns indicated that >40% of the call volume occurs in the core of the city which constitutes 11% of the city's area. There are more calls in this core area and fewer calls in other parts of the city during all hours of the week. Twelve months of call volume data were examined, and the five units stationed in the core of the city accounted for 33% of the total call volume. This statistic does not include calls into the core area taken by units in adjoining districts.

Therefore, if there are only four or five units available in the system, it would seem that positioning them based upon expected call volume rather than their station number might improve the likelihood of a rapid response. However, by using the fixed-deployment strategy, an ambulance may remain in the northwestern part of the city while there are no ambulances stationed in the central portion of the city, where calls are statistically most likely to occur.

The dispatch system does not track the locations of units outside of their station, so it is impossible to determine from CAD records what percentage of the time the closest unit actually is sent to an emergency call. One medic observed that accuracy might be improved if dispatchers were trained "to know the locations of the hospitals as well as the stations and know when a medic unit is not in their station." Several interviewees noted that it was not uncommon for units to exchange calls among themselves in order to avoid passing each other.

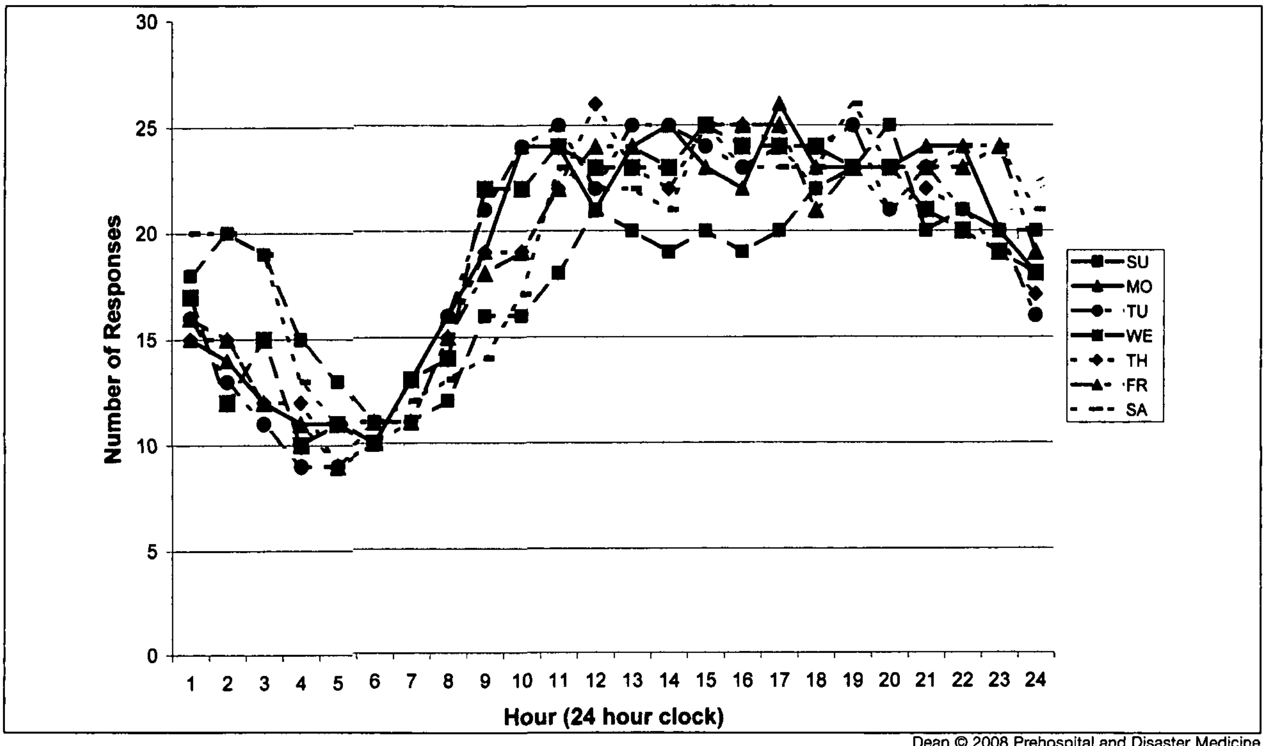
If the emergency is life-threatening or if the ambulance response will be delayed, a fire department engine or ladder company is sent to the call by the fire dispatcher who sits near the EMS dispatcher.

When the system runs low on available ambulances, "overload medic" units are placed in service. These units may be basic life support (BLS) or ALS. In 2003, there were 783 transports by overload medic units.

Response Time Performance

Ambulance response-time performance is shown in Figure 3. This system produces response times of 13:34 minutes with 90% reliability to all emergency calls. These statistics were compiled from dispatch data from August 2003. The EMS Division does not track response time performance using a frequency distribution, so calculations were made using their data. The Dispatch Center also does not distinguish life-threatening from non-life threatening emergencies in the dispatch data.

Part of the problem confronting the system is the amount of time it takes to dispatch calls, and then the



Dean © 2008 Prehospital and Disaster Medicine

Figure 2—Temporal demand analysis

amount of time it takes crews to start responding once they are notified of the call (en route time) (Table 1). Unit travel times are ≤8 minutes with 84% reliability. Dispatch policy generally is that the unit will return to its home station regardless of the number of units available, hour of the day, or day of the week. However, there are certain areas of town, such as the central area of downtown that are more likely to generate a call than other areas.

Another problem in the system is the manner in which staff changes are scheduled and performed. The same number of units is scheduled for each hour. All shifts start and end at the same time. There are two shifts each day with an equal number of units assigned to each shift. Day shifts are scheduled from 07:00 to 17:00 h, and night shifts from 17:00 to 07:00 h. However, ambulance crews realize that 17:00 h is a peak time and that they may get off late, so they make informal arrangements among themselves to change shifts between 15:30 and 16:00 h (starting their shift at 17:30 or 18:00 h instead of 17:00 h) so they may get off on time. Management does not schedule orderly shift changes. If all crews are trying to get back to their stations at 17:00 h, theoretically the system has no available units for some period of time while this is accomplished. The crews have informally staggered their end of shift times for their convenience, not to produce rapid response times.

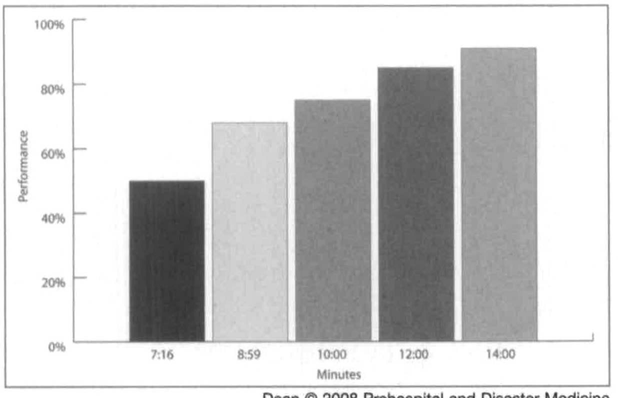
Discussion

The results were surprising because it was thought that response times could be improved through the use of peak-load staffing and perhaps through the use of dynamic deployment. Managers were concerned about how these proposed changes would be received by the workforce, and particularly by the union. However, the discovery that the system is unable to dispatch the closest ambulance to emergency calls and is unable to determine retrospectively whether the closest available ambulance was dispatched to a particular emergency was unexpected. Management stated that the system's policy was to send the closest ambulance. However, a recent survey of EMS systems found that only 55% reported that they actually sent the closest ambulance to each call.[14]

Increasing the number of units available for dispatch may not improve response times if the dispatcher is unable to determine which unit is nearest to the call. In fact, giving the dispatcher more choices mathematically decreases the odds of guessing correctly.

There also are safety and operational advantages to sending the closest ambulance. Reducing the length of time units are responding with lights and siren may reduce the risk of crashes. It also is dangerous for emergency vehicles to be driving past each other on emergency responses because drivers become confused when they are trying to locate and yield to ambulances that are approaching from multiple directions. Operationally, reducing response times decreases out-of-service times, which increases the number of available units.

Emergency medical services systems have been using computers to assist in dispatching the closest ambulance since 1981.[15] However, all CADs are not able to provide dispatchers with information about the location of the nearest available ambulance. In this case, the CAD was designed for use with a fire department in which it is assumed units will either be in their stations or in their first due areas. When ambulances are available at hospitals, the CAD is unable to provide the dispatcher with an accurate recommendation. Crews reported that it was not uncommon for units to pass each other responding to calls. The



Dean © 2008 Prehospital and Disaster Medicine

Figure 3—August response time performance

NFPA 1710 Standard (time in minutes)	Reliability Standard	Actual Reliability
Dispatch Time (1:00)	90%	9%
Enroute Time (1:00)	90%	43%
Travel Time (8:00)	90%	84%

Dean © 2008 Prehospital and Disaster Medicine

Table 1—Response time performance vs. NFPA 1710 Standard

procedure for reassigning calls was described as, and observed to be, cumbersome and time-consuming.

The results of this study are limited by the fact it is a study of only one system and one particular CAD.

Conclusions

Opportunities for improving response time performance were identified in this system including the use of peak-load staffing and dynamic deployment. However, the inability of the dispatch center to always send the closest ambulance to the emergency and their inability to track how often the closest ambulance was sent to a call has created a serious obstacle to improving response time performance. The medical director of the system did not have direct authority over the Dispatch Division. The assumption that a dispatch center always sends the closest ambulance to an emergency was not correct in this system, and may not be correct in other systems using fixed deployment and a similar type of dispatch process and computer software.

Further study is needed to determine how prevalent this problem is because a large percentage of EMS systems use fixed deployment.[16] Adding units and personnel to systems in which the dispatch process is flawed may not produce expected improvements in response time performance. At a minimum, medical directors and managers should obtain reports from their dispatch center showing the number of incidents to which the closest ambulance was dispatched. Medical directors should also have access to back-up detail for each response showing the actual location and status of each ambulance in the system at the time each response is dispatched.

References

1. Pons PT, Haukoos JS, Bludworth W, *et al*: Paramedic response time: Does it affect patient survival? *Acad Emerg Med* 2005;12:594–600.
2. Blackwell TH, Kaufman JS: Comparison of response time and survival in an urban emergency medical services system. *Acad Emerg Med* 2002;9:288–295.
3. Clawson J: Priority Dispatch Response. In: Kuehl (eds), *Prehospital Systems Medical Oversight*. Dubuque, IA: Kendall/Hunt, 2002, pp 208–228.
4. International Liaison Committee on Resuscitation: 2005 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. *Circulation* 2005;112(24s1).
5. Savas ES: Simulation and cost-effectiveness analysis of New York's emergency ambulance service. *Management Science* 1969;15:608–627.
6. Holloway R: New York City's experience in improving ambulance service. *Health Services Reports* 1972;87:445–450.
7. Ryan JL: Quality Management. In: Kuehl A (ed), *Prehospital Systems Medical Oversight*. St. Louis: Mosby, 1994, pp 217–246.
8. Overton J: *High Performance and EMS: Market Study 2002*. Richmond: North American Association of Public Utility Models, 2002.
9. Overton J, Stout J: System Design. In: Kuehl A (ed), *Prehospital Systems Medical Oversight*. St. Louis: Mosby, 2002, pp 114–131.
10. Arreola-Risa C, Mock CN, Lojero-Wheatly L, *et al*: Low-cost improvements in prehospital trauma care in a Latin American city. *J Trauma* 2000;48:119–124.
11. Peleg K, Pliskin JS: A geographic information system simulation model of EMS: Reducing ambulance response time. *Am J Emerg Med* 2004;22:164–170.
12. Maguire BJ, Hunting KL, Smith GS, *et al*: Occupational fatalities in emergency medical services: A hidden crisis. *Ann Emerg Med* 2002;40:625–632.
13. Clawson J: Emergency Medical Dispatch. In: Kuehl A (ed), *Prehospital Systems Medical Oversight*. Dubuque: Kendall/Hunt, 2002, pp 172–207.
14. Williams DM: 2006 JEMS 200-City Survey. *JEMS* 2007;32:38–54.
15. Fitch JJ: Strategic Deployment. *JEMS* 2002;27:36–45.
16. Williams DM: 2005 JEMS 200-City Survey. *JEMS* 2006;31:44–61,100–101.