Triage in Medicine, Part I: Concept, History, and Types

Kenneth V. Iserson, MD, MBA
John C. Moskop, PhD

From the University of Arizona, Tucson, AZ (Iserson); and the Brody School of Medicine at East Carolina University and the University Health Systems of Eastern Carolina, Greenville, NC (Moskop).

This 2-article series offers a conceptual, historical, and moral analysis of the practice of triage. Part I distinguishes triage from related concepts, reviews the evolution of triage principles and practices, and describes the settings in which triage is commonly practiced. Part II identifies and examines the moral values and principles underlying the practice of triage. [Ann Emerg Med. 2007;49:275-281.]

0196-0644/$-see front matter
Copyright © 2007 by the American College of Emergency Physicians.
doi:10.1016/j.annemergmed.2006.05.019

SEE RELATED ARTICLE, P. 282.

INTRODUCTION

When the needs or demands for medical treatment significantly outstrip the available resources, decisions must be made about how to distribute these resources, recognizing that not all needs will be satisfied immediately and some may not be satisfied at all. Decisions about distributing scarce health care resources can arise at all levels, from societal choices within a national health care system (macroallocation) to individuals allocating immediate emergency treatment and transport among the multiple severely injured survivors of a motor vehicle crash or industrial accident (microallocation). Several terms, including "triage," "rationing," and "allocation," are used to refer to the distribution of scarce resources in different health care contexts. This article will focus on "triage," the term most commonly used to mean the sorting of patients for treatment priority in emergency departments (EDs) and in multicasualty incidents, disasters, and battlefield settings. Most discussions about triage address practical questions, such as when the process should occur and which techniques are most effective. Commentators rarely consider the essential characteristics of triage, the historical evolution of the practice, or the ethical justification for selecting those who will receive priority treatment—or any treatment—among a large group of acutely ill and injured patients. In essence, triage discussions usually focus on when and how to cut the resource "pie," not whether providers should be using a particular tool to do the cutting—or whether they should be cutting the pie at all.

This 2-article series seeks to remedy the relative neglect of the conceptual, historical, and moral foundations of triage. In part I, we first explicate the concept of triage and distinguish it from related concepts. Next, we review the development of various triage systems and plans. We then describe the most common settings in which triage is practiced. In part II, we offer a moral analysis of different triage systems, examining their underlying values and principles.

WHAT IS TRIAGE?

"Triage," "rationing," and "allocation" are terms commonly used to refer to the distribution of medical resources to patients. Although these terms are sometimes used interchangeably, there are clear differences among them. The broadest of the 3, allocation, describes the distribution of both medical and nonmedical resources and does not necessarily imply that the resource being distributed is scarce. For example, a host may allocate seats to the guests at a dinner party.

Rationing also refers to resource distribution but implies that the available resources are not sufficient to satisfy all needs or wants. It also implies that some system or method is being used to guide this distribution, such as the card systems used to ration gasoline and food in the United States during World War II.

The term “triaje” is the narrowest in scope. Derived from the French word trier, to sort, it was originally used to describe the sorting of agricultural products. 1 “Triage” is now used almost exclusively in specific health care contexts. Though “triaje” may be used in an extended sense to refer to any decision about allocation of a scarce medical resource, we believe that use of the term in its primary sense (which we will use in this article) requires that 3 conditions be satisfied:

1. At least a modest scarcity of health care resources exists. The degree of scarcity can vary considerably, from modest, as in a hospital ED where not every patient who presents for care can be served immediately, to dire, as after a catastrophic disaster in which hundreds or thousands of people may experience severe injuries in a short time. Thus, in circumstances in which resources are sufficient to address all patients’ needs without delay, no triage is necessary. At the other extreme, if there are no health care resources available, triage is pointless (Table).
2. A health care worker (often called a “triage officer”) assesses each patient’s medical needs, usually based on a brief examination. This assessment distinguishes the practice of triage, in which microallocation decisions are made about specific individuals according to face-to-face encounters, from the process of macroallocation, such as decisions made by legislators or administrators when allocating health care funds or other resources to different population groups.

3. The triage officer uses an established system or plan, usually based on an algorithm or a set of criteria, to determine a specific treatment or treatment priority for each patient. This condition distinguishes triage from purely ad hoc or arbitrary decisions about distribution of health care resources.

The third condition suggests an important distinction between the concepts of triage and triage planning. If a triage officer makes use of an established plan, some person or group must have developed the plan, and someone must have chosen to use that plan for making specific triage decisions in that situation. Triage planning involves developing and adopting a system or plan to prioritize patient treatment in particular contexts. The level of social order that exists determines, in part, the type of triage plan that can be implemented (Table).

**HISTORY OF TRIAGE**

The practice of triage arose from the exigencies of war, and it remains closely associated with military medicine. The earliest documented systems designed to distribute health care systematically among wounded and sick warriors date back only to the 18th century. Ancient and medieval armies made little or no formal effort to provide medical care for their soldiers, and the care provided was likely to be ineffective. Injured soldiers usually relied on their comrades for aid, and most died of their wounds. Beginning in the 18th century, military surgeons developed and implemented the first battlefield triage rules in the West; little is known about triage elsewhere.

---

**Table. Continuum of triage scenarios: most resources, most social order, to fewest resources, chaos.**

<table>
<thead>
<tr>
<th>Circumstances</th>
<th>ED “Daily”</th>
<th>ICU</th>
<th>Multicasealty Incident</th>
<th>Battlefield</th>
<th>Disaster, Localized</th>
<th>Disaster, Widespread (eg, Weapons of Mass Destruction)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resources available</strong></td>
<td>Relatively plentiful</td>
<td>Relatively plentiful</td>
<td>Good locally; usually transport patients to hospital with plentiful resources</td>
<td>Fair locally; usually transport to plentiful resources</td>
<td>Sparse initially; increases to plentiful over short to medium time period</td>
<td>Sparse for prolonged time</td>
</tr>
<tr>
<td><strong>Social order</strong></td>
<td>Intact</td>
<td>Intact; possible local confusion</td>
<td>Variable; military command structure often intact</td>
<td>Variable; military command structure often intact</td>
<td>Temporary, localized diminished social order</td>
<td>Chaos, possibly for a long period</td>
</tr>
<tr>
<td><strong>Resource-to-patient ratio</strong></td>
<td>High for sickest patients; high to moderate for others</td>
<td>High</td>
<td>Low to moderate; high at hospital</td>
<td>Low on battlefield; higher at treatment facility</td>
<td>Initially low to moderate; later high</td>
<td>Extremely low</td>
</tr>
<tr>
<td><strong>Patient arrival pattern</strong></td>
<td>Linear</td>
<td>Linear</td>
<td>Grouped</td>
<td>Linear or grouped</td>
<td>Grouped, then linear</td>
<td>Linear</td>
</tr>
<tr>
<td><strong>Triage method(s)</strong></td>
<td>Sickest treated first (and sometimes least sick if they are triaged to urgent care clinic); then patients treated on a first-come, first-served basis</td>
<td>Variable</td>
<td>Best possible outcomes first (or selected for most rapid transport method); rarely use “expectant” category</td>
<td>Modern armies: best possible outcomes first (or selected for most rapid transport method); expectant category used when relative limitations of resources exist; guerilla/third-world armies: those able to return to battle first; expectant category used</td>
<td>Best possible outcomes first; “expectant” category used at least until additional resources obtained and social order restored</td>
<td>None; minimal treatment, sporadically</td>
</tr>
</tbody>
</table>
Most scholars attribute the first formal battlefield triage system to the distinguished French military surgeon Baron Dominique-Jean Larrey, chief surgeon of Napoleon’s Imperial Guard. Larrey recognized a need to evaluate and categorize wounded soldiers promptly during a battle. His system was to treat and evacuate those requiring the most urgent medical attention, rather than waiting hours or days for the battle to end before treating patients, as had been done in previous wars. Acting on this recognition, Larrey performed hundreds of amputations on the battlefield while the battle was still raging; he also designed light carriages, which he called “flying ambulances,” to rapidly transport the wounded. In his memoirs on the Russian campaign (1812), Larrey articulated a clear rule for sorting patients for treatment: “Those who are dangerously wounded should receive the first attention, without regard to rank or distinction. They who are injured in a less degree may wait until their brethren in arms, who are badly mutilated, have been operated on and dressed, otherwise the latter would not survive many hours; rarely, until the succeeding day.”

Commentators credit British naval surgeon John Wilson with the next major contribution to military triage. In 1846, Wilson argued that, to make their efforts most effective, surgeons should focus on those patients who need immediate treatment and for whom treatment is likely to be successful, deferring treatment for those whose wounds are less severe and those whose wounds are probably fatal with or without immediate intervention.

The US Army was slow to implement triage systems. In the early days of the Civil War, for example, the medical services were understaffed and poorly organized, and there was no uniform method of sorting casualties. Working as a “wound dresser” for Union troops, poet Walt Whitman described the order of treatment as follows: “The men, whatever their condition, lie there, and patiently wait till their turn comes to be taken up.” Whitman’s description indicates that the guiding principle was “first come, first served.” This method does establish treatment priority, but it does not take into account relative urgency, patient salvageability, or effective use of available resources. After a disastrous first year, the Union Medical Corps greatly decreased mortality by combining triage procedures with front-line medical care and ambulance services. Much of the credit for this goes to Jonathan Letterman, medical director of the Army of the Potomac from 1862 to 1864.

Military surgeons continually refined their triage protocols, widely using the term “triage” for the first time during World War I. The introduction in World War I of deadly new weapons, including machine guns and poison gases, created an unprecedented number of potentially treatable mass casualties requiring triage. This description of a triage situation from a World War I–era military surgical manual offers a slightly different approach to prioritization for treatment from that of Larrey or Wilson:

“A hospital with 300 or 400 beds may suddenly be overwhelmed by 1000 or more cases. It is often, therefore, physically impossible to give speedy and thorough treatment to all. A single case, even if it urgently requires attention—if this will absorb a long time—may have to wait, for in that same time a dozen others, almost equally exigent, but requiring less time, might be cared for. The greatest good of the greatest number must be the rule.”

The approach proposed in this manual clearly differs from Larrey’s dictum that priority goes to the most seriously injured. It also goes beyond Wilson’s proposal that the hopelessly injured not be treated. It asserts that a critical and treatable patient should not be given priority for treatment if the time required to provide that treatment would prevent treatment for other patients with critical but less complicated injuries. This approach explicitly recognizes that, when resources are limited, some patients who could be saved may be allowed to die to save others.

Other World War I triage planners offered a quite different approach to battlefield triage; rather than deferring treatment of the less severely wounded, some suggested giving priority to this group because they could be treated quickly and returned to combat duty. One medical handbook cited by Winslow listed the 2 objectives of triage as “1st, conservation of manpower; 2nd, the conservation of the interest of the sick and wounded.”

World War II saw the introduction of additional weapons, including improved tanks and air support, and of new treatments, including plasma and penicillin. Military physicians developed new, more detailed protocols to assess and triage patients. Beecher recounts a well-known example of a controversial World War II decision about allocation of the extremely limited supply of penicillin. When the first shipment of penicillin arrived in North Africa in 1943, US military physicians decided to use it to treat and return to duty soldiers with gonorrhea rather than soldiers with infected war wounds. Similarly, German military physicians, in the Russian campaign of 1941, used the principle of maximizing the fighting strength by treating those who could most quickly be returned to action with the least expenditure of time and resources. Another example of this approach to military triage can be found in a 1958 North Atlantic Treaty Organization military handbook that describes 3 triage categories: (1) those who are slightly injured and can return to service, (2) those who are more seriously injured and in need of immediate resuscitation or surgery, and (3) the “hopelessly wounded” or dead on arrival.

Today, primarily covert, guerilla, and developing world armed forces lack the resources to treat severely injured combatants. Scarcity of medical resources has become much less likely in modern armed forces that can quickly evacuate large numbers of critically wounded combatants from the battlefield to fully equipped, high-level medical facilities that are able to treat all casualties under most circumstances. Rapid evacuation of the wounded began with basic aeromedical transport (without in-air medical care) in the Korean War and progressed to sophisticated multicasualty helicopter transport with airborne treatment in Vietnam. The average time from injury to definitive care decreased from 12 to 18 hours in World War II, to 2 to 4 hours in Korea, and to
less than 2 hours in Vietnam. In the 2 Iraq conflicts, mobile field hospitals, ideally within 10 miles of the battlefield, kept evacuation times relatively short. In modern military conflicts, triage often is a matter of deciding who should be evacuated to definitive care first, with the dead being evacuated last.

The use of nuclear weapons in World War II and the continuing threat of nuclear, chemical, and biological weapons of mass destruction pose special challenges for triage and triage planning. In a limited attack with weapons of mass destruction, triage planning for major disasters may help providers distribute limited resources among injured survivors. After the widespread use of such weapons or a major natural disaster, however, the number of casualties and the destruction of available resources and of the social order may be so great that effective medical care, including meaningful triage, becomes impossible (Table).

It is often mentioned that military triage systems have been adapted for triage in civilian contexts, including disasters and EDs, but there has been little discussion of the history of triage in these civilian contexts. Based on a comprehensive review of United States disasters, Auf der Heide reported that, despite the existence of triage systems, most disaster casualties do not undergo out-of-hospital triage, because victims are found and transported directly to hospitals by bystanders. It was not until 1964 that Weinerman et al published the first systematic description of civilian EDs’ use of triage. Individual institutions, local and regional emergency medical systems, and federal agencies have subsequently developed and refined triage systems for most ED and disaster situations. The following section briefly describes several of these systems.

TRIAGE: TYPES AND SYSTEMS

As noted above, triage in its primary sense is the sorting of patients for treatment in situations of at least modest resource scarcity, according to an assessment of the patient’s medical condition and the application of an established sorting system or plan. Defined in this way, the most common types of triage include ED triage, inpatient (ICU) triage, incident (multicasualty) triage, military (battlefield) triage, and disaster (mass casualty) triage.

Although each of these types of triage has distinctive elements, all of them satisfy the 3 basic conditions for triage described above, and some have additional features in common. One can, in fact, represent the types of triage as points on a continuum from relatively resource-rich situations in a stable social environment, as in EDs, to the almost total lack of resources and social chaos experienced during or after severe widespread disasters. This continuum is based on the ratio of resources to the number of patients who must be evaluated and treated simultaneously. EDs have the highest resource-to-patient ratio, and large-scale weapons of mass destruction incidents have the lowest, although these ratios often change as a situation progresses (Table).

ED Triage

In modern US EDs, triage officers, usually nurses, routinely assess all patients who present for treatment to sort and prioritize them. ED triage systems are typically designed to identify the most urgent (or potentially most serious) cases to ensure that they receive priority treatment, followed by the less urgent cases on a first-come, first-served basis. In routine ED triage, resources are available to treat every patient, although those who are less severely ill or injured must wait longer. Some patients choose to leave the ED rather than continue waiting for treatment. Some ED triage systems are designed to identify patients with very minor problems and refer them for treatment at clinics or by their own physicians. Commentators have criticized this practice as both morally and medically perilous.

For routine on-site triage, EDs in the United States generally use a 3-level system, although 5-level systems are gaining acceptance as they prove themselves to be more reliable. Other countries, such as Canada, Spain, the United Kingdom, and Australia, have already adopted 5-level systems for ED use. Several methods of 5-level triage are in use. The Emergency Severity Index, developed in the United States, designates the most acutely ill patients as level 1 (highest level) and 2 and uses the number of resources a patient needs to determine levels 3 to 5 (lowest level). The Manchester Triage Scale, used widely in Great Britain, uses 52 algorithms based on the patient’s chief complaint to determine the triage level. The Canadian Triage and Acuity Scale uses an extensive list of clinical descriptors to place patients in one of 5 triage levels. Each level has an associated time required for physician assessment, with all level 1 patients needing to be treated immediately. These methods have good, but not excellent, intrarater reliability, making it unclear whether these are flawed systems, whether those using them are not up to the task, or whether other-than-medical criteria are influencing some decisions.

Inpatient (ICU) Triage

When a patient requires hospitalization, additional decisions must be made about what level of hospital care the patient should receive. In the optimal situation with abundant hospital resources, the patient can immediately receive any and all services that reason suggests may be beneficial. In the more common situation of relative scarcity of at least some hospital-based resources, decisions must be made about who will receive priority access to those services. If these decisions are based on assessment of the patient’s condition and are made according to some system or plan, they are triage decisions. The most common inpatient triage decisions in US hospitals involve access to intensive care. In theory, these decisions allocate ICU beds to those who can most benefit from this level of treatment. In less affluent nations with limited hospital services, inpatient triage decisions are routinely made about priority access to surgery and diagnostic imaging, as well as intensive care.
Incident (Multicase) Triage

This type of triage is designed to respond to an incident that creates multiple casualties, as, for example, a multiple-vehicle collision, a major residential fire, or a commercial airliner crash. In such events, many injured patients, including some with severe injuries, place significant stress on, but typically do not overwhelm, a local emergency medical system. Emergency caregivers at the scene and in the ED triage patients to identify the most critically injured for priority transportation and treatment. Although some on-scene confusion may occur, social stability is not an issue. Additional physicians and other medical and support personnel may be called to help treat the large number of patients with urgent needs, and those with minor injuries and illnesses (the “background noise” of the ED) must wait longer than usual for treatment, but all can eventually receive care.

Military (Battlefield) Triage

As noted, military physicians were the first to implement formal systems of triage to determine treatment priorities for wounded soldiers. Military triage has several distinctive features. The triage officers and treating professionals are typically members of a military service, and the patients are usually, but not always, also military personnel. As military personnel, these health care professionals and patients may have obligations, allegiances, and expectations that are not shared by other health care professionals or by the general public. For example, military personnel typically give up certain rights and liberties and assume an obligation to obey their superior officers’ orders. Military personnel may also be willing to accept life-threatening assignments according to, in part, the expectation that they will receive optimal medical care if they are injured in the line of duty. Furthermore, in addition to the internal medical objective to act in the patient’s best interest, external objectives related to accomplishing a strategic or military mission may influence military triage systems. These systems may, for example, define which patients they may treat, such as combatants and civilians injured by their actions, and whom they may not, typically all other civilians. Finally, international laws, such as the Geneva Conventions about treatment of the wounded in war, define legitimate and illegitimate practices when different categories of wounded soldiers and civilians are treated.

Disaster (Mass Casualty) Triage

In its policy titled “Disaster Medical Services,” the American College of Emergency Physicians offers the following description of a medical disaster: “A medical disaster occurs when the destructive effects of natural or man-made forces overwhelm the ability of a given area or community to meet the demand for health care.” As this description suggests, disaster triage can be roughly distinguished from incident triage by the trigger event’s magnitude of destruction. Because a medical disaster creates demands that overwhelm the capacity of the local health care system, at least some demands cannot be satisfied, and triage can be used to determine who will receive treatment and who will not. Criteria used for triage after natural or manmade disasters may vary, depending on the anticipated number of casualties and the severity of their injuries, the geographic area involved, and the expected arrival time of additional resources. Therefore, to make optimal disaster triage decisions, in addition to rapid patient assessment skills and knowledge of triage systems, triage officers also need accurate information about the cause and extent of the disaster, as well as the location, capabilities, and functional status of nearby health care facilities.

The World Medical Association has recommended that clinicians categorize disaster victims with a system that has been adopted worldwide in some form and which involves the following triage criteria:

a) Those who can be saved but whose lives are in immediate danger, requiring treatment immediately or within a few hours (red triage tag: “immediate”; priority 1)
b) Those whose lives are not in immediate danger but who need urgent but not immediate medical care (yellow triage tag: “delayed”; priority 2)
c) Those requiring only minor treatment (green triage tag: “minimal”; priority 3)
d) Those who are psychologically traumatized and might need reassurance or sedation if acutely disturbed (no specific triage tag)
e) Those whose condition exceeds the available therapeutic resources, who have severe injuries such as irradiation or burns to such an extent and degree that they cannot be saved in the specific circumstances of time and place, or complex surgical cases that oblige the physician to make a choice between them and other patients (black triage tag: “expectant”; no priority)

This last category, “expectant,” which encompasses those who are dead or who are “beyond emergency care,” carries the most emotional and ethical baggage for individuals doing triage. Yet, it is a vital part of disaster triage systems. As the World Medical Association points out, “It is unethical for a physician to persist, at all costs, at maintaining the life of a patient beyond hope, thereby wasting to no avail scarce resources needed elsewhere.”

Alternative categorization methods have been adopted for disaster triage. Among these are Simple Triage and Rapid Treatment (START) and JumpSTART, the more prescriptive and specific methods adopted by disaster medical assistance teams in the United States. Developed at Hoag Hospital in Newport Beach, CA, START is an expedient triage system designed to assist minimally trained first responders to identify the most seriously injured patients and to triage multiple victims in 30 seconds or less, according to primary observations about respiration, perfusion, and mental status. Although it has been field-proven in mass-casualty incidents and in disasters, its ease of use is offset by its high overtriage rate. JumpSTART, a modification of START for pediatric patients, emphasizes the
The presence of respiratory arrest, a common problem in critically injured children.

Supervising editor: Robert K. Knopp, MD

Funding and support: The authors report this study did not receive any outside funding or support.


Presented at the Second Congress of Emergency Medicine, May 2004, Buenos Aires, Argentina.

Address for reprints: Kenneth V. Iserson, MD, University of Arizona, 1501 N. Campbell Avenue, POB 245057, Tucson, AZ 85724; fax 520-626-2480; E-mail kvi@u.arizona.edu.

REFERENCES

**IMAGES IN EMERGENCY MEDICINE**

(continued from p. 272)

**DIAGNOSIS:**

* Cecal volvulus. The CT revealed a markedly dilated right colon with a thickened, irregular wall and a small amount of free fluid and free air (Figures 1 and 2). Cecal volvulus was confirmed during surgery, and ileocectomy was performed. The patient recovered well. Unlike sigmoid volvulus, which occurs more often in elderly patients, incidence of cecal volvulus peaks at age 25 to 35 years. It is associated with hypofixation of the cecum and other parts of the intestine to the posterior abdominal wall, which results in hypermobility, often around the ileocecal artery’s mesenteric pedicle, and can be provoked by neoplasms, inflammation, or previous surgery. Marathon runners seem to have higher rates of cecal volvulus, possibly because of a thin elastic mesentery. The characteristic “coffee bean” finding is not always seen on plain radiograph. Expeditious evaluation is essential because mortality is 10% to 15% if the bowel is viable and up to 40% if the bowel has infarcted. Although successful reduction by barium enema has been reported, there are higher rates of perforation, and the standard of care is almost always operative, with either cecopexy or right-sided colectomy.

**REFERENCES**
