ORIGINAL RESEARCH

Why Is Treatment Urgency Often Overestimated? An Experimental Study on the Phenomenon of Over-triage

Nils Ellebrecht, MA

ABSTRACT

In the 19th century, triage emerged as an administrative concept to overcome the unjust and medically unreasonable consequences of an unsystematic adhoc selection of casualties. Until today, however, triage concepts are often applied incorrectly. High over-triage rates are a well-known phenomenon, which increase mortality rates. In order to examine their frequent occurrences, the article discusses different reasons and presents results of an experimental study. Two triage exercises were conducted: a paper-based triage exercise and a real-world simulation. Both exercises used the same case-vignettes consisting of 5 pairs. Each pair described a patient with the same injury pattern and vital parameters but with differing behaviour (calm/highly excited). Different behavior has a minor but no significant effect on over-triage rates. Over-triage is significantly higher in the real-world simulation than in the paper exercise. This is explained by the characteristics of face-to-face situations themselves: they are more complex and ambiguous, and hold more normative power. Accordingly, over-triage is understood as a means to resolve unclear situations ("better to over- than to under-triage") and to comply with normative demands "within" the strict margins of an administrative concept.

Key Words: triage, societies, multi-agency coordination, leadership, emergency responders

Triage has been a solution to the difficulties inherent in sorting masses of casualties and treating them with the correct priority. From its very beginning, triage has occurred as a well-founded concept that overcomes the dire consequences of an instant and unsystematic medical selection of patients. In his records, 19th century military physician Ernst von Bergmann described the challenges that triage promises to solve: "Who screams the most and is still able to help himself, is transported first; however, in most cases, this is only a slightly wounded person. For whom the bullet penetrated his lung, he is losing his voice with his breath" (Buchholtz et al. p. 269).[1] Because slightly injured soldiers were still able to bring attention to their medical needs, they received care before their severely injured, dying comrades. From a modern perspective, this type of "first come, first served" sorting was a medically unreasonable and ethically unjust selection of patients. Therefore, the Russian military surgeon Pirogov was convinced "that only the well-shaped administration on the battle field as well as on the dressing station is much more important and beneficial than the pure medical activity" (Pirogov p. 37).[2] With the triage, an organizational regime was deployed on the dressing station, discriminating rigorously between medical needs.

In the 19th century, the organizational efforts of triage manifested in category systems distinguishing among different treatment urgencies and sorting casualties into tiers of exigency. Since the 1980s, different triage algorithms have been developed to support the process of correct sorting. Yet, there is still evidence of imperfect categorization, particularly of over-triage.[3,4] Over-triage means labeling noncritical casualties as critical or, in a broader definition, to assign casualties a higher treatment priority than is justified. Because over-triage decreases the chance of a truly critical casualty to quickly receive and benefit from the necessary treatment, it increases the mortality rate of the affected casualty collective as a whole.[5-7] Until today, it is not clear whether and to what degree over- (and under-) triage results from deficient triage concepts, incorrectly applied concepts, or decisions consciously ignoring formal guidelines. So far, research has nearly exclusively focused on the difficulties of hospital (emergency department) triage and triage training[8]; mis-triages in the preclinical field are scarcely examined.

Table 1 lists some possible reasons for over-triage, although they might not in reality be separated sharply but mixed and may even amplify each other. The

TABLE 1

Over-triage: Possible Reasons

Deficient triage concept
- Too complicated to apply correctly
- Imprecise
- Only suitable for trauma patients

Incorrect application of concepts
- Careless mistake
- Incorrect measuring / calculation of vital parameters

Deliberate deviation
- The applied triage concept is considered as wrong (for a particular case)
- Sympathy with the patient / patient is known
- In case of uncertainty: "over-triage is better than under-triage"
- Normative expectation to help: each over-triage does "a little bit more" for the patient (without leaving the formal frame of triage)

current state of research on the reasons for over- and under-triage has not developed beyond the level of assumptions.[9] Evidently, triage seems to be conducted more correctly when simplified concepts are used.[10] What we also know from a German questionnaire is that, during triage particularly, young EMS physicians show difficulties refraining from administering cardiopulmonary resuscitation to a patient who might benefit from resuscitations efforts.[11] Although this general discomfort to not apply therapeutic measures (except head-tilt-chin-lift and managing bleedings), at least while triage is still ongoing, is not directly connected to the phenomenon of over-triage. It rather shows that deviations from triage concepts might be performed deliberately. In this study we investigate, whether the presence and behavior of a patient affect triage decisions and can explain over-triage to some extent.

METHODS

Concerning the aforementioned historical background of triage as an organizational regime overcoming a treatment order based on mere situational perception, we designed an experimental triage exercise to study to what extent the presence of a patient as well as his/her expressed behavior affect triage results. In short, the test design was informed by 2 hypotheses:

H1 The *more expressive* (or insistent) the behavior of a patient, the more frequently s/he is over-triaged;

H2, During triage the *very presence* of patients influences classification to the effect that they are over-triaged more frequently.

To test both hypotheses, 5 case vignettes of noncritical casualties were created. We paid attention that the triage category (TC) of each casualty was *technically* very clear. We then duplicated each case vignette. Each pair kept the vital parameters (pulse, breath, capillary refill time, pain score [numerical rating scale {NRS}-11], appearance of injuries). However, the behavior of both patients differed substantially. While the first case vignette described the patient's behavior as highly excited, the second one characterized the patient's expression as very calm (see example in Table 2).

TABLE 2

Example of a Case Vignette Pair

Case Vignette 5a	**Case Vignette 5b**
Impalement injury, TC 2 ("yellow tag"), highly excited expression	*Impalement injury, TC 2 ("yellow tag"), calm expression*
You are approaching an injured woman lying on the ground and holding her lower right leg. Her feet show several bloody cuts. An iron item sticks in her lower leg. The person is looking up with a face contorted with pain (NRS = 7) and calls on you to help her. She shouts at you: "The thing is sticking right in me! Take the thing out!" Pulse is 120 beats/min; breathing rate is 29/min. Capillary refill happens under 2 s. While taking vital parameters, you notice the person is shaking.	In front of you lies a woman staring at you the whole time. She has an impalement injury to her lower left leg. Apart from an item sticking in her leg, she shows several cuts on her feet and ankles. The injured person has a pulse of 120 beats/min and a respiratory rate of 29/min. The person groans quietly (NRS = 7), tries to stand up but fails because it causes too much pain. Being asked what has happened, she answers quietly and with a voice distorted by pain: "I don't remember. Everything happened so fast." The nail bed test showed a capillary refill time under 2 s.

The 2×5 case vignettes were used in 2 different triage exercises conducted by 16 and 20 German *Rettungsassistenten* (~ paramedics), respectively. In the first exercise, the paramedics read the scripted form of each case vignette in a seminar room, whereas in the second exercise, 10 paid professional actors assumed the patient role described by the case vignettes (Figure 1). All paramedics worked in the Frankfurt area and were recruited by means of e-mail distribution lists and posters looking for paramedics interested in participating in a "scientific emergency exercise." The study took place in the training center of Frankfurt's fire brigade. The ethical board of the University of Freiburg, Germany, gave the study its ethical approval.

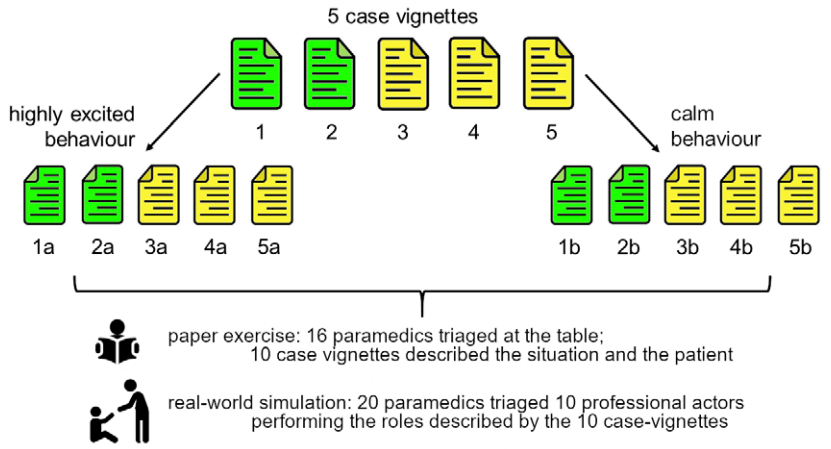
Each paramedic was handed the PRIOR (Primary Ranking for Initial Orientation in Rescue) triage algorithm to determine the TC of each casualty. The PRIOR algorithm differentiates between 3 TCs, where TC I (red tag) demands the "fastest possible treatment," TC II (yellow tag) requires a "fast treatment," and TC III (green tag) determines a delayed treatment[12] (see Supplementary Materials, which are available online). Three of 5 case vignettes described casualties with a TC II ("yellow"), the other 2 casualties were clearly assigned to TC III ("green"). As it is not possible to over-triage a patient with a TC I ("red"); therefore, no case vignette described such a casualty.

All triage decisions were compared with predetermined TCs. If a chosen TC differed by 1 (or 2) categories from the predetermined TC, the deviation was counted as 1 (or 2). Under-triages

https://doi.org/10.1017/dmp.2019.74 Published online by Cambridge University Press

FIGURE 1

Case Vignettes and Triage Exercises (Icons Made by Freepik From www.flaticon.com).



were noted as negative deviation (−1). SPSS 35 (© IBM 2017) and Excel (© Microsoft 2016) were used for statistical evaluation.

To evaluate hypothesis 1, we compared the mean deviation from the correct TC for each calm patient and his/her "excited" duplicate. As it was possible to compare each TC of a calm casualty with the TC of its "excited" duplicate triaged by the same paramedic, the dependent sample t-test could be used to evaluate statistically whether both patient groups were categorized differently. In a sub-analysis, this was also done for each type of exercise (paper/real).

To validate hypotheses 2, the mean deviation of all triage results from the paper exercise was compared statistically with the mean deviation of all triage results from the real-world simulation. Due to the independence of these data sets (both exercises were done by different paramedics), an unpaired (2-sample) t-test was used for this calculation. Finally, we calculated the statistical power of the study.

RESULTS

In total, 359 triages were conducted, 126 (35.1 %) of which were incorrect. A total of 123 (34.3%) mis-triages were too high (over-triage).

As Table 3 shows, differences in patients' presentations lead to different results. However, they are not statistically significant. On average, a highly excited expression leads to a slightly higher over-triage (mean = 0.51 ± 0.61 TC) than a calm expression (0.45 ± 0.61). This difference is a bit stronger in the real-world scenario (0.64 ± 0.64 vs 0.57 ± 0.62) than in the paper exercise (0.35 ± 0.53 vs 0.30 ± 0.56). Thus, the largest discrepancy is found between calm casualties in the paper exercise (0.30 ± 0.56) and patients in a highly excited state in the simulation (0.64 ± 0.64), where the error rate is more than twice as high.

The evaluation shows a significantly higher over-triage in the real-world setting (0.61 ± 0.63 TCs) compared with the paper exercise (0.33 ± 0.55 TCs), with $t(357) = -4.28$, $P = 0.001$ (Table 4). In other words, in the real-world setting on average nearly 2 of 3 casualties were assigned to a TC one level too high, whereas in the paper exercise a little more than 1 of 3 casualties had been assigned 1 level too high. Thus, over-triage was nearly twice as high in the real-world scenario compared with the paper-based exercise.

With Cohen's $d = 0.47$ (effect size) and a power = 0.9, it would need 79 triage decisions per exercise (paper/real) to achieve a significant result with a 1-sided unpaired t-test ($\alpha = 0.05$). With 159 and 200 triage decisions, the statistical power of this statement (H2 = true) is very high.

DISCUSSION

With greater distance to the patient and lower patient expressiveness, paramedics sort casualties more correctly. Direct encounters with patients, no matter whether they express their situation very excitedly or not, affect triage decisions and result in significantly higher over-triage rates. In the field, triage concepts cannot entirely assert their logic and "coldness" (or impartiality) against the demands for help raised by the appearance of injured bodies and verbal and nonverbal behavior. The question is why? High over-triage rates should not be belittled as a product of pure compassion. Two further explanations can be given (cp. Table 1).

https://doi.org/10.1017/dmp.2019.74 Published online by Cambridge University Press

TABLE 3

Paired Samples Statistics and Paired Samples Test for Case Vignette Pairs (Calm/Expressive)

		Mean	N	Std. Deviation	Std. Error Mean
Both studies	Calm	0.45	179	0.61	0.05
	Expressive	0.51	179	0.61	0.05
Paper-exercise	Calm	0.30	79	0.56	0.06
	Expressive	0.35	79	0.53	0.06
Real-world simulation	Calm	0.57	100	0.62	0.06
	Expressive	0.64	100	0.64	0.06

		Paired Differences							
					95% CI of the Difference				
		Mean	SD	SEM	Upper	Lower	T	df	Sig. (2-tailed)
Both studies	Calm - Expressive	−0.061	0.72	0.05	−0.17	0.04	−1.142	178	0.255
Paper-exercise	Calm - Expressive	−0.051	0.64	0.07	−0.19	0.09	−0.705	78	0.483
Real-world simulation	Calm - Expressive	−0.070	0.78	0.08	−0.23	0.09	−0.895	99	0.373

TABLE 4

Group Statistics and Independent Sample Test for Both Studies (Paper Exercise/Real-World Simulation)

		N	Mean	SD	SEM
Study	Paper-exercise	159	0.33	0.55	0.04
	Real-world simulation	200	0.61	0.63	0.04

		Levene's test for Equality of Variances		*t*-Test for Equality of Means						
									95% CI of the Difference	
		F	Sig.	T	df	Sig. (2-Tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Study	Equal variances assumed	11.238	0.001	−4.285	357	0.000	−0.27	0.06	−0.40	−0.15
	Equal variances not assumed			−4.356	354	0.000	−0.27	0.06	−0.39	−0.15

First, high over-triage rates can partly be explained by a tacit action pattern that can be named "better to over- than to under-triage." Assumedly, this work concept is always followed in situations of uncertainty. If the category is unclear (for whatever reasons), the paramedic will always opt for a higher treatment urgency. Consequences of a false under-triage are seen as being much more detrimental.

Second, and from a more sociological standpoint, over-triage can be understood as fulfilling the medical commitment for patients signaling and insisting on their need for help. Already through their mere presence normative expectations are built up. With each over-triage, the medical professional does a little bit "more" for her/his client than is allowed during triage. In doing so, the paramedic still performs a triage without actually performing a therapeutic act. However, simultaneously s/he meets the expectation to do more by classifying the patient too high. Thus, both situational and organizational demands are satisfied. The abstract consequences for the collective of patients, ie, a higher mortality rate, are neglected in the concrete situation.

One can argue that the significant discrepancy between both exercises regarding the mean over-triage rate can, at least partly, also be traced back to the higher complexity and ambiguity of a face-to-face situation compared with a shallow paper exercise. Similar to studies comparing the application of text-based scenarios with multimedia-based scenarios in medical education,[13,14] the higher error rate of the real-world simulation in our study can be understood as a result of a lack of clarity. This point is further strengthened by the argument that PRIOR, because of its indistinct criteria, is said to encourage over-triage.[15] However, if we consider that situational complexity and ambiguity would have been expected to also

https://doi.org/10.1017/dmp.2019.74 Published online by Cambridge University Press

increase the number of under-triages, but at least in this study did not, this argument can be interpreted in favor of the first explanation. If lacking clarity is predominantly resolved by decisions rigorously estimating the treatment urgency of a patient too high, the medical commitment to the *individual* patient is taken into account too much, although abstaining from it is advised for the time of triage.

Which changes can lead to a stricter application of triage concepts? At this point, one would usually call for a better qualification or further training of paramedics: this study actually shows how paramedics might experience the difficulties of triage by means of 2 different exercises and learn from it. Additionally, paramedics could be better sensitized to the abstract goal of triage by highlighting the negative effect of over-triage on the mortality rate of the whole group. However, as organizational sociologists, we want to emphasize another option: The presented triage study was accompanied by a tele-medical project, which was able to show that triage algorithms were applied far more correctly if a telemedical physician was connected to the paramedic on the scene by means of radio and camera. Together they were able to triage more strictly, with the doctor committing the paramedic to exactly follow the algorithm. For the remote doctor in his/her office, it seemed to be much easier to comply with the formal guideline (on the screen). This disciplining effect may to some extent also occur if 2 paramedics conduct triage together on the scene. However, a "team triage" might be an unrealistic working mode in a true mass casualty situation. Therefore, the only promising external help might be a triage algorithm that is so clear and simple, that any deviation is made extremely difficult, just because of the fact that any deviation becomes blatantly obvious to the paramedic him-/herself. Further studies should compare to what extent different algorithms amplify/mitigate the medical propensity to over-triage.

About the Authors

Centre for Security and Society, Albert-Ludwigs-University, Freiburg, Germany (Mr Ellebrecht).

Correspondence and reprint requests to Nils Ellebrecht, Centre for Security and Society, Albert-Ludwigs-University, Freiburg, Germany (e-mail: nils.ellebrecht@soziologie.uni-freiburg.de)

Funding

The study was funded by the German Ministry of Education and Research (BMBF) as part of the research project AUDIME (funding number 13N13262).

Supplementary Material

To view supplementary material for this article, please visit https://doi.org/10.1017/dmp.2019.74

REFERENCES

1. Buchholtz A, Bergmann Ev. Ernst von Bergmann. *Mit Bergmanns Kriegsbriefen von 1866, 1870 / 71 und 1877*. Leipzig: Vogel; 1911.
2. Pirogov N. *Grundzüge der allgemeinen Kriegschirurgie. Nach Reminiscenzen aus den Kriegen in der Krim und dem Kaukasus und aus der Hospitalpraxis*. Leipzig: F.C.W. Vogel; 1864.
3. Kahn CA, Schultz CH, Miller KT, et al. Does START triage work? An outcomes assessment after a disaster. *Ann Emerg Med*. 2009;54:424-430.
4. Kilner T. Triage decisions of prehospital emergency health care providers, using a multiple casualty scenario paper exercise. *Emerg Med J*. 2002;19:348-353.
5. Hupert N, Hollingsworth E, Xiong W. Is overtriage associated with increased mortality? Insights from a simulation model of mass casualty trauma care. *Disaster Med Public Health Prep*. 2007;1:S14-S24.
6. Armstrong JH, Hammond J, Hirshberg A, et al. Is overtriage associated with increased mortality? The evidence says "yes". *Disaster Med Public Health Prep*. 2008;2:4-5.
7. Frykberg ER. Medical management of disasters and mass casualties from terrorist bombings: how can we cope? *J Trauma*. 2002;53:201-212.
8. Lentz BA, Jenson A, Hinson JS, et al. Validity of ED: addressing heterogeneous definitions of over-triage and under-triage. *Am J Emerg Med*. 2017;35:1023-1025.
9. Kahn CA, Schultz CH, Miller KT, et al. In reply. *Ann Emerg Med*. 2010;55:580-581.
10. Lehmann RK, Arthurs ZM, Cuadrado DG, et al. Trauma team activation: simplified criteria safely reduces overtriage. *Am J Surg*. 2007;193:630-634; discussion 634-635. doi: 10.1016/j.amjsurg.2007.01.017
11. Ellebrecht N. Die Realität der Sichtung: Ergebnisse einer Befragung zur Sichtungsausbildung und MANV-Erfahrung von Notärzten und Rettungsassistenten. *Notf Rett Med*. 2013;16:369-376.
12. Bubser F, Callies A, Schreiber J, et al. PRIOR: Vorsichtungssystem für Rettungsassistenten und Notfallsanitäter. *Rettungsdienst*. 2014;37:730-734.
13. Shen L, Li F, Wattleworth R, et al. The promise and challenge of including multimedia items in medical licensure examinations: some insights from an empirical trial. *Acad Med*. 2010;85:S56-S59.
14. Chang TP, Schrager SM, Rake AJ, et al. The effect of multimedia replacing text in resident clinical decision-making assessment. *Adv Health Sci Educ Theory Pract*. 2017;22:901-914.
15. Heller RA, Salvador N, Frank M, et al. Diagnostische Güte von Vorsichtungsalgorithmen für den Massenanfall von Verletzten und Erkrankten. *Anaesthesist*. 2017;66:762-772.

https://doi.org/10.1017/dmp.2019.74 Published online by Cambridge University Press