Criteria for recommendations after perioperative sentinel events

Bos, K.; Dongelmans, D.A.; Groeneweg, J.; Legemate, D.A.; Leistikow, I.P.; Laan, M.J. van der

BMJ Open Quality

Criteria for recommendations after perioperative sentinel events

Kelly Bos [ID],[1,2] Dave A Dongelmans,[2,3] Jop Groeneweg,[4,5] Dink A Legemate,[1] Ian P Leistikow,[6,7] Maarten J van der Laan [ID] [2,8]

[1]Department of Surgery, Amsterdam UMC location AMC, Amsterdam, the Netherlands
[2]Impulse Institute, Amsterdam, the Netherlands
[3]Department of Intensive Care Medicine, Amsterdam UMC location AMC, Amsterdam, the Netherlands
[4]Delft University of Technology, Delft, the Netherlands
[5]Leiden University, Leiden, the Netherlands
[6]Dutch Health and Youth Care Inspectorate, Utrecht, the Netherlands
[7]Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, the Netherlands
[8]Department of Surgery, University Medical Center Groningen, Groningen, the Netherlands

Correspondence to
Dr Kelly Bos;
k.bos@amsterdamumc.nl

ABSTRACT

Background The recurrence of sentinel events (SEs) is a persistent problem worldwide, despite repeated analyses and recommendations formulated to prevent recurrence. Research suggests this is partly attributable to the quality of the recommendations, and determining if a recommendation will be effective is not yet covered by an adequate guideline. Our objectives were to (1) develop and validate criteria for high-quality recommendations, and (2) evaluate recommendations using the criteria developed.

Methods (1) Criteria were developed by experts using the bowtie method. Medical doctors then determined if the recommendations of Dutch in-hospital SE analysis reports met the criteria, after which interobserver variability was tested. (2) Researchers determined which recommendations of Dutch perioperative SE analysis reports produced from 2017 to 2018 met the criteria.

Results The criteria were: (1) a recommendation needs to be well defined and clear, (2) it needs to specifically describe the intended changes, and (3) it needs to describe how it will reduce the risk or limit the consequences of a similar SE. Validation of criteria showed substantial interobserver agreement. The SE analysis reports (n=115) contained 442 recommendations, of which 64% failed to meet all criteria, and 28% of reports did not contain a single recommendation that met the criteria.

Conclusion We developed and validated criteria for high-quality recommendations. The majority of recommendations did not meet our criteria. It was disconcerting to find that over a quarter of the investigations did not produce a single recommendation that met the criteria, not even in SEs with a fatal outcome. Healthcare providers have an obligation to prevent SEs, and certainly their recurrence. We anticipate that using these criteria to determine the potential of recommendations will aid in this endeavour.

INTRODUCTION

The WHO is urging healthcare providers to improve patient safety globally.[1] Unintended healthcare events that cause death or significant injury, so-called sentinel events (SEs), have a major impact on patients, relatives and healthcare workers and result in the loss of 64 million disability-adjusted life years per year worldwide.[2 3] The majority of reported in-hospital SEs are associated with surgical procedures.[4] Wrong-site surgery (WSS), for example—defined as invasive procedures on the wrong patient, at the wrong site or using the wrong procedure—is a persistent problem in most countries.[5–8] WSS and unintended retention of a foreign body have both been listed as obvious unacceptable errors or 'never events'. These never events are deemed preventable and are a major concern to both patients and healthcare workers.[9] Many countries have implemented a reporting system for SEs, and in the Netherlands, healthcare organisations are obliged to notify the Dutch Health and Youth Care Inspectorate (DHI) of all SEs, analyse each SE and send the analysis report to the DHI.[2] Based on this analysis, healthcare organisations formulate and implement recommendations aimed at preventing recurrence. Despite these and similar efforts, recurrent SEs such as WSS remain a persistent and widespread problem.[6–8] A recent survey among Dutch academic hospitals confirmed the recurrence of similar SEs, but also cast doubt on the quality of a significant proportion of the recommendations.[10] This does not solely apply to Dutch healthcare, as Hibbert *et al* have previously shown that only 8% of Australian recommendations are of high-quality.[11] However, there is currently no adequate guideline or strategy to determine if a recommendation will be effective, and selecting recommendations for implementation is still based on subjective criteria such as estimated feasibility.[10 12]

The DHI advises formulating recommendations based on Specific, Measurable, Achievable, Realistic and Timely (SMART) principles.[13] However, using SMART does not necessarily result in recommendations that reduce the risk of SE recurrence. For instance, if an SE occurs after a specific protocol was ignored, a frequently encountered recommendation is to improve awareness by sending the specific protocol via email to all healthcare workers within 1 month. While this recommendation may be formulated according to SMART principles, it is not an effective, sustainable intervention. Clearly, the content of the protocol needs to be scrutinised and the reason why it was ignored

should be investigated. Merely sending a reminder via email will not resolve the underlying problems that made the SE possible.

In this study, we (1) developed and validated basic criteria to determine the potential effect of recommendations designed to reduce the risk of recurrent SEs, and (2) evaluated the recommendations found in Dutch perioperative SE analysis reports using the newly developed criteria. Although implementation of recommendations and the monitoring of effects are critical steps in the learning process, this was considered beyond the scope of the present study.[10 14]

METHODS

Part 1a: developing criteria

An expert-based consensus meeting was organised that comprised three medical doctors specialised in the handling of SEs in healthcare, together with two experts specialised in dealing with the 'human factor' during incident analysis in other safety-critical industries. The 'human factor' field involves the study of humans interacting with their environment.[15] The goal of the meeting was to develop simple and easily applicable criteria to determine the potential effect of recommendations. The starting point for developing criteria was the bowtie method as used in barrier management. Barriers are physical or non-physical methods implemented by an organisation to prevent or mitigate SEs caused by hazards.[16]

The bowtie method, shown in figure 1, distinguishes barriers that aim to 'prevent', thus reducing the risk of a similar SE, or 'mitigate', representing an attempt to limit the consequences of an SE. Ideally, recommendations should interact with these barriers, an example of which might be that a patient dies as a result of a fire in an operating theatre due to a malfunctioning apparatus. Replacing the malfunctioning machine would be a preventive action, as it reduces the risk of a similar SE and thus belongs on the left side of the bowtie. Installing better fire detection equipment would be a corrective action, as it limits the consequences of an SE and therefore belongs on the right side of the bowtie. Components of recommendations which appear to enhance
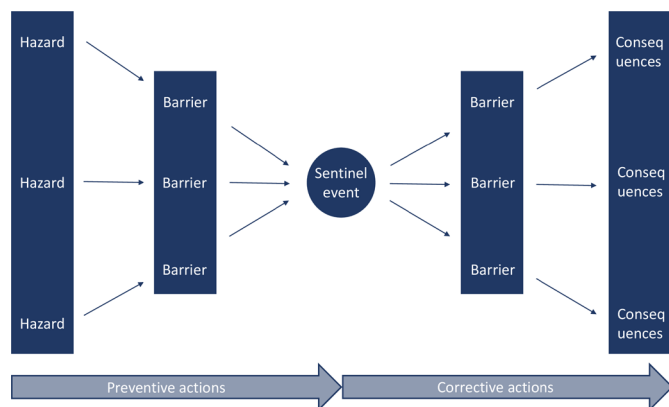
the effectiveness of recommendations in preventing SEs, the 'core component' or 'active component', are viewed as another basic condition.[17] Modifying a protocol, for example, is a meaningless recommendation without elaborating on the content. In order to distinguish an active component, the seven main questions comprising the steps 'searching for earlier effective recommendations' and 'clarifying the recommendation' (such as 'what is the core component of an earlier effective recommendation' and 'how can that component be integrated into the organisation') need to be answered.[18] The criteria were reviewed by an academic cognitive psychologist specialised in Safety Science, a senior DHI inspector, a former chief medical officer now specialised in SEs in healthcare and a healthcare implementation fellow. The criteria were adjusted based on their suggestions.

Part 1b: validating criteria

Validation of criteria facilitates appropriate use when selecting recommendations with the potential to reduce the risk of recurrent SEs. All Dutch healthcare organisations are obliged to report SEs to the DHI within 3 days after the occurrence or discovery of the event.[2] All correspondence between healthcare organisations and the DHI regarding the reported SE, including the results of the analysis and recommendations, is collected in the DHI database and the DHI assesses the quality of the analysis reports.[2] Ten complete analysis reports of general in-hospital SEs, assessed and approved by the DHI between 1 October 2017 and 1 January 2018, were randomly selected from the database for use in validating the criteria. Three medical doctors individually determined if the recommendations met the criteria, after which interobserver variability was tested.

Part 2: evaluating recommendations

In the DHI database all perioperative SEs that occurred between 1 July 2017 and 1 July 2018 were selected. A perioperative SE was defined as an SE that occurred prior, during or after an intervention executed in the operating theatre performed by a medical doctor in collaboration with an anaesthesiologist. All recommendations found in these SE analysis reports were noted. In addition, type of hospital, analysis method, analysis period and outcome of the SE were recorded. Two researchers (KB and A-FT (Acknowledgements)) determined, in consensus, which recommendations met the criteria.

Statistical analysis

Interobserver agreement was tested by calculating Fleiss' kappa (κ) and 95% CIs. Values ≤ 0 were interpreted as indicating no agreement, 0.01–0.20 as none to slight agreement, 0.21–0.40 as fair agreement, 0.41–0.60 as moderate agreement, 0.61–0.80 as substantial agreement and 0.81–1.00 as almost perfect agreement.[19] Descriptive data were reported as numbers with percentages (%) or median with IQR. The χ^2 test or Fisher's exact test was used when appropriate to compare proportions. The

Figure 1 Bowtie method.[16]

Mann-Whitney U test and Kruskal-Wallis test were used to compare continuous variables between the groups. $P<0.05$ was considered statistically significant. IBM SPSS Statistics for Windows (V.26.0, IBM) was used to perform statistical analysis.

Patient and public involvement
Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

RESULTS
Part 1a: developing criteria
The following three criteria for recommendations were developed:
► Based on the active component[17] → the recommendation needs to be well defined and clear.
► Based on the relevant questions[18] → the recommendation needs to specifically describe what will be modified in the future.
► Based on the bowtie method[16] → the recommendation needs to describe how it will reduce the risk or limit the consequences of a similar SE.

Part 1b: validating criteria
The 10 analysis reports on general in-hospital SEs randomly selected to validate the criteria yielded 38 recommendations. The three observers individually determined the same outcome for 29 of the 38 recommendations after applying the criteria. Fleiss' kappa (κ) was 0.67 (95% CI 0.48 to 0.85), resulting in substantial interobserver agreement.[19]

Part 2: evaluating recommendations
Between 1 July 2017 and 1 July 2018, a total of 115 perioperative SEs were reported to the DHI. Seventy-six (66%) SEs resulted in major injury and 39 (34%) in death. Eighty-five (74%) SEs occurred in the operating theatre, including 18 (21%) WSS cases and 16 (19%) cases of unintended retention of a foreign body.

The 115 perioperative SE analysis reports contained 442 recommendations in total. The median number of recommendations per SE was 4 (IQR 2–5) and the median number of recommendations per SE which met the criteria was 1 (IQR 0–2). Hospital type, analysis method, analysis period or the outcome of the SE did not significantly influence the mean number, as shown in online supplemental file 1. In total, 281 (64%) recommendations did not meet the three criteria.

Of the 115 SE analysis reports, 32 (28%) did not result in any recommendation that met the criteria. Eighty-three (72%) SE analysis reports produced at least one recommendation that met the criteria, as shown in online supplemental file 2. Of the 442 recommendations, five did not meet any of the three criteria. Half of the recommendations (244) met one of the three criteria, 32 recommendations met two of the three criteria and 161 met all three criteria. Almost every recommendation (437) had a

well-defined and clear goal. However, more than half of the recommendations did not specifically describe what needed to be modified (260) or how it might reduce the risk or limit the consequences of a similar SE (270) in the future.

DISCUSSION
In this study, a basic set of criteria was developed and validated that can be applied to determine the potential effect of recommendations to reduce the risk of recurrent SEs. Thereafter, an evaluation of the recommendations of Dutch perioperative SE analysis reports using the criteria developed was performed. The three criteria were easy to use and showed substantial interobserver agreement. Evaluation of the recommendations found in SE analysis reports indicated that most do not meet the three criteria and over a quarter of all investigations did not include a single recommendation that met the criteria.

Critical appraisal of our findings
Over the period 2014–2016 there were 100 reported cases in the Netherlands of WSS or unintended retention of a foreign body.[7] In this study, 34 of these cases were found within the selected year. Assuming that an approximately equal number of surgical procedures are performed each year, this number is consistent with expectations. The majority of recommendations did not meet our three criteria, which suggests that current recommendations are either not specifically aimed at reducing SEs or are not sufficiently well designed to reduce the risk of recurring never events.[9]

The number of reported SEs alone is not an adequate measure of the quality of delivered healthcare. The occurrence of an SE is influenced by a variety of factors, including awareness of patient safety and SEs, and the learning potential in dealing with SEs. However, the mere fact of a recurrence of a similar SE is in our view a reflection on the quality of delivered healthcare and an indication of how and whether we can learn and improve. Criteria that provide insight regarding the fundamentals of a recommendation could aid in decreasing the number of similar SEs and thus improve patient safety.

The quality of recommendations is only one step in the process of learning from SEs. A failure to draw appropriate conclusions from an SE is most likely multifactorial in nature, as the learning potential consists of many steps and each step is crucial. However, if recommendations fall short, subsequent steps in the learning process will be negatively affected and the ultimate goal of improvement will not be achieved.[10 14] Recent studies concluded that high-quality recommendations represent only a small proportion of the total and suggested this might be the reason why similar SEs persist.[10 11]

In the Netherlands, recommendations in relation to the analysis of SEs must be formulated according to SMART.[13] However, using SMART does not necessarily result in recommendations aimed at reducing the risk of recurrent SEs, as

SE analysis might highlight shortcomings in a healthcare system that did not directly contribute to that particular SE. For example, a patient dies following unintended retention of a foreign body. During the SE analysis, the analysis team notices technical shortcomings of the operating table and formulates a recommendation to ensure correct functioning of the operating table so that the table does not present a risk in future operations. While this recommendation is clearly aimed at improving patient safety, it is unlikely to have any effect on reducing the risk of an SE due to unintended retention of a foreign body.

To improve objectivity, members of the analysis team are frequently drawn from departments other than the department in which the SE occurred. This means that the home team (professionals working within the involved department) is not included in formulating the recommendations, and extensive information regarding specific medical knowledge or workflow might be taken less into account. As a result, formulating adequate recommendations becomes more challenging, since the exact circumstances in the department are likely to be unfamiliar to SE analysis team members. This could consequently result in recommendations that lack detail and have less impact on patient safety than intended. If our goal is to decrease the number of similar SEs—the reason why we actually analyse SEs—we must be objective but also take this possible reduced impact of the recommendations formulated by 'outsiders' into account. The developed criteria can support in formulating high-quality recommendations without the knowledge of the home team.

Our criteria were derived using the bowtie method. The bowtie method helps distinguish barriers that may eliminate or reduce the risk of a similar SE,[16] and an ideal recommendation should therefore interact with these barriers. When evaluating currently formulated recommendations, the majority fail to meet these criteria and as such show little potential to reduce the risks underlying the SE. It is not a prerequisite to be familiar with the bowtie method, we deconstructed this method and used it in our criteria. Keeping the criteria in mind can help maintain focus, and as such a departmental head or the executive board could use the criteria to select recommendations, and the analysis team could formulate risk-reducing recommendations based on the criteria.

Limitations

A limitation of our study is the fact that the criteria were not prospectively tested. Therefore, the impact and sustainability of the criteria on the quality of recommendations, the number of similar SEs and ultimately patient safety has yet to be proven in future prospective research. However, the criteria are based on available literature and validation resulted in substantial interobserver agreement. As there is currently no adequate guideline or strategy to determine if a recommendation will be effective, our easily applicable criteria could aid in filling this gap.

The bowtie method used here was developed in safety-critical environments other than healthcare. The number of studies within the medical field concerning the basic conditions of a recommendation is therefore limited, which might call into question the relevance of literature from other safety-critical industries to the medical field. However, the petrochemical industry, for example, played a prominent role in developing the Dutch reporting system that is used to this day.[20]

For practical reasons, the consensus meeting included a small panel of experts. Organising a broader team might have yielded different results, all team members included though were knowledgeable and experienced in dealing with SEs.

During the validation process we noticed that SE context is important, and in order to apply the criteria detailed information on the SE is essential. We expect this will be less of a challenge to a team analysing an SE than to the medical doctors involved in the validation process of our study, as the former will have access to detailed information on the SE, the department and the organisation where the SE took place.

In this study, we focused on recommendations found in Dutch perioperative SE analysis reports, and although we subsequently validated the criteria using recommendations from general in-hospital SE analysis reports, this might still limit the relevance in other healthcare systems. However, we are confident that the quality of recommendations is a universal problem and that our observations are potentially relevant to other healthcare systems and other safety-critical industries.[11]

CONCLUSION AND IMPACT

The goal of SE analysis is to provide recommendations that result in the improvement of patient safety. Nonetheless, similar SEs do recur and research suggests that the quality of recommendations might be partly responsible. In this study, we developed and validated three basic criteria that can be used to determine the potential effect of recommendations on reducing the risk of recurrent SEs, and we subsequently evaluated recommendations contained in 115 Dutch perioperative SE analysis reports. The majority of recommendations did not meet the three formulated criteria, and it was especially disconcerting to find that over a quarter of all investigations did not lead to a single recommendation that met the criteria. In other words, despite the fact that a patient had suffered serious harm or had even died, one out of four investigations did not effectively address the causes of the SE. It must be recognised that it is very well possible that the highest quality recommendations cannot be implemented due to valid practical reasons, such as financial constraints or lack of other resources. However, healthcare providers have an obligation to both their patients and colleagues to prevent SEs, and certainly their recurrence. We expect that using the three criteria formulated in this study will

help determine the value of safety recommendations and thus aid in this endeavour.

ORCID iDs
Kelly Bos http://orcid.org/0000-0001-8548-5286
Maarten J van der Laan http://orcid.org/0000-0002-3125-1342

REFERENCES

1 WHO. World patient safety day. World Health Organization (WHO), 2020. Available: https://www.who.int/patientsafety/world-patient-safety-day/en/ [Accessed Jan 2020].
2 Leistikow I, Mulder S, Vesseur J, *et al*. Learning from incidents in healthcare: the journey, not the arrival, matters. *BMJ Qual Saf* 2017;26:252–6.
3 WHO. Patient safety, global action on patient safety, report by the director-general. World Health Organization (WHO), 2020. Available: https://apps.who.int/gb/ebwha/pdf_files/WHA72/A72_26-en.pdf [Accessed July 2020].
4 de Vries EN, Ramrattan MA, Smorenburg SM, *et al*. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008;17:216–23.
5 The Joint Commission. Comprehensive accreditation manual for hospitals (CAMH) update 2. The Joint Commission, 2020. Available: https://www.jointcommission.org/-/media/deprecated-unorganized/imported-assets/tjc/system-folders/assetmanager/camh_24_se_all_currentpdf.pdf?db=web&hash=CAD6AB3AC78EAFD220CF9ACDD13772C1 [Accessed Apr 2020].
6 Geraghty A, Ferguson L, McIlhenny C, *et al*. Incidence of wrong-site surgery list errors for a 2-year period in a single National health service board. *J Patient Saf* 2020;16:79–83.
7 Leistikow IP, Bos ME. Links-rechtsverwisseling hardnekkig fenomeen. *Medisch Contact* 2017:18–21.
8 Seiden SC, Barach P. Wrong-side/wrong-site, wrong-procedure, and wrong-patient adverse events: are they preventable? *Arch Surg* 2006;141:931–9.
9 Lembitz A, Clarke TJ. Clarifying "never events and introducing "always events". *Patient Saf Surg* 2009;3:26.
10 Bos K, Dongelmans DA, Greuters S, *et al*. The next step in learning from sentinel events in healthcare. *BMJ Open Qual* 2020;9:e000739.
11 Hibbert PD, Thomas MJW, Deakin A, *et al*. Are root cause analyses recommendations effective and sustainable? an observational study. *Int J Qual Health Care* 2018;30:124–31.
12 Testik Özlem Müge, Shaygan A, Dasdemir E, *et al*. Selecting health care improvement projects: a methodology integrating cause-and-effect diagram and analytical hierarchy process. *Qual Manag Health Care* 2017;26:40–8.
13 Guideline sentinel event report. Dutch health and youth care Inspectorate. Available: https://www.igj.nl/onderwerpen/calamiteiten/documenten/richtlijnen/2016/01/01/richtlijn-calamiteitenrapportage [Accessed Mar 2020].
14 Drupsteen L, Groeneweg J, Zwetsloot GIJM. Critical steps in learning from incidents: using learning potential in the process from reporting an incident to accident prevention. *Int J Occup Saf Ergon* 2013;19:63–77.
15 Carayon P, Schoofs Hundt A, Karsh B-T, *et al*. Work system design for patient safety: the SEIPS model. *Qual Saf Health Care* 2006;15 Suppl 1:i50–8.
16 Sklet S. Safety barriers: definition, classification, and performance. *J Loss Prev Process Ind* 2006;19:494–506.
17 Dyreborg J, Lipscomb HJ, Olsen O, *et al*. Protocol: safety interventions for the prevention of accidents at work. *Campbell Syst Rev* 2015;11:1–70.
18 Herrera-Sánchez IM, León-Pérez JM, León-Rubio JM. Steps to ensure a successful implementation of occupational health and safety interventions at an organizational level. *Front Psychol* 2017;8:2135.
19 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
20 Willems R. Hier werk JE veilig, of JE werkt hier niet. Sneller Beter – de veiligheid in de zorg. *Eindrapportage Shell Nederland* 2004.