Research and reporting methodology

OPEN

Standard Elements in Studies of Adverse Events and Medical Error: the SESAME statement

Richard T. Griffey,[1] David C. Stockwell,[2] Lee M. Adler,[3] Rachel Ancona,[1] David Classen,[4,5] Ryan M. Schneider,[1] Anne WS. Rutjes,[6] Christopher R. Carpenter,[7] Maria Unbeck,[8,9] The SESAME Development Team

Abstract

Variability and persistent gaps in reporting have been consistently observed across studies evaluating adverse events in healthcare, dating back to the early days of the patient safety movement. Incomplete descriptions—particularly missing details regarding study design, conduct, methodologies and applied definitions—impede interpretation, hinder critical appraisal and limit reproducibility. These deficiencies not only obstruct a shared understanding of findings but also diminish the overall impact such studies may have on patient safety initiatives and policy development. To address these challenges and enhance the quality of reporting, we developed the SESAME (Standard Elements in Studies of Adverse Events and Medical Error) reporting guideline. SESAME is a 44-item checklist designed to support comprehensive and standardised reporting in studies evaluating potential patient harm. The guideline was developed over a 2-year period by an international, multidisciplinary panel through a consensus-driven process aligned with EQUATOR (Enhancing the QUAlity and Transparency Of Health Research Network) standards.

INTRODUCTION

In the 30+ years since the 'To Err is Human' report, studies evaluating patient harm and our understanding of medical error, adverse event (AE) analysis, patient safety, quality measurement and improvement have evolved and expanded considerably.[1–3] The body of literature in patient safety and quality has extended to explore numerous settings, conditions and aspects of adult and paediatric care, using various approaches to assessing error, harm and preventability. However, multiple editorials, systematic reviews and meta-analyses have observed incomplete or missing reporting and wide variability in the quality of reporting of basic but important methodological features. This has negative downstream consequences, impairing comparisons and the synthesis of knowledge.[4–12] Items such as definitions of fundamental outcomes, standards of causation used, details about reviewers, review approach, measures of reliability and other elements are missing in up to 50% of studies.[11] Subtle variations in approaches to detection and characterisation of safety events impact findings and may not be readily apparent unless comparing studies side by side.[13 14] Further, lack of explicit definitions hinders advancement in our understanding and characterisation of foundational constructs such as preventable harm.[15–17] Standard, transparent and complete reporting of methodological components is important for reproducibility, actionability and shared understanding of findings. Unclear or unreliable methods risk loss of credibility by clinicians and policy makers, which is ultimately self-defeating for improvement efforts. Researchers are in need of better guidance to facilitate complete reporting, which is essential for information synthesis, the advancement of methods and may ultimately translate to improved patient outcomes.

[1]Department of Emergency Medicine, Washington University in St. Louis School of Medicine, St. Louis, Missouri, USA; [2]Anesthesiology and Critical Care Medicine, Johns Hopkins University, Baltimore, Maryland, USA; [3]Clinical Quality and Safety, AdventHealth, Altamonte Springs, Florida, USA; [4]Division of Epidemiology, University of Utah, Salt Lake City, Utah, USA; [5]PascalMetrics, Washington, DC, USA; [6]Departmental Faculty of Medicine, UniCamillus - Saint Camillus International University of Health and Medical Science, Rome, Lazio, Italy; [7]Emergency Medicine, Mayo Clinic, Rochester, Minnesota, USA; [8]School of Health and Welfare, Dalarna University, Falun, Sweden; [9]Department of Clinical Sciences, Danderyd Hospital, Karolinska Institutet, Stockholm, Sweden

Correspondence to Richard T. Griffey; griffeyr@wustl.edu

Additional supplemental material is published online only. To view, please visit the journal online (https:// doi. org/ 10. 1136/ bmjqs- 2025-019458).

To cite: Griffey RT, Stockwell DC, Adler LM, et al. BMJ Qual Saf Epub ahead of print: [please include Day Month Year]. doi:10.1136/ bmjqs-2025-019458

© Author(s) (or their employer(s)) 2026. Re-use permitted under CC BY. https:// creativecommons.org/licenses/by-nc/4.0/ Published concurrently by both BMJ Publishing Group Limited in BMJ Quality & Safety and by Wolters Kluwer Health, LLC. in Journal of Patient Safety & Pediatric Quality and Safety. The article is identical in each publication except for minor stylistic and spelling differences in keeping with each publication's style. Citations from either of the two publications can be used when citing this article.

This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by-nc/4.0/.

Received for publication August 21, 2025; Accepted December 15, 2025.

Published online September 11, 2026

DOI: 10.1097/pq9.0000000000000931

To improve completeness and transparency, we set out to develop a consensus-based reporting guideline called SESAME (Standard Elements in Studies of Adverse Events and Medical Error) using the process outlined by the Enhancing the QUAlity and Transparency Of Health Research (EQUATOR) network.[18] Since 2006, the EQUATOR network, an international initiative to enhance the transparency and quality of health research, has fostered the development and use of reporting guidelines such as Consolidated Standards of Reporting Trials,[19] Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)[20] and Preferred Reporting Items for Systematic Reviews and Meta-Analyses,[21] to promote the inclusion of relevant data elements for an array of study types, settings and clinical areas.[22] Multiple studies have demonstrated improvement in reporting quality of manuscripts that use these reporting guidelines.[23–25] In addition, reporting guidelines are increasingly felt to be even more beneficial at earlier drafting and planning stages of research, with some agencies requiring their use for grant applications.[26] EQUATOR currently lists numerous guidelines that help ensure completeness in reporting and are increasingly encouraged or required by medical journals for manuscript submissions.[26–30] While several guidelines touch on AEs, there is none whose primary focus is on studies of AEs, preventable harm or medical error.[31] Our aims were to convene a diverse team of experts in patient safety research and methodology to review the AE literature and derive a consensus-based checklist of essential reporting standards in studies of AEs and medical error.

WHAT IS ALREADY KNOWN ON THIS TOPIC

- The quality of reporting of studies of adverse events and medical errors is suboptimal, hampering the interpretation of study findings, limiting not only the reproducibility of study findings but also their potential impacts for directing efforts towards patient safety and quality improvement.

WHAT THIS STUDY ADDS

- Using a modified Delphi procedure and an international team of experts, we created a checklist of essential reporting standards in studies of adverse event evaluation, which author teams should use to facilitate complete and transparent reporting.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- Standard, transparent and complete reporting of methodological components is important to ensure reproducibility, actionability and shared understanding of findings. Adherence to the SESAME Statement will improve the clarity of reporting, with positive downstream consequences including more accurate translation of study findings for patient safety measures and policy decisions.

METHODS

Overview

In March 2023, we began work generating the SESAME statement. We first conducted a literature search described below, then communicated with the EQUATOR network to identify existing related reporting guidelines. As we found little or no overlap with items we envisioned for our tool, we registered SESAME as a guideline in development. We established a steering committee and published our comparative findings.[31] We followed the outline in the EQUATOR network's Guidelines for Reporting Health Research manual, which includes a stepwise process for reporting standard development, divided into three phases: Pre-Meeting Phase; Meeting Phase and Post-Meeting Phase.[26] We have used this organisation to describe our work below.

Pre-Meeting Phase

Narrative review

We worked with a research librarian to perform a narrative review.[32] We searched Embase.com from 1947 using both controlled vocabulary and free text terms to find systematic reviews, scoping reviews or meta-analyses about AEs, medical errors, preventable harm and reporting practices. The search was run on 5 May 2023, with no additional limits to find 421 results. Two authors (RG, RMS) then performed a title and abstract review, which, along with citation pearl searching,[33] led to the identification of key manuscripts with candidate items for a draft checklist. Further detail of the search strategy is provided in online supplemental table 1.

Steering committee

While recruiting an international team to work on this project, we learned of a separate group in the early stages of a similar effort. We communicated and agreed to collaborate on this work and established a steering committee (nine members). We created a website where we could post the checklist, our task list, video, audio or other summaries of activity, including links to other reporting guidelines and host any intragroup dialogue (www.sesamestatement.org). We established a monthly meeting cadence for the steering committee, decided on the methods to generate the reporting guidance and began work on a manuscript outlining the need for a checklist. We decided unanimously that this checklist would be a stand-alone standard rather than being constructed as an extension of another checklist, namely of STROBE,[20]

since we anticipated including multiple new items specific to studies of AEs, and since we did not want this guidance to be limited to specific study designs.

Recruitment of the development team

In March–May 2023, the steering group recruited researchers with deep experience and expertise in AE evaluation and medical error. We first contacted authors of key review manuscripts highlighting gaps and variability in reporting and how these affected harm estimates and used snowball sampling for further recruitment.[34] We aimed to recruit a diverse group both geographically (and ultimately had representatives from the USA, Canada, Europe, Asia and Australia), and across disciplines, including experience in methodology, biostatistics, epidemiology, systematic reviews, the Delphi method and reporting guideline development. We also sought out those with backgrounds in various clinical specialties including nursing and pharmacy, and physicians in different areas of medicine and surgery. We recruited members with expertise in diagnostic error, acute, critical and postacute care, and informatics as well as expertise in topics with well-developed bodies of literature (eg, paediatrics, adverse drug events, ambulatory care) that might represent areas for future reporting guideline extensions. We recruited a patient representative, a researcher focused on health disparities, and input/representation from leaders at the US Office of the Inspector General, the Institute for Healthcare Improvement and Press Ganey. Many team members have editorial or other leadership roles as well that could assist with dissemination.

Meeting planning

We sent members of the development team an information packet that included key manuscripts, our draft checklist and a password to the website explaining our plan for a series of video call meetings interposed with rounds of voting, a list of tasks and proposed deliverables. We also described the rationale for the SESAME standard and included a conceptual framework with anticipated key reporting items.[31] We proposed a modified Delphi consensus process with an initial round of voting based on the findings of our review of systematic reviews plus items solicited by the development team, planning three rounds of voting with additional rounds as needed. We emphasised that it was not our intention in this project to establish common definitions (which, though much needed, would be a much larger undertaking) but rather to compel researchers to be explicit with the definition they used for concepts such as AE or preventability. However, because in order to draft a checklist, we had to agree on a small range of terms, we did create a terminology table (table 1), which provides a list of terms we use in the checklist, with alternate terms and definitions that are commonly used. This is not meant to be exhaustive nor definitive.

Meeting Phase

Development team group calls

We held a series of six web-based video calls beginning in late 2023 through early 2025 in which we discussed and deliberated various items for inclusion and wording. Meetings were scheduled to optimise participant availability across the international group. We recorded meetings and posted these on a password-protected area of our website, or transcript summaries were sent out to the group. On the first organising call, we made introductions, reviewed a statement of purpose and reviewed key literature by author/members on our team demonstrating gaps, variability and missing definitions. We also reviewed the recommended steps from EQUATOR for guideline development and an agenda of pre-meeting work that had been accomplished, including our conceptual framework, a tour of the website, a description of the consensus process and voting and finally an overview of the draft checklist. We began the item-by-item review on our second call. Each call that followed a vote summarised the data from the prior voting round, including free text comments received.

Voting

Our reporting on voting follows the ACCORD statement (Accurate Consensus Reporting Document: a reporting guideline for consensus methods in biomedicine developed via a modified Delphi) in reporting on our modified Delphi voting process.[35] We used a modified Delphi process to identify and prioritise items for the guideline. The Delphi method is a technique originated by the RAND in the 1950s in which a panel contributes ideas, ranking suggestions in successive rounds until consensus is reached.[36 37] Team members' contributions and voting are anonymous, but in each round, responses are informed by summary feedback from previous rounds. ACCORD describes many different consensus techniques and various modifications to these.[35] As opposed to what may be a more classic version of the Delphi technique, the items in our initial draft checklist were derived from our literature review with input from the steering committee. Subsequent items that were added were done so anonymously, as were proposed wording changes. On calls, we did openly discuss items, which can lead to biases such as social desirability bias.[38] However, we felt the risk of this was outweighed by the importance of having a firm and shared understanding of the import of items being considered that would result from open discussion.

We conducted modified Delphi rounds using a web-based English-language survey in REDCap (Research Electronic Data Capture), with links that were emailed to the development team (online supplemental table 2). REDCap is a secure, software platform designed to support data capture for research studies with this application hosted by Washington University in St Louis.[39 40] In each round, the denominator of development team voting was 30. No incentives were provided for survey completion.

Table 1. Terminology table

Terms used and alternate terms and definitions	
Act of commission	• An action that leads to some safety event (eg, prescribing, procedurally, in patient care). • Doing the wrong thing, at either the planning or execution phase.[34]
Act of omission	• A failure to act when there is broad consensus that an action is warranted in the circumstances present. This may include delays in treatment, diagnostic delays or failure to rescue. • Failing to do the right thing, at either the planning or execution phase.[34] • A failure to act that leads to a safety event.
AEs Harm, harm-associated incident, healthcare-associated harm	• An incident not due to the patient's underlying condition that resulted in physical harm to a patient. • Some definitions require that an intervention took place in response to an event in order to consider it a 'harm'.[35] • Some definitions require a causation standard of 'caused by' healthcare while others may use 'due to or contributed to by healthcare'.[13 14] • Impairment of structure or function of the body and/or any deleterious effect arising therefrom. Harm includes disease, injury, suffering, disability and death. Harm arising from or associated with plans or actions taken during the provision of healthcare, rather than an underlying disease or injury.[34] • Some definitions include psychological/emotional harm.[36]
Cascading event Related event	• A chain of events initiated by an unexpected result or other incident that may result in patient harm. • A series of events having a common root or in which one caused the next.[35]
Error	• Failure to carry out a planned action as intended or application of an incorrect plan.[34]
Event Patient safety incident Safety event Learning event	• Often used as an umbrella term that includes near misses, non-harm events, AEs and occurrences being evaluated to see whether harm occurred. • An event or circumstance that could have resulted, or did result, in unnecessary harm to a patient. An incident can be a reportable circumstance, near miss, no harm incident or AE.
Near miss Intercepted error	• Conditions are present that make an AE more likely. • An error or event that had the potential to cause harm and is identical to an AE except for the outcome. • An incident which did not reach the patient.[34] • An error or event that occurs that does not hit the patient and so does not produce harm.
Non-harm event No harm incident	• An error or event that hits the patient but does not cause harm or that may require monitoring or intervention in order to prevent harm from occurring. • An incident which reached a patient, but no discernible harm resulted.[34]
Preventability	• Accepted by the community as avoidable in the particular set of circumstances.[34] • An event that could have been prevented if adequate actions had been taken during the patient's contact with healthcare. • Preventability is often considered around the details of the sharp end of a patient interaction but should include consideration of latent features such as institutional policies, workflows, environment and culture, tasks and equipment in addition to healthcare workers and whether preventability assumes existing conditions or ideal/improved conditions.

AEs, adverse events.

We voted on items section by section, only after all the items in a section (eg, Methods, Results) had been discussed during a call and after we arrived at the wording choice we were going to vote on. Up to three reminders were sent to non-responders. Surveys were identifiable only to the analyst (RA), which allowed reminders to be sent to non-responders. No other members of the team had access to the REDCap database or response data, making the survey anonymous to members of the SESAME steering committee and development team.

Modified Delphi voting focused on (1) inclusion of the item in the checklist and (2) wording of the item. Voting for inclusion of each item in the checklist offered the following responses on a 1–5 Likert scale: 'Not Important', 'Slightly Important', 'Neutral', 'Important' and 'Very Important'. Additionally, team members were asked to score their agreement with the wording of each checklist item, also on a 1–5 Likert scale: 'Strongly Disagree', 'Disagree', 'Neutral', 'Agree' and 'Strongly Agree'. The voting form also included a free text field where team members could propose suggested wording, organisational or other changes. The survey was pilot tested by all members of the steering committee prior to launch to confirm that receipt and completion functioned correctly and that results were able to be analysed to produce results as anticipated. We performed a search of other common guidelines but found no 'standard' threshold for determining consensus. The steering committee thus defined the consensus threshold as ≥75%, which was one common option. We collapsed the options 'Important' and 'Very Important' for determining inclusion of an item and 'Agree' and 'Strongly Agree' for determination of consensus for wording.

Wording suggestions made in free text comments during voting were considered by the steering committee, both for items achieving consensus and those not. Small wording changes could be made and approved by the steering committee, and larger changes were discussed with the development team. Items where consensus on wording was not

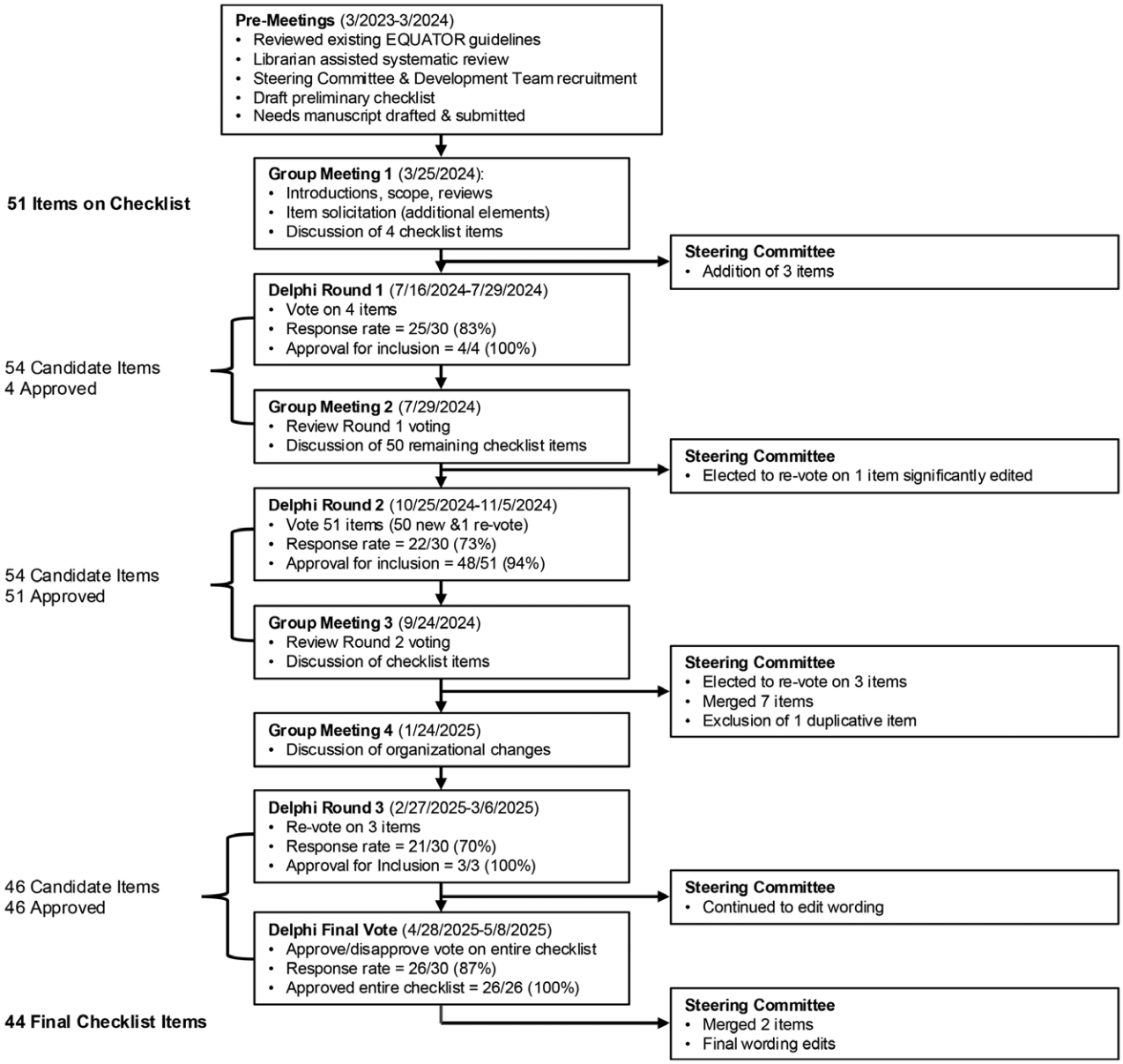

Figure 1. An overview of our consensus process, consisting of group meetings and rounds of Delphi voting in developing the SESAME checklist. EQUATOR, Enhancing the QUAlity and Transparency Of Health Research; SESAME, Standard Elements in Studies of Adverse Events and Medical Error.

achieved were discussed on following calls for modification and included in a subsequent round of voting. For items with persistent disagreement on the calls as to inclusion of specific wording, we voted in a yes/no fashion whether to include the specific wording in question. Items achieving consensus were not included in subsequent rounds for revoting unless there was some substantial wording change or if we were voting on merging two items. Our final round of voting consisted of a simple final yes/no vote whether to approve the checklist as crafted.

Trialling the checklist

Several team members applied the checklist to existing manuscripts with various study designs, including prospective studies, studies using real-time or near-real time computerised surveillance approaches, on their own manuscripts or grants in progress and asked colleagues to trial the checklist and to provide feedback on usability, clarity and usefulness.

Post-Meeting Phase and dissemination plan

Preparation of manuscripts and dissemination

All members of the development team were invited to author sections describing the rationale for the work and elaboration, which was accomplished using Google Docs to allow real-time collaboration. To disseminate our work more widely, we sought the interest of journal editors at leading patient safety journals in simultaneous

Table 2. SESAME: Standard Elements in Studies of Adverse Events and Medical Error

Elements	Item #	Guidance for reporting	Reported on page #
Title, abstract and keywords			
Title	1	Include a commonly used term (eg, adverse event, near-miss) that conveys the study topic in the title.	
Abstract	2	Provide an informative and balanced summary of background, objectives, methods, results and conclusion.	
Keywords	3	Use informative wording, including keywords, to facilitate correct indexing and retrievability in bibliographic databases.	
Introduction			
Background/ rationale	4	Describe the scientific and clinical background and rationale for undertaking the study (eg, to characterise harm, measure the adverse event rate, test or compare a surveillance tool).	
Objectives	5	State specific objective(s).	
Methods			
Study design	6a	Describe the study design (eg, retrospective or prospective; observational or interventional).	
	6b	Provide details of any patient or public involvement during any phase of the study (from design to dissemination) or state no involvement.	
Setting details	7a	Describe the country/countries in which the study takes place or that is the source for data used.	
	7b	Describe the setting (eg, academic, community, both, urban, rural).	
	7c	Describe the facility or unit type (eg, hospital, ambulatory, intensive care unit, nursing home) and volume (eg, single or multicentre, number of inpatient beds, annual visit volume, catchment area).	
Population details	8a	Describe any eligibility criteria related to a target medical condition or specialty (eg, oncology, orthopaedic, acute/elective care).	
	8b	Describe any eligibility criteria related to participant demographics (eg, adult, age, race, ethnicity, gender).	
Exclusion criteria	9	Describe any exclusion criteria for participants or events.	
Case finding/ sampling strategy	10a	Provide details for the study period (eg, dates, recruitment period, duration).	
	10b	Describe the method of case finding used (eg, all records meeting screening criteria; triggered records, computerised surveillance).	
Reviewers and training	11a	Describe reviewer recruitment, qualifications and experience in event adjudication.	
	11b	Describe reviewer training and what performance standards for reviewers (if any) were required prior to independent review.	
Review process	12a	Describe how reviews were conducted (eg, by an individual, multiple reviewers working independently or in phases, or together as a group) and any consensus technique used.	
	12b	Describe any time limit on the duration of reviews, as this may impact outcome estimates (eg, a maximum of 20 min per review).	
	12c	Describe how individual events were determined (eg, implicit or explicit review, use of an algorithm, use of automated/machine learning approaches).	
Event scope and definition	13a	Indicate the outcome(s) of interest (eg, all-cause harm, preventable or ameliorable harm, errors, near misses, non-harm events) and include explicit definitions for the outcome(s) of interest. If harm is the outcome, indicate whether the definition of harm requires that an intervention take place.	
	13b	Describe the causation standard used when determining events (eg, 'caused by' vs 'resulting from or contributed to by' health care) and specify whether a scale or rating system was used to characterise the confidence of causation and if so, describe this and its origin.	
	13c	Provide the time frame(s) for event inclusion (eg, present on arrival, during the index visit, some period of time after an intervention).	
	13d	Indicate whether event capture included acts of omission, acts of commission or both.	
	13e	Indicate whether 'cascading' events, where one event caused the next, were counted as a single event or each event was counted separately.	
Characterisation of events	14a	Describe the taxonomy or classification system used to describe the event and/or contributing factors (eg, surgical/procedural events, health care-acquired infections, adverse drug events).	
	14b	Describe whether a severity assessment was made and if so, what scale was used and its origin.	
Preventability	15a	Indicate whether preventability was assessed and if so, provide the explicit definition used.	
	15b	Describe any preventability scale or rating system used, its origin and the language used for various ratings (eg, definitely not preventable, possibly vs potentially preventable).	
	15c	Describe the reviewer process used in determining preventability (eg, single or tiered review, single or dual reviewers per tier, group review).	
Quality assurance	16a	Indicate any deviations from the intended study plan.	
	16b	Describe any quality assurance mechanisms to improve validity and reliability (eg, monitoring/auditing, feedback provided to reviewers).	
	16c	Include assessment(s) of inter-rater reliability for outcomes in the study (eg, whether an event represented harm, its preventability and categorisations) and describe how assessments were conducted and calculated.	
	16d	For any automated electronic data capture, describe whether there was any process to validate data against manual review.	
Data analysis	17a	Describe all statistical methods used, including whether this study includes a post hoc analysis and how missing data were handled.	
	17b	Explain how the sample size was determined	
	17c	Specify the unit(s) used for any presentation of event rates (eg, number of events, number of events per patient, proportion of visits with one or more events, number of events per visit/hospitalisation, events per 100 visits/hospitalisations or per 1000 patient days).	
	17d	Indicate whether any scales that were used to assess causation or preventability were collapsed for analysis.	
	17e	Describe procedures used to assess and control for bias or confounding.	

(Continued)

Table 2. Continued

Elements	Item #	Guidance for reporting	Reported on page #
Title, abstract and keywords			
Results			
Participants	18	Report the numbers of participants at each stage of the study—those eligible, excluded (with reasons for exclusion), included and analysed, included but not analysed. Consider including a flow diagram to depict the design and patient flow.	
Descriptive data	19	Report data on characteristics of study participants (eg, demographic, clinical, social determinants of health).	
Study findings	20	Report the findings of the study including all outcomes described in the Methods section.	
Discussion			
Key results	21	Summarise key results with reference to study objectives and discuss the findings in relation to existing literature.	
Strengths and limitations	22	Discuss strengths and limitations of the study including sources of potential bias, and internal and external validity.	

publication agreements. To increase impact, we also planned follow-up editorials introducing and advertising the checklist and discussed using our SESAME website to inform a larger group of stakeholders.

RESULTS

Our development team of 31 individuals was 60% male; 40% female with 20 members from the USA, 8 from Europe and 1 from Australia, Canada and Asia respectively (more detailed member data in online supplemental table 2).

During the first call, the scope of the SESAME guidance and checklist was decided. We decided that the guidance should apply to various types of study designs, including observational and experimental designs, of prospective or retrospective nature. We decided that we would first develop a generic tool for use in studies evaluating AEs and medical error and would consider adaptations or extensions after publication of the tool, to accommodate studies focusing on, for example, drug-related harm or diagnostic error. The target users were determined as authors of primary studies, editors and peer reviewers.

We held a total of four rounds of individual voting. We began with a list of 51 candidate items. In round one, we approved four items but revoted on one in the second round due to substantial wording modifications. Three new items solicited in the first group meeting were added prior to round two. In that round, we thus voted on 51 items altogether, approving 48 for inclusion but with 9 requiring wording modifications. On subsequent group and steering committee calls, we merged several items, eliminated a duplicative item and made wording modifications to nine items. In round three, we had a total of nine voting items: revoting on inclusion of three previously discussed items, four wording changes and two yes/no votes on inclusion of specific phrasing that where members were previously split. After merging two other items previously discussed, this yielded a 44-item checklist (22 elements with subitems totalling 44). There were no deviations from our study plan. Details of meetings and of voting are displayed in figure 1.

Four development team members tested the draft SESAME checklist (online supplemental material 1) on existing manuscripts or manuscripts in development. This also included solicitation from one colleague who was not a member of the development team. We received favourable feedback from those who applied the checklist, with no major gaps identified and reports that most items were applicable across study types trialled. The checklist (table 2) is downloadable and designed to be filled in with the section/subsection or line number where the respective item is addressed in the manuscript or to place an N/A (not applicable) note in the space if this is warranted.

The SESAME Statement will be jointly and simultaneously published in *BMJ Quality and Safety*, *The Journal of Patient Safety* and *Pediatric Quality and Safety*. The explanation of how to apply the tool, together with a comprehensive description of each item in the SESAME checklist and illustrative examples from the literature, is presented in a separate publication.[41] Other activities to enlarge the potential impact of our tool will be completed after these publications.

DISCUSSION

Studies that measure harm or error are needed to evaluate progress or highlight areas for improvement in patient safety efforts. The details of how such events are measured have significant import for the findings of such studies, the lessons learnt and ability to compare to other similar studies. Not uncommonly, studies report an AE rate in some area of healthcare and a proportion that was found to be preventable without providing a clear definition of what is meant by AE, harm or by preventability and without adequate detail needed to reproduce or compare main study findings. It is our intent that the SESAME checklist is a solid first step towards improved reporting, transparency and shared understanding.

We closely adhered to the process recommended by the EQUATOR network. We confirmed a lack of existing reporting guidelines to address this area and created a consensus-derived checklist with 22 elements and a

total of 44 items. Because of the importance of the topic area and the need for a standard that applies across study designs, team members felt strongly that this checklist should be a stand-alone standard rather than an extension of another guideline. While many studies of AEs use retrospective, observational methodology, there are many examples of prospective or other study methods, increasingly with use of computerised surveillance and even real-time AE detection or prediction. We feel that this checklist is still largely responsive to different study designs and can be used in addition to many of the existing checklists that are specific to certain methodologies.

Aware of the challenges experienced in a large prior effort at the WHO in 2009 to standardise language, it was never our intention in this project to establish common definitions. Though much needed, this would be a significantly larger undertaking. While we hope that use of the checklist nudges researchers towards use of common conceptions and language, it was not our intent to derive consensus definitions for terminology. Rather, our goal was to compel researchers to be explicit with the definitions they use for concepts such as AE or preventability in the hopes that this would improve transparency and allow for clearer comparisons.

Strengths and limitations

We used a systematic approach to generate initial candidate items in addition to soliciting checklist items from the entire development team based on expert opinion. The breadth and depth of experience and expertise in our international multidisciplinary development team was a key strength, though we did not have geographical representation from South America, Africa, the Middle East and diverse areas of Asia. Our development team covered a wide range of expertise, although for any future updates, we would add a human factors engineer. All participants remained on the development team throughout, without anyone withdrawing. The response rate was good, with little variability in voting participation. During our first call, when we discussed the scope of the SESAME guidance, there was some interest expressed in being mindful of newer conceptions of harm that include emotional harm. We decided as a group not to pursue this in the first version of the checklist but to include this as a future consideration. The focus in this checklist was also decidedly on a Safety I approach, attempting to learn from AEs rather than a Safety II approach, focusing on adaptability and resilience of teams in ensuring safe outcomes.[42] This might also represent an opportunity for future work. We discussed but did not specifically decide on areas for extensions of SESAME.

CONCLUSION

The SESAME checklist was created using a rigorous process consistent with the EQUATOR guidelines for development of reporting standards. SESAME enhances complete and clear reporting in studies evaluating patient harm and patient safety. Use of the checklist can significantly improve transparency and reliability in this important area of research.

COLLABORATORS

The SESAME Development Team: Lee M Adler DO, Andrew Auerbach MD, MPH, MHM, David Bates MD, MSc, Benjamin Brooke MD, PhD, Andrew Carson-Stevens MRGCP, PhD, Christopher R Carpenter MD, MSc, David Classen MD, MS, Ruth Ann Dorrill MPA, Tejal Ghandi MD, MPH, Richard T Griffey, MD, MPH, Peter D Hibbert MD, Dorothy O Klein PhD, Akira Kuriyama MD, MPH, PhD, Edmund Kwok MD, MHA, MSc, Patricia McGaffigan MS, RN CPPS, Kenneth Michelson MD, MPH, M Hassan Murad MD, MPH, Sarah Musy PhD, MSc, Maria Panagioti PhD, Anne W S Rutjes PhD, MSc, Ryan M Schneider ACNP, CPPS, CPHQ, Lucy Schulson MD, MPH, Rene Schwendimann PhD, RN, Sue Sheridan, MIM, MBA, DHL, Michael Simon PhD, RN, Hardeep Singh MD, MPH, David C Stockwell MD, Eric Thomas MD, MPH, Maria Unbeck PhD, RN, Almut Winterstein, RPH, PhD.

CONTRIBUTORS

All authors contributed to the conception of this manuscript. RTG primarily drafted and all authors edited the manuscript. RA assisted with data management and editing. RTG is the guarantor of the manuscript overall. We used AI (ChatGPT) to revise our abstract. AI was not used in any other portion of the manuscript.

FUNDING

The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

COMPETING IN TERESTS

None declared.

PATIENT CONSENT FOR PUBLICATION

Not applicable.

ETHICS APPROVAL

Not applicable.

PROVENANCE AND PEER REVIEW

Not commissioned; externally peer reviewed.

DATA AVAILABILITY STATEMENT

All data relevant to the study are included in the article

SUPPLEMENTAL MATERIAL

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID IDs

Richard T Griffey https://orcid.org/0000-0002-6279-7380
David C Stockwell https://orcid.org/0000-0001-6074-5731
Anne W S Rutjes https://orcid.org/0000-0001-9782-779X
Christopher R Carpenter https://orcid.org/0000-0002-2603-7157
Maria Unbeck https://orcid.org/0000-0002-5090-0352

REFERENCES

1. Stelfox HT, Palmisani S, Scurlock C, et al. The "To Err is Human" report and the patient safety literature. *Qual Saf Health Care* 2006;XXX:174–8. doi:10.1136/qshc.2006.017947
2. Bates DW, Singh H. Two Decades Since To Err Is Human: An Assessment Of Progress And Emerging Priorities In Patient Safety. *Health Aff (Millwood)* 2018;37:1736–43. doi:10.1377/hlthaff.2018.0738
3. Kohn LT, Corrigan J, Donaldson MS. *To err Is human: building a safer health system*. xxi. Washington, D.C: National Academy Press, 2000:287.
4. Hodkinson A, Tyler N, Ashcroft DM, et al. Preventable medication harm across health care settings: a systematic review and meta-analysis. *BMC Med* 2020;18:313. doi:10.1186/s12916-020-01774-9
5. Panesar SS, deSilva D, Carson-Stevens A, et al. How safe is primary care? A systematic review. *BMJ Qual Saf* 2016;25:544–53. doi:10.1136/bmjqs-2015-004178
6. Anderson O, Davis R, Hanna GB, et al. Surgical adverse events: a systematic review. *Am J Surg* 2013;206:253–62. doi:10.1016/j.amjsurg.2012.11.009
7. Panagioti M, Khan K, Keers RN, et al. Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis. *BMJ* 2019;366. doi:10.1136/bmj.l4185
8. Schildmeijer KGI, Nilsson L, Arestedt K, et al. The assessment of adverse events in medical care; lack of consistency between experienced teams using the Global Trigger Tool'. *BMJ Qual Saf* 2013;22:271–2. doi:10.1136/bmjqs-2012-001587
9. Ferner RE, Aronson JK. Preventability of drug-related harms - part I: a systematic review. *Drug Saf* 2010;33:985–94. doi:10.2165/11538270-000000000-00000
10. Schwendimann R, Blatter C, Dhaini S, et al. The occurrence, types, consequences and preventability of in-hospital adverse events - a scoping review. *BMC Health Serv Res* 2018;18:521. doi:10.1186/s12913-018-3335-z
11. Hibbert PD, Molloy CJ, Hooper TD, et al. The application of the Global Trigger Tool: a systematic review. *Int J Qual Health Care* 2016;28:640–9. doi:10.1093/intqhc/mzw115
12. 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. doi:10.1136/qshc.2007.023622
13. Thomas EJ, Studdert DM, Runciman WB, et al. A comparison of iatrogenic injury studies in Australia and the USA. I: Context, methods, casemix, population, patient and hospital characteristics. *Int J Qual Health Care* 2000;12:371–8. doi:10.1093/intqhc/12.5.371
14. Runciman WB, Webb RK, Helps SC, et al. A comparison of iatrogenic injury studies in Australia and the USA. II: Reviewer behaviour and quality of care. *Int J Qual Health Care* 2000;12:379–88. doi:10.1093/intqhc/12.5.379
15. Hakkarainen KM, Andersson Sundell K, Petzold M, et al. Methods for assessing the preventability of adverse drug events: a systematic review. *Drug Saf* 2012;35:105–26. doi:10.2165/11596570-000000000-00000
16. Nabhan M, Elraiyah T, Brown DR, et al. What is preventable harm in healthcare? A systematic review of definitions. *BMC Health Serv Res* 2012;12:128. doi:10.1186/1472-6963-12-128
17. Klein DO, Rennenberg RJMW, Koopmans RP, et al. A Systematic Review of Methods for Medical Record Analysis to Detect Adverse Events in Hospitalized Patients. *J Patient Saf* 2021;17:e1234–40. doi:10.1097/PTS.0000000000000670
18. Gould KA. The EQUATOR Network: A Resource for Authors. *Dimens Crit Care Nurs* 2016;35:350. doi:10.1097/DCC.0000000000000213
19. Hopewell S, Chan A-W, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. *Lancet* 2025. doi:10.1016/S0140-6736(25)00672-5
20. Elm E von, Altman DG, Egger M. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;335:805–35. doi:10.1136/bmj.39386.490150.94
21. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372. doi:10.1136/bmj.n71
22. Simera I, Moher D, Hirst A, et al. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med* 2010;8:24. doi:10.1186/1741-7015-8-24
23. Prager R, Bowdridge J, Kareemi H, et al. Adherence to the Standards for Reporting of Diagnostic Accuracy (STARD) 2015 Guidelines in Acute Point-of-Care Ultrasound Research. *JAMA Netw Open* 2020;3:e203871. doi:10.1001/jamanetworkopen.2020.3871
24. Korevaar DA, van Enst WA, Spijker R, et al. Reporting quality of diagnostic accuracy studies: a systematic review and meta-analysis of investigations on adherence to STARD. *Evid Based Med* 2014;19:47–54. doi:10.1136/eb-2013-101637
25. Hopewell S, Ravaud P, Baron G, et al. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. *BMJ* 2012;344:e4178. doi:10.1136/bmj.e4178
26. Moher D, Schulz KF, Simera I, et al. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010;7:e1000217. doi:10.1371/journal.pmed.1000217
27. Golub RM, Fontanarosa PB. Researchers, readers, and reporting guidelines: writing between the lines. *JAMA* 2015;313:1625–6. doi:10.1001/jama.2015.3837
28. Arundel C, James S, Northgraves M, et al. Study reporting guidelines: How valid are they? *Contemp Clin Trials Commun* 2019;14:100343. doi:10.1016/j.conctc.2019.100343
29. Hirst A, Altman DG. Are peer reviewers encouraged to use reporting guidelines? A survey of 116 health research journals. *PLoS One* 2012;7:e35621. doi:10.1371/journal.pone.0035621
30. International Committee of Medical Journal Editors. Recommendations on the conduct, reporting, editing, and publication of scholarly work in medical journals. 2018. Available: https://www.icmje.org/recommendations/archives/2018_dec_urm.pdf
31. Carpenter CR, Griffey RT, Rutjes AWS, et al. The problem with the existing reporting standards for adverse event and medical error research. *BMJ Qual Saf* 2025;34:273–8. doi:10.1136/bmjqs-2024-017491
32. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J* 2009;26:91–108. doi:10.1111/j.1471-1842.2009.00848.x
33. Ramer SL. Site-ation pearl growing: methods and librarianship history and theory. *J Med Libr Assoc* 2005;93:397–400.

34. Heckathorn DD. SNOWBALL VERSUS RESPONDENT-DRIVEN SAMPLING. *Sociol Methodol* 2011;41:355–66. doi:10.1111/j.1467-9531.2011.01244.x
35. Gattrell WT, Logullo P, van Zuuren EJ, et al. ACCORD (ACcurate COnsensus Reporting Document): A reporting guideline for consensus methods in biomedicine developed via a modified Delphi. *PLoS Med* 2024;21:e1004326. doi:10.1371/journal.pmed.1004326
36. Dalkey NC, Brown BB, Cochran S. The delphi method. Santa Monica, Calif Rand Corporation Memorandum RM-5888-PR, RM-5957-PR, RM-6115-PR, RM-6118-PR; 1969.
37. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;32:1008–15.
38. Piedmont RL, McCrae RR, Riemann R, et al. On the invalidity of validity scales: evidence from self-reports and observer ratings in volunteer samples. *J Pers Soc Psychol* 2000;78:582–93. doi:10.1037//0022-3514.78.3.582
39. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019;95:103208. doi:10.1016/j.jbi.2019.103208
40. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81. doi:10.1016/j.jbi.2008.08.010
41. Griffey RT, Adler LM, Stockwell DC. Standard elements in studies of adverse events and medical error (SESAME): explanation and elaboration. *BMJ Qual Saf* 2025. doi:10.1136/bmjqs-2025-019458
42. Braithwaite J, Wears RL, Hollnagel E. Resilient health care: turning patient safety on its head. *Int J Qual Health Care* 2015;27:418–20. doi:10.1093/intqhc/mzv063