# EMERGENCY MEDICAL SERVICES: SYSTEM DESCRIPTION

Edward H. Blum

October 1975 WP-75-137

## Introduction

This working paper presents a draft of a forthcoming report. The report will contain more detail (including diagrams) and numerous references to existing literature. Both the working paper and the report are intended to serve as the basis for a chapter in the forthcoming monograph, Analysis, Planning and Management of Urban Emergency Services.

The author welcomes comments. Responses to the following two questions would be especially appreciated:

(a) How can the description of the emergency medical services system be sharpened and clarified, (i) for analytically oriented readers interested in the analytical structure of the system, and (ii) for managerially oriented readers interested more in how the system works and where a particular EMS system may best be improved?

(b) How can the description be made both more general and more precise, so that it will apply as well as possible to actual systems in all IIASA's member countries?

# Emergency Medical Services: System Description

Edward H. Blum

## I. Emergency Medical Services in Health Care

As the name implies, emergency medical services (EMS) respond to and treat medical emergencies--situations that arise quickly and need immediate attention to reduce injury or preserve life. They may also provide a variety of related services, such as the transportation of disabled patients. As such, then, emergency medical services are a conspicuous means of health care delivery.

But they are certainly not the only means of delivering health care, and what they do strongly affects what some of the other means do, can do, or try to do, and vice versa. In formal terms, this means that EMS are imbedded in a larger health care system (whether or not that system is explicitly acknowledged or managed as such), and EMS policies and resources both supplement and compete with those of other parts of the system.

At the broadest level, for example, there are important policy tradeoffs between prevention and treatment (EMS are predominantly part of the treatment subsystem). Many of the most serious emergencies with which EMS deal stem from traumatic accidents and cardiac problems; these can be partially prevented, and the number and severity of incidents reduced. Accident prevention, for example, entails extensive public education about possible hazards and safe practices, and infrastructure changes (such as addition of technological safety features, modifying working environments, regulating or influencing alcoholic consumption, changing traffic regulations and patterns, etc.). Reduction of cardiac emergencies entails extensive screening and monitoring of potential

victims, with appropriate preventive treatment and medication, as well as long-term supporting research (since so many key elements are yet ill-understood).

Viewed from the perspective of systems analysis, preventive activities in general involve reaching and/or screening large numbers of people, few of whom are immediately at risk, at a relatively low cost per person (but perhaps a high total cost), with potentially great leverage but quite uncertain effects. There are serious tradeoff problems in setting thresholds between statistical type I and type II errors. Emergency response and treatment, on the other hand, deals with many fewer people, a very high percentage of whom are immediately at risk, at a high cost per person, with low-to-high leverage (depending often on circumstances beyond medical control) but relatively certain effects. There are some slight threshold problems, but in general EMS are extremely risk-averse.

There are also many alternative ways of delivering (or obtaining) treatment. In most health care systems, patients have a choice of entry points, organizational forms, specialized facilities, centralized or decentralized operations, etc. Organizationally, for example, there may be individual doctors or groups of doctors, clinics, hospitals, intermediate facilities at places of work and public assembly, etc. These organizational units may be oriented toward in-patients (e.g., with beds and their supporting facilities) or out-patients; they may be stationary (with patients coming to them) or mobile (travelling to the patient), etc.

Within any of these orientations, units may specialize according to disease or type of medical problem (to achieve managerial, medical, and technological concentration), or by target populations or clientele (e.g., children, workers, women, elderly people, etc.). Within these subdivisions, they may also specialize on short-term versus long-term care,

for example, or (though the policies are rarely stated so explicitly) on morbidity (loss of productivity) versus mortality, or vice versa.

Whatever their specific focus, these units all use very similar scarce resources: doctors and other skilled (and semi-skilled) personnel; pharmaceuticals; medical equipment and technology; medically suitable treatment space (e.g., beds, operating rooms); research personnel and facilities; laboratories; etc. They thus implicitly supplement and compete with one another, whether or not they are formally considered to do so.

Indeed, to the extent that their potential clientele overlap, they compete for patients as well as resources. Where family doctors or out-patient clinics are inaccessible, inconvenient, or expensive (in terms of money, waiting time, or whatever rationing mechanism the society uses), people are likely to turn more often to emergency rooms and to ambulances. Where EMS are regarded as less adequate than alternative services, people are likely to substitute the alternative services at the margin, until the probable costs of the alternatives (perhaps in terms of potential inability to respond promptly to a serious condition) are perceived to outweigh the benefits.

The demand for EMS in a particular city thus depends not only on the intrinsic numbers of serious medical situations but also on how people perceive the quality of the EMS relative to what they consider alternative sources of treatment for their conditions. A city perceived to have excellent EMS, such as Moscow, thus would be expected to have relatively high EMS use per capita. A priori, one would anticipate that people will tend to call the EMS in Moscow more often than they would in other cities for moderately serious or medically marginal problems. This substitution effect can be quite substantial, and thus

must be included in EMS demand analyses that compare several cities or examine one city over an extended period of time.

Just as the volume and composition of demand for EMS depend on the rest of the health care system, so does the ultimate quality of health care EMS deliver. A cardiac victim kept alive and brought in by ambulance, for example, must eventually be turned over to heart specialists, perhaps for intensive care immediately upon arrival or for less drastic treatment and long-term follow-through on an in-patient or out-patient basis. Initial high quality emergency care can be vitiated if the subsequent care is poor. Maintaining high standards throughout the care sequence requires good communication, follow-through, and conscious, consistent quality control.

## II. EMS Entry Points

Most people enter the emergency medical system in one of four ways:

(1) By a call to the emergency ambulance service (made by someone else if the person is incapacitated), or to the police (who may transport some people directly);

(2) By a call to an emergency counselling or referral service--such as a poison control center, a suicide prevention (or personal "crisis") counselling service, a telephone medical advisory service, such as Vienna's Ärtzenotdienst, etc.;

(3) By coming to the admissions station of an out-patient or in-patient emergency room (that is, transported privately, rather than by ambulance); or

(4) By the request of a physician. The physician may be the patient's private doctor (or the clinic equivalent) or part of a counselling, referral,

or advisory service. He or she may request emergency or quasi-emergency service (such as a house-call by a qualified professional or team), or transport service--usually for someone who is unable to travel on his own or should have attention while being transported between hospitals or between home and a medical facility, etc.

To some extent, these entry points are complementary; a large fraction of the demands for EMS could enter by any of several routes, depending on what the patient (or those taking responsibility for him) choose to do first. This choice, of course, is influenced by people's perception and understanding of the possibilities open to them, and the relative opportunitites (one may be nearer one's car than a telephone, etc.) at the decisive moment. Again, substitution at the margin of one alternative for another may be significant, and must be accounted for in planning or studies focussing on one entry route, such as ambulance calls.

Notwithstanding, this paper focusses primarily on entry method number one--ambulance calls and their consequences. Peripherally, it also concerns entry method number four--doctors' requests for transport services. Entry points numbers two and three are important in many cities, but are of less immediate concern for the work described here.

## III. Ambulance Services

As noted earlier, ambulance services divide generally into two types: emergency or rescue services, where time is of the essence and the medical problems may be severe; and transport services, which are essentially specialized

taxis. Some cities combine the two services, using the same vehicles and crews, allocating them as needed and appropriate; in the field, these combined services are often called "one-tier" systems. Other cities separate the services, using vehicles with less equipment and less specialized crews for transport. Often the transport service trains drivers and technicians for the rescue service. These separated services are termed "two-tier" systems. Which organization is better continues to be argued.

This paper focusses on the emergency ambulance service, which contributes much more to overall medical care and raises many more important planning, managerial, and analytical issues. The main issues in planning and managing transport services are demand forecasting and control, coordination among physicians, facilities, and patients, and scheduling and routing the vehicles.

Key emergency ambulance issues and opportunities for planning and managerial improvement are highlighted by considering the "flowchart" or time-sequence of events that an emergency medical incident can generate:

1. Prevention: Before an incident occurs, technical social, and economic means are applied to try to prevent it or reduce its potential severity. Most of the ideas for preventive activities derive from retrospective analyses of past incidents. The linkage between preventive programs and effects (if any) is still poorly understood, and few systematically designed experiments have been attempted.
2. Occurrence of Problem: Failures of preventive efforts occur as problems, which may be organic malfunctions or diseases, accidents, vaguely defined malaise, or other medical problems (e.g., imminent childbirth, allergic attack). The nature and severity of the problem may not be apparent or clear to those affected or on the scene.

3. Detection of Problem: Detection is the classic problem of extracting a signal from "noise"--i.e., trying to distinguish the situation at hand from the many perturbations that are "normal" or not truly serious. People having cardiac trouble, for example, may not recognize it, or if they are incapacitated, others around them may not recognize the gravity of their distress; people who are hurt may not be noticed, etc. The time to detection is often one of the more critical delays. It is most effectively reduced by public awareness, reinforced by public training in at least problem recognition and first-aid.
4. On-the-Scene Preliminary Diagnosis and Treatment: Once a problem has been detected, the victim or those around him attempt some form of preliminary diagnosis and treatment. Some people may attempt little--simply stand around and watch, and see if the victim asks for help. At the other extreme, with suitable training, others may begin artificial respiration or external cardiac massage. In very serious cases, this immediate on-the-scene attention may mean the difference between life and death, keeping the victim alive long enough to reach professional care. It is most effective, of course, when supported by an extensive public training program, as noted in number three, above.
5. Notification of the EMS: As soon as the problem is detected, and while or before the preliminary diagnosis and treatment are carried out, the EMS should be notified. Delays at this stage are best reduced by having widely available points or means of notification (e.g., public telephones or street-corner boxes wired directly to the police and/or fire department), with the easiest possible access (easily remembered number, widely publicized and posted coin-free emergency dialing, etc.).

6. Problem Identification: The call for assistance is usually received at a central dispatching station, which may serve several emergency services or EMS alone. The dispatcher answering the telephone attempts to elicit as much useful information as possible from the caller, to identify at least
   (a) Where the problem is, in enough detail that the ambulance can find the spot quickly;
   (b) Who is involved--e.g., how many people, what age, etc.;
   (c) What assistance, if any, is already being rendered (e.g., are the police already on the scene); and
   (d) What the problem seems to be.

   The dispatcher's aim is to be able to assign quickly to the call
   (i) a priority (especially if there are so many calls that resources must be rationed), taking into account an estimated assessment of need for EMS and an estimate of importance relative to other calls, and
   (ii) the appropriate response, taking into account possible need for specialized services (e.g., a cardiac-care unit) or multiple units (e.g., for a building collapse or train derailment).
7. Allocation Decision: Once the problem has been identified as well as possible, given the caller's information and abilities and the time constraints, a dispatcher must decide which EMS units to allocate. This dispatcher may be the same one who answered the telephone (in a one-stage dispatching system) or a different one (in a two-stage system, where one group deals with the public and another with the operational units). Tradeoffs in queueing, and in specialization versus continuity, determine which system structure is better for a particular city.

The immediate allocation decision is constrained by longer-range decisions that have set the numbers of EMS units of different types that are stationed at various locations at various times, and by recent demands for service that have made some of those units unavailable for the present dispatch. In addition, the dispatcher may want to consider demands that might be expected to come in while the units to be dispatched now will still be unavailable--especially if, for example, a minor catastrophe or series of serious incidents has already nearly exhausted a region's EMS resources and the present call does not seem especially urgent or serious. In such situations, this decision involves difficult tradeoffs under uncertainty and risk.

If one region has or will shortly become depleted, the dispatchers also must consider how to cover it. Options here include temporary relocation of EMS units from other regions, activation of auxiliary or reserve units, relocation of units from neighboring jurisdictions, etc.

Algorithms to assist dispatchers with the issues of problem identification and allocation can be (and for most issues, have been) developed. Many of these are so complex that a computer is needed to do the calculations in real time, especially in services having large numbers of units; approximate, simpler algorithms (which may also be more robust) usually can be developed for the most frequent decisions, however.

8. Dispatch: Once the allocation decision has been made, the actual dispatch consists of communicating the decision as an order to the appropriate units. Any rapid, reliable, and secure communication means can suffice: telephone, internal public address system,

teletype, coded digital signals (bells, buzzers, etc.), radio-telephone, etc. Some form of acknowledgement or feedback from the units is often incorporated for control; automatic systems (such as a signalling treadle in the vehicle's exit way) are used in many cities.

9. Response: Once the dispatch has been made, the next critical factor is getting service to the scene as promptly as possible. Some help can be provided immediately, in selected cases, if the dispatching center has qualified medical personnel trained to give preliminary advice and counselling over the telephone. The caller can be instructed in the most elementary first-aid ("Is he breathing? If not, or with difficulty, look for..."), queried for information that may help the crew begin service immediately upon arrival ("...is there a history of diabetes, then?), or asked to help the crew find an obscure location.

Major aid, of course, awaits the arrival of the vehicles(s) dispatched. What assistance the crew can provide on arrival depends on the number and skills of the personnel, the capacities and capabilities of the equipment on board, and the support (via communication links) offered in difficult cases by the stationary parts of the EMS system.

The time taken to respond, a large part of the total "response time" from call to the beginning of professional assistance at the scene, depends on the location of the vehicle(s) relative to the incident and, to a lesser extent, the travel speed (experience in many cities shows remarkably similar average emergency vehicle speeds under widely varying traffic conditions, etc.). The quantitative value of decreased response time is still debated; it is unanimously agreed, however, that for

serious emergencies, the sooner the arrival, the better the chance of saving the victim's life or reducing the effects of his injuries.

10. On-the-Scene Diagnosis and Treatment: On arrival, the crew must quickly assess the situation and decide upon appropriate action. At this stage, incidents (or through triage, the individuals in a multiple victim situation) are effectively placed into one of the following classes:

(a) Serious medical emergency ("emergent");
(b) Moderate medical emergency ("urgent");
(c) Medical problem, but not urgent;
(d) No real medical problem, but good reason for caller to have thought there was;
(e) Possibly once a medical problem, but situation now normal;

The real world being what it is, an incident may also be found to be:

(f) No medical problem, and no reason to have thought there was--e.g., a call for transport in the guise of an emergency call;
(g) Malicious or prank call.

The distinction between calls ranking below (c) in medical importance matters little for action at the site, but does matter for broader policy, such as public education for events number three, four, and five.

Having assessed the situation, for classes (a) to (c) the crew has four basic options:

(i) Perform appropriate minor treatment and release the patient, either outright or to the care of a non-EMS physician;
(ii) Perform appropriate treatment and stabilize the patient for transport to an EMS treatment facility;

(iii) Perform extended, on-the-scene medical treatment prior to transport to an EMS facility for further treatment;

(iv) Conclude that the victim is beyond saving or already dead, and transport to facility for appropriate treatment. Rarely exercised so overtly, this last option often must be exercised, de facto, when there are multiple victims who must be assigned initial treatment priority.

The option selected depends upon the crew's abilities and assessment of bureaucratic rewards and risks as well as on the medical situation. Crews with little medical training, or in organizations that punish risk-taking, nearly always choose option (ii). Usually, they fear to release patients who might later be found to have needed medical attention, and are rarely qualified or inclined to undertake major treatment at the site (which entails medical risks and delayed transport to the hospital). The greater the crew's training and the more flexible the organization, the greater is the range of options likely to be chosen.

The range can also be expanded, and the service quality within options improved, by good communication links to central expertise. A basic two-way radio link to the hospital can help provide supplementary medical advice and answer urgent questions. At a more advanced level, biotelemetry of vital functions (perhaps supported by real-time computer processing) can further assist diagnosis and help monitor the course of treatment.

11. Transport and In-transport Treatment: In options (ii) and (iii), above, patients are loaded aboard the ambulance for transport to an appropriate treatment facility. Though there is a sizeable literature (with considerable

controversy) concerned with design and operation of ambulances, and substantial variation in practice from country to country, the determinants of quality in this service stage appear straightforward and essentially invariant.

Depending on the severity of the patient's condition, and the treatment performed at the incident site, further treatment may be desirable enroute. Because of the motion, only relatively simple care can be provided. The quality of this care, like that at the scene, depends on the skills of the crew, the equipment on board, and the communications (possibly including telemetry) to additional expertise.

When the patient's condition is moderately serious or worse, prudent speed is vital. A radio link to the receiving facility can further reduce the time until effective in-hospital treatment begins, by notifying staff of the incoming emergency and giving them a head start on assembling the people and equipment needed to deal promptly with it.

In some cities, the time from scene to treatment facility also depends importantly on the organization and location of emergency receiving rooms. Centralized facilities may have medical advantages, but may take so long to reach from certain areas that the delay offsets some of the nominal gains.

12. Transfer to Treatment Facility: When the ambulance arrives at the treatment facility, the patient must be effectively transferred from the vehicle to a suitable location in the facility. This transfer is facilitated by good physical design of the receiving area, rapid and efficient admitting procedures, and responsive

emergency room organization. Where there are large numbers of incoming patients (not all of whom need arrive by ambulance, as noted in section II), entry triage (medical priority assignment) may be necessary to ensure that only non-essential cases are forced to queue. Such triage is especially important here where it has not been done at the scene (i.e., where crews tend to choose mainly option (ii) at step ten).

13. Final Diagnosis and Treatment: In the hospital, the patient's condition is again diagnosed, this time presumably in the best possible medical circumstances. Where this final diagnosis often disagrees significantly with those attempted at earlier stages (by the caller, as interpreted by the dispatcher, and by the crew arriving at the scene), operational improvements (supported by analysis) are indicated.

Public education and training programs, noted as important at earlier stages, may help improve the information callers supply. Better training for dispatchers can sharpen the questions they ask and their interpretations of the replies. Improved training for crews in diagnostic procedures and in recognizing serious emergency conditions can speed application of correct treatment. All these can be focussed by detailed analysis and comparison of the diagnoses at successive stages, using methods such as decision theory and factor analysis to relate quantitatively what is or can be known at the earlier stages to the most probable, final best estimate.

Obviously, once the condition has been suitably diagnosed, appropriate treatment should begin as soon as possible. The better the crew's diagnosis and the better the communications from the scene and enroute,

the sooner the right treatment at the facility is likely to start. And the better the ability of the public and the crew to supply immediate treatment at the scene, the more likely it is that victims with serious problems will reach the hospital in condition to benefit effectively from the treatment it can provide.

The treatments that can or should be provided, and the means needed to achieve them, are the province of emergency medicine, per se--a large, burgeoning field, with which this work is not directly concerned.

14. Follow-through: Once the patient has entered the hospital for care, particularly if he enters in serious condition, the emergency medical service's concern with him cannot end completely. It is important for EMS management to know the patient's ultimate progress and condition, both to assess relations with subsequent treatment stages (and perhaps to influence their quality) and to ascertain whether ultimate failures (deaths, lasting severe disabilities) may have been due to EMS deficiencies or to exogenous circumstances over which they might have some influence or control.

Follow-through is also an integral part of quality control--both inside EMS and of EMS by higher or regulatory authorities. Detailed studies of deaths and other instances of undesired or unanticipated outcomes can yield insight into potential improvements (or at least advance understanding of emergency medicine). Re-examination of cases sampled according to a systematic design can help monitor the quality of various service features and reinforce continuing education for staff, dispatchers, and crews.

Moreover, higher authorities may be interested in emergency cases that would have been expected to enter the system by particular routes (e.g., ambulance calls) but did not. Such cases can highlight the nature and extent of service substitution (see sections I and II), and illuminate why it has occurred.