Verheul, M.L.M.I., Dückers, M.L.A., Visser, B.B., Beerens, R.J.J., Bierens, J.J.L.M. Disaster exercises to prepare hospitals for Mass Casualty Incidents. Does it contribute to preparedness or is it ritualism? Prehospital and Disaster Medicine: 2018, 33(4), 287-393



# Disaster-exercises to prepare hospitals for Mass Casualty Incidents. Does it contribute to preparedness or is it ritualism?

VERHEUL, M.L.M.I., DÜCKERS, M.L.A., VISSER, B.B., BEERENS, R.J.J., BIERENS, J.J.L.M.

**ABSTRACT**

**Introduction**

The central question this study sought to answer was whether the team members of Strategic Crisis Teams (SCTs) participating in mass-casualty incident (MCI) exercises in the Netherlands learn from their participation.

**Methods**

Evaluation reports of exercises that took place at two different times were collected and analyzed against a theoretical model with several dimensions, looking at both the quality of the evaluation methodology (three criteria: objectives described, link between objective and items for improvement, and data-collection method) and the learning effect of the exercise (one criterion: the change in number of items for improvement).

**Results**

Of all 32 evaluation reports, 81% described exercise objectives; 30% of the items for improvement in the reports were linked to these objectives, and 22% of the 32 evaluation reports used a structured template to describe the items for improvement. In six evaluation categories, the number of items for improvement increased between the first (T1) and the last (T2) evaluation report submitted by hospitals. The number of items remained equal for two evaluation categories and decreased in six evaluation categories.

**Conclusion**

The evaluation reports do not support the ideal-typical disaster exercise process. The authors could not establish that team members participating in MCI exercises in the Netherlands learn from their participation. More time and effort must be spent on the development of a validated evaluation system for these simulations, and more research into the role of the evaluator is needed.

## INTRODUCTION

Disasters, human-made or natural, can occur at anytime, anywhere in the world. In 2015, 346 disasters were reported worldwide, with 98,580,793 people affected; 22,773 people lost their lives; and the events resulted in an economic damage of US$ 66.5 billion. [1] Mass-casualty incidents (MCIs) are a specific disaster type. The National Emergency Medical Services Information System (NEMSIS; Salt Lake City, Utah USA) states the following definition of an MCI:

A mass-casualty incident is defined as an event which generates more patients at one time than locally available resources can manage using routine procedures or resulting in a number of victims large enough to disrupt the normal course of emergency and health care services and would require additional non-routine assistance. [2]

In most countries, hospitals are legally required to prepare for disasters, including MCIs. Dutch health care institutions are, under Dutch law, obliged to guarantee a constant level of health care in every situation, including disasters. [3, 4] Since 2008, the Dutch Ministry of Health (The Hague, Netherlands) reserves an annual budget of 11 million Euros to stimulate disaster preparedness activities of the health care system. [5] There are 82 hospitals in the Netherlands situated in 11 trauma regions. In each trauma region, a coordinator is responsible for the annual disaster preparedness budget (one million Euros per year, per region). [5]

An estimated 10% of this budget is attributed to exercises focusing on the role of Strategic Crisis Teams (SCTs) in hospitals during an MCI. At the time of a disaster, the SCT is responsible for the overall coordination. In addition, the SCT has the authority to intervene in the regular hospital processes (eg, cessation of [part of] the elective operations program) to provide an optimized reception of the disaster victims, and it is also responsible for the external communication with networking partners and the public. It is common practice for hospitals to involve a specialized training agency in the organization, execution, and evaluation of the exercise. The evaluation usually results in a written evaluation report. The format of those reports differs greatly: some agencies use a structured format, others use a free narrative. Reports always contain some form of points of improvement, formulated as remarks, tips, or comments.

It is estimated that 10 million Euros have been spent on this type of exercise in the Netherlands since 2008. So far, the effects of this preparedness program have not been evaluated systematically, making it impossible to determine whether participating in these exercises contributes to disaster preparedness. The global idea - an ideal in line with key models in quality management literature - is that organizations will learn from experiences and improve their task performance over time. [6, 7] An exercise is an instrument to train for, assess, practice, and improve performance in prevention, protection, response, and recovery capabilities in a risk-free environment. Exercises can be used for testing and validating policies, plans, procedures, training, equipment, and interagency agreements; clarifying and training personnel in roles and responsibilities; improving interagency coordination and communications; identifying gaps in resources; improving individual performance; and identifying opportunities for improvement. [8] Thus, MCI simulation exercises are a common method to prepare for disaster in hospitals. [6] In an MCI simulation exercise, the hospital simulates the influx of a large number of patients and the hospitals response to this influx. Ideally, the exercise process is shaped according to

the Plan-Do-Study-Act cycle and starts by describing the desired outcomes (exercise objectives), followed by the design of the exercise itself, and its execution. After the exercise, the evaluation, reflected in an evaluation report with items for improvement, allows the organization to be informed about what can be improved. By implementing these suggestions for improvement, the disaster preparedness of the organization should improve. [7, 9–11] Only if the objectives are stated clearly is it possible to assess accurately the degree to which the participants have fulfilled those goals. [12, 13]

Since the first principle of learning is change, learning from these exercises would imply changes in behavior, or at least in the range of potential behaviors. [14, 15]

Literature shows the importance of the evaluation findings being based on appropriate, credible, and reliable information enabling the drawing of any conclusions, and of evaluators making selective use of both qualitative and quantitative data collection tools and strategies. [16, 17] It is therefore relevant to identify what data collection methods are used in the Dutch preparedness program.

Additionally, existing literature seems rather pessimistic about the learning effect of exercises in disaster preparedness. [18, 19] The lack of objective data (such as pre- and post-intervention tests) makes it difficult to prove any effect.

The central question this study sought to answer with the ambition to contribute to the international knowledge base, was:

"Do SCT members participating in MCI exercises in the Netherlands learn from their participation?"

To answer this question, this study firstly explored the quality of the evaluation methodology of evaluation reports used in the Netherlands, adopting the methodology applied by Beerens and Tehler. [20] Secondly, this study assumed that the learning behavior in individuals, teams, or organizations participating in the exercises is accompanied by a decrease in the number of items for improvement in exercise evaluations over time. This decrease is considered an indication for learning behavior and, as such, the change in number of items for improvement over time when comparing the reports from two moments in time was calculated as a proxy for learning.

## METHODS

### Study Design and Data Collection

The authors conducted a retrospective descriptive study of MCI exercise evaluation reports provided by Dutch hospitals. Between June 2016 and October 2016, the 11 coordinators of the Dutch trauma regions received an email requesting to contribute two evaluation reports of past disaster exercises about MCIs in which the SCT participated, from each hospital in their region. A reminder was sent after eight weeks. For the analysis in this paper, reports were excluded if they were from exercises before 2008 (in 2008, the Dutch government started the program to stimulate disaster preparedness in health care); if a hospital submitted only a single report; and if the reports were not about exercises, not about an MCI, or if the reports did not describe the STC. If a hospital submitted more than two reports, the oldest and the newest reports were included, excluding the reports in between.

In order of submission, every hospital was assigned a letter, and the included reports were put in a data extraction form (Microsoft Excel spreadsheet Version 2010;

Microsoft Corporation; Redmond, Washington USA). The first evaluation report of each hospital was called T1, and the second report T2.
Thereafter, evaluation categories were defined using 11 evaluation categories that had been validated in previous studies. [21 – 23] The items for improvement in four randomly selected reports were test-scored by one author. As not all items for improvement could be scored in this test, four additional categories (plans and task cards, collaboration, team composition, and network partners) were defined for the final list of categories used in this study (Table 1).

**[TABLE 1]**

Subsequently, two investigators (MV and BV) independently attributed each of the items for improvement to one of the evaluation categories. Differences in the attributions were discussed until full agreement was reached. The evaluation category "irrelevant, theme specific" (n = 2) had been excluded, as these details were not generic and therefore not relevant for the outcome of this study.

### Analysis

In accordance to the method proposed by Beerens and Tehler, [20] the following data were extracted: hospital reference; the presence of exercise objectives in the reports (per report); the link between the exercise objectives and items for improvement (per remark); and the data collection method (per report), items for improvement, evaluation-category (per remark), and time between reports (in months). The objectives of the reports and items for improvement concerning the SCT from each included report were copied directly from the reports into the extraction form. Data were anonymized. A complete file with the translated contents of the items for improvement can be obtained from the first author. The full reports used for the study cannot be disclosed because they contain information about individuals.

## RESULTS

In total, 98 reports were submitted by 82 hospitals. After applying the exclusion criteria, 32 reports were included (Figure 1). The reports in T1 contained 80 items for improvement, in total, and the reports in T2 described 84 items for improvement. The exclusion of one evaluation category (n = 2; eg, [translated] *"the team composition is meagre"* or *"unclear who coordinates with whom")* led to 79 items for improvement for T1 and 83 items for improvement for T2.

**[FIGURE 1]**

The time between the two reports varied between 10 and 49 months, with a mean of 26.1 months (IQR = 21).
In 26 (81%) reports, exercise objectives were described. Of all items for improvement, 48 (30%) were linked to these objectives, and seven reports (22%) used a structured template to describe the items for improvement compared to 78% of the 32 evaluation reports (n = 25) written as an unstructured narrative. When comparing the quality criteria from T1 and T2, a slight improvement was visible (Table 2).

**[TABLE 2]**

The items for improvement were distributed over the evaluation categories as shown in Table 3.

There was an increase in the number of items for improvement between T1 and T2 for six evaluation categories; two evaluation categories remained equal and six decreased (Figure 2).

**[FIGURE 2]**

An additional finding was that 10 of the 16 sets of reports showed one or more items for improvement formulated in almost the exact same wording in both reports (T1 and T2). For example, [translated] "*make a list of all abbreviations that are applicable for this team*" was found three times. Also, the items for improvement were often formulated vaguely, ungainly, generic, and multi-interpretable (eg, [translated] "*The secretary who stood next to the flip-over wrote everything on one sheet. Therefore, nobody could read what she had written because she was standing in the line of sight,*" or "*Scenario thinking,*" or "*It seems the chairman and the secretary may need some additional training*").

## DISCUSSION

This paper intended to answer the question whether SCT members and their hospital organizations in the Netherlands learn from their participation in MCI exercises. The authors did so by considering the methodology of the evaluation and by assessing the learning effect through calculating the change in the critical number of items for improvement.

Of the evaluations, 78% were written as a free-format narrative, which may have caused biased and incomplete evaluations. In addition, although 81% of the reports state objectives, only 30% of the items for improvement are linked to the objectives. If clearly stated goals or objectives are conditional on an objective assessment of the participants performance, these findings lead to the conclusion that the methodology of evaluation is inadequate compared to the state of the art in literature.

Second, in six out of 14 evaluation categories, the number of items for improvement increased over time. For two categories, the numbers did not change. A decrease would have been expected, given the assumption that a decrease in items for improvement would indicate learning, reflecting changes in behavior. Without a structured framework for the evaluation, it is difficult to interpret increases or decreases in this number, but the authors can hypothesize with Lundberg that evaluators will always have items for improvement if they are asked to look for them. [24] Since no existing literature validates the choice of indicator, further research is needed to validate it.

Third, it was noticed that a large number of items for improvement were so similar in wording that it seemed to be copied directly from previous reports. The similarity in items for improvement might lead to the conclusion that the situation had not changed. This could mean, for instance, that SCTs or hospitals are not learning, or that the authors of the reports simply copy-pasted parts of the old reports to later reports.

Finally, one could argue that the content of the items for improvement lacks professionalism: items for improvement are multi-interpretable and formulated

vaguely, ungainly, and are very generic. This does not meet with the general standard that requires evaluations to consist of reliable, appropriate, and credible information. These findings raise questions about the evaluators themselves. How are they selected and trained? How much do they know about the basic rules of evaluation? Future research is necessary to shed a light on these questions.

This study shows that researchers have been unable to assess if the SCT members participating in MCI exercises in the Netherlands learn from their participation. As such, this study echoes the conclusions of other studies. The systematic literature review performed by Hsu, et al concludes that due to the lack of objective data (eg, the data of hospital responses to actual MCIs are rarely made available to the public), the effectiveness of disaster drills as a tool for hospital disaster preparedness is difficult to determine. One author even states that "simulation-enabling facilitators and designers learn more from the exercise than the intended participants," and in 2009, Thomas Birkland called the evaluation reports "fantasy documents because they are created and disseminated for rhetorical purposes, even if their authors somehow believe that learning has really occurred." [25 , 26]

More importantly, the present study results in evidence-based concern whether the 10 million Euros spent on preparation of SCTs for MCIs in the Netherlands has had any positive effects on disaster preparedness. If exercises would indeed contribute to preparedness, this should be through learning, and learning implies a cyclic approach. There should be areas for improvement indicated after the intervention (the exercise), serving as input for the next intervention. Only in this way, it may be possible that the individual participants, the teams, and the organizations involved can learn.

## LIMITATIONS

This study has several limitations. The final number of reports included was small. At the same time, the authors were able to compare one coherent, uniform group of MCI exercise evaluations.

The reports were written by many different evaluators from various agencies. Differences in personal preferences, knowledge, experience, and intents may also account for differences in items for improvement. This advocates the standardization of the evaluation.

By lack of a validated measure of the learning effect of exercises, the authors chose to calculate the change in number of items for improvement as indicator for that learning effect. However, if the number of items for improvement is not regulated or limited beforehand, it is difficult to contribute a change in that number to a learning effect.

Finally, it is likely that there were significant changes in the participants of the SCT, including trainers and evaluators, between T1 and T2. The authors could not control for the possibility that this might have affected the findings. On the other hand, this type of variation is part of reality, and it at least underscores the questionability of an MCI exercise to prepare a SCT.

## CONCLUSION

Without clear objectives for a disaster exercise, a validated and consistent system of evaluation, and objective tools for measuring the learning effects, it is not possible to evaluate whether MCI simulations contribute to a better preparedness of the SCT of a

hospital. The results of this study, based on the assumption that a reduction in number of items for improvement would reflect a learning effect of disaster simulations, suggest that there is only limited (or even no) objective learning effect for the SCT to handle an MCI. The development of a validated evaluation system for these simulations, in order to facilitate comparison of team performance over time, between teams, and against outside criteria will be the first step to understand the return on investment of disaster exercises. The findings concerning the quality of the evaluation methodology indicate that further research into the role of the evaluator is needed.

**Acknowledgements**

The authors thank all hospitals that contributed to this study by sharing their reports so freely with them. They also thank Willem Verheul, Emma Berkelbach van der Sprenkel, Willy-Anne van Stiphout, and Annemiek van de Weg, Dagmar Droogsma, and all colleagues from the University Medical Center of Utrecht (Utrecht, The Netherlands) who provided them with feedback and supported them in the process of writing.

**REFERENCES**

1. United Nations Office for Disaster Risk Reduction web site. 2015 disasters in numbers. http://www.unisdr.org/files/47804_2015disastertrendsinfographic.pdf. Published 2016. Accessed September 12, 2017
2. National Association of State EMS Officials. Extended Definition Document NEMSIS/NHTSA 2.2.1 Data Dictionary. https://nemsis.org/media/nemsis_v2/documents/Data_Managers_Council_-_Data_Definitions_Project_Final_Ve.pdf. Accessed November 22, 2017
3. Wet toelating zorginstellingen. http://wetten.overheid.nl/BWBR0018906/2016-08-01. Accessed September 12, 2017
4. Kwaliteitswet zorginstellingen. http://wetten.overheid.nl/BWBR0007850/2015-01-01. Accessed September 12, 2017
5. Landelijke convenant Opleiden Trainen Oefenen 2008. https://www.rijksoverheid.nl/documenten/convenanten/2008/10/16/convenant-inzake-opleiden-trainen-en-oefenen-ter-voorbereiding-op-rampen-en-crises. Accessed September 12, 2017
6. Tur-Kaspa, I. Preparing hospitals for toxicological mass-casualties events. Crit Care Med. 1999;27(5):1004-1008. CrossRef |
7. Reinertsen, JL. Physicians as leaders in the improvement of health care systems. Physicians as Leaders in Improving Health Care Series. Berwick DM, Nolan TW, (series eds). Ann Intern Med. 1998;128(10):651-656
8. FEMA - HSEEP Glossary, 2013. https://www.fema.gov/media-library-data/20130726-1914-25045-8890/hseep_apr13_.pdf. Accessed September 12, 2017
9. Taylor, MJ, McNicholas, C, Nicolay, C, Darzi, A, Bell, D, Reed, JE. Systematic review of the application of the plan–do–study–act method to improve quality in healthcare. BMJ Qual Saf. 2014;23(4):290-298. CrossRef
10. Dückers, MLA, Wagner, C, Vos, L, Groenewegen, PP. Understanding organizational development, sustainability, and diffusion of innovations within hospitals participating in a multi-level quality collaborative. Imp Sci. 2011;6(18).
11. Beerens, R, Abraham, P, Glerum, P, Kolen, B. Flood preparedness training and exercises. In: Bierens J, (ed). Handbook on Drowning. Berlin, Germany: Springer; 2013: 1009-1016.
12. St. Leger, AS, Walsworth-Bell, JP. Change Promoting Research for Health Services. Buckingham: Open University Press; 1999.
13. Dausey, DJ, Moore, M. Using exercises to improve public health preparedness in Asia, the Middle East, and Africa. BMC Res Notes. 2014;7(1):474.

14. Huber, GP. Organizational learning: the contributing processes and the literatures. Organ Sci. 1991;2(1):88-115.
15. Alexander, PA, Schallert, DL, Reynolds, RE. What is learning anyway? A topographical perspective considered. Educ Psychol. 2009;44(3):176-192.
16. Stufflebeam, DL, Coryn, CLS. Evaluation Theory, Models, and Applications. 2nd ed. San Francisco, California USA: Jossey-Bass; 2007.
17. Calidoni-lundberg, F. Evaluation: Definitions, Methods, and Models. Östersund, Sweden.
18. Lagadec, P. Learning processes for crisis management in complex organizations. J Contingencies Cris Manag. 1997;5(1):24.
19. Hsu, EB, Jenckes, MW, Catlett, CL, et al. Effectiveness of hospital staff mass-casualty incident training methods: a systematic literature review. Prehosp Disaster Med. 2004;19(3):191-199.
20. Beerens, R, Tehler, H. Scoping the field of disaster exercise evaluation - a literature overview and analysis. Int J Disaster Risk Reduct. 2016;19(May):413-446.
21. Juffermans, J, Bierens, JJLM. Recurrent medical response problems during five recent disasters in the Netherlands. Prehosp Disaster Med. 2010;25(2):127-136.
22. Savoia, E, Preston, J, Biddinger, PD. A consensus process on the use of exercises and after-action reports to assess and improve public health emergency preparedness and response. Prehosp Disaster Med. 2013;28(3):305-308.
23. Savoia, E, Agboola, F, Biddinger, PD. Use of after action reports (AARs) to promote organizational and systems learning in emergency preparedness. Int J Environ Res Public Health. 2012;9(8):2949-2963.
24. Lundberg, J, Rollenhagen, C, Hollnagel, E. What-You-Look-For-Is-What-You-Find - The consequences of underlying accident models in eight accident investigation manuals. Saf Sci. 2009;47(10):1297-1311.
25. Borodzicz, E, van Haperen, K. Individual and group learning in crisis simulations. J Contingencies Cris Manag. 2002;10(3):139-147.
26. Birkland, TA. Disasters, lessons learned, and fantasy documents. J Contingencies Cris Manag. 2009;17(3):146-156.

## TABLES AND FIGURES

Table 1. Evaluation Categories
Abbreviation: SCT, Strategic Crisis Team.

| Category | Short Definition | Examples |
|---|---|---|
| Structured Decision Making | All critical remarks that concern the SCT's technique of decision making. | More structured decision making is needed. Prioritize. |
| Command Structure | All critical remarks that concern the chain of command. | Who is in charge of what? What is the hierarchy? |
| Logging | All critical remarks that concern the way the meeting was logged and all critical remarks concerning secretarial work. | Log all relevant information, share the log, keep logging, even if it becomes hectic. |
| Practical Circumstances | All critical remarks that concern practical matters. | The size of the room, or the lack of ventilation, use name-tags. |
| Information Sharing | All critical remarks that concern the sharing of information within the SCT or between SCT and other teams. | The way the team (member) structures information, whom they share with, whether the information is relevant, gapless, and timely. |
| Network Partners | All critical remarks that concern knowledge of the SCT concerning network partners. | Not knowing the network-partners, not knowing their role and responsibility. |
| Team Composition | All critical remarks that concern who should or should not be on the SCT. | People missing on the team, people on the team that should not be there. |
| Scenario Thinking | All critical remarks that concern the SCT's ability to develop scenarios. | Think ahead, develop scenarios, take time to think of scenarios. |
| Plans and Task Cards | All critical remarks that concern crisis-plans and task-cards. | Adjust the plan, elaborate on the plan, use the plan, make task-cards. |
| Knowledge | All critical remarks that concern the knowledge of the SCT members about crisis management. | Involve experts if you miss knowledge in the team, no knowledge of regional procedures, lack of knowledge of each other's processes in the team. |
| Collaboration | All critical remarks that concern the way the SCT works together during the exercise. | Share feelings of unease in the team, make sure every team member is heard during the meeting. |

Verheul © Prehospital and Disaster Medicine

Table 1. Evaluation Categories *(continued)*

| Category | Short Definition | Examples |
|---|---|---|
| Crisis Communication | All critical remarks that concern crisis-communication: communication about the incident directed at the public. | Inform parents of patients, make sure crisis-communication-manager gets a chance to speak during the meeting, have a press-statement ready. |
| Alerting System | All critical remarks that concern the way the SCT and/ or hospital is alerted. | People missing from the alert-list, the list being outdated, people not knowing who is on the alert-list. |
| Verification of Information | All critical remarks that concern the verification of information. | Team working with assumption. |
| Irrelevant/Theme Specific or Unclear | All critical remarks that concern a specific detail or that the researchers deem irrelevant or unclear in a generic sense. | "Due to building-activities outside, it was hard to hear each other" or "the chairman had a broken leg and his wheelchair didn't fit under the table." |

Figure 1. Selection of Evaluation Reports.
Abbreviations: MCI, mass-casualty incident; SCT, Strategic Crisis Team.

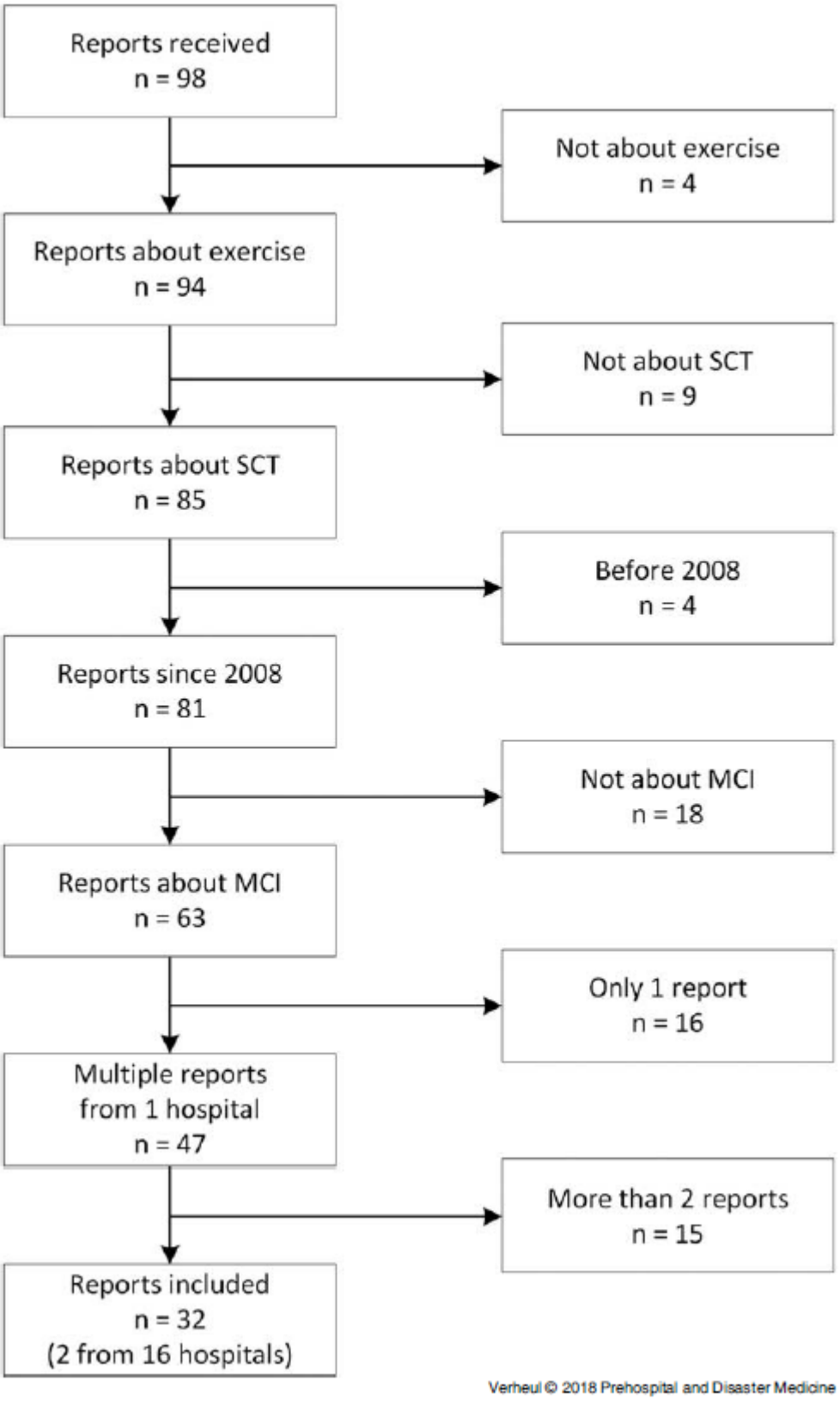

Table 2. Quality of Evaluation Methodology: Overall, at T1, and at T2

| Criterion | n (%) Meet Criterion - Overall | n (%) Meet Criterion at T1 | n (%) Meet Criterion at T2 |
|---|---|---|---|
| Report describes objectives for the exercise n = 32 | 26 (81.0%) | 13 (81.0%) | 13 (81.0%) |
| Evaluators work with structured performance indicators n = 32 (T1 = T2: n = 16) | 7 (21.0%) | 2 (12.5%) | 5 (15.6%) |
| Items of improvement are linked to an objective in the exercise's design n = 162 (T1: n = 79, T2: n = 83) | 48 (29.6%) | 22 (27.8%) | 26 (31.3%) |

Verheul © 2018 Prehospital and Disaster Medicine

Table 3. Number of Items for Improvement per Category, T1 and T2

| Evaluation Category | T1 | T2 |
|---|---|---|
| Structured Decision Making | 21 | 24 |
| Command Structure | 10 | 5 |
| Logging | 9 | 6 |
| Practical Circumstances | 7 | 11 |
| Information Sharing | 3 | 2 |
| Network Partners | 5 | 5 |
| Team Composition | 5 | 6 |
| Scenario Thinking | 2 | 5 |
| Plans and Task Cards | 2 | 7 |
| Knowledge | 5 | 4 |
| Collaboration | 5 | 2 |
| Crisis Communication | 2 | 2 |
| Alerting System | 1 | 3 |
| Verification of Information | 2 | 1 |
| Total | 79 | 83 |

Verheul © 2018 Prehospital and Disaster Medicine

Figure 2. Change in Number of Items for Improvement, Per Category.