

Systematic review of grading systems for adverse surgical outcomes

Saba Balvardi, MD
Etienne St-Louis, MD
Yasmine Yousef, MD
Asra Toobaie, MD
Elena Guadagno, MLIS
Robert Baird, MD
Dan Poenaru, MD

Presented at the Canadian Surgical Forum
2018, Sept. 13–15, 2018, St. John's, Nfld.

Accepted May 5, 2020

Correspondence to:

D. Poenaru
McGill University Health Centre
B04.2022.2–1001 Décarie Blvd
Montréal QC H4A 3J1
dan.poenaru@mcgill.ca

DOI: 10.1503/cjs.016919

Background: Grading scales for adverse surgical outcomes have been poorly characterized to date. The primary aim of this study was to conduct a systematic review to enumerate the various frameworks for grading adverse postoperative outcomes; our secondary objective was to outline the properties of each grading system, identifying its strengths and weaknesses.

Methods: We searched 9 databases (Africa Wide Information, Biosis, Cochrane, Embase, Global Health, LILACs, Medline, PubMed and Web of Science) from 1992 (the year the Clavien–Dindo classification system was developed) until Mar. 2, 2017, for studies that aimed to develop or improve on an already existing generalizable system for grading adverse postoperative outcomes. Study selection was duplicated as per PRISMA recommendations. Procedure-specific grading systems were excluded. We assessed the framework, strengths and weaknesses of the systems qualitatively.

Results: We identified 9 studies on 8 adverse outcome grading systems with frameworks generalizable to any surgical procedure. Most systems have not been widely incorporated in the literature. Seven of the 8 systems were produced without including patients' perspectives. Four allowed the derivation of a composite morbidity score, which had limited tangible significance for patients.

Conclusion: Although each instrument identified offered its own advantages, none satisfied the need for a patient-centred tool capable of generating a composite score of all possible postoperative adverse outcomes (complications, sequelae and failure) that enables comparison of noninterventional and surgical management of disease. There is a need for development of a more comprehensive, patient-centred grading system for adverse postoperative outcomes.

Contexte : Jusqu'ici, les systèmes de classification des issues postopératoires indésirables n'ont pas encore fait l'objet d'une analyse comparative. Cette étude avait pour objectif principal de recenser, au moyen d'une revue systématique de la littérature, les divers systèmes de classification des issues postopératoire indésirables, et pour objectif secondaire de dégager les propriétés, les forces et les faiblesses de chaque système.

Méthodes : Nous avons interrogé 9 bases de données (Africa Wide Information, Biosis Previews, Cochrane, Embase, Global Health, LILACS, Medline, PubMed et Web of Science) pour trouver des articles publiés entre 1992 (année de la mise au point du système de classification de Clavien–Dindo) et le 2 mars 2017. Ces articles devaient porter sur la création d'un système généralisable de classification des issues postopératoires indésirables, ou l'amélioration d'un système existant. La sélection des études a été faite en double, conformément aux recommandations PRISMA. Les systèmes de classification visant une seule intervention ont été exclus. Nous avons évalué, d'un point de vue qualitatif, le cadre, les forces et les faiblesses des systèmes retenus.

Résultats : Nous avons retenu 9 études sur 8 systèmes de classification accompagnés d'un cadre pouvant être appliqué à n'importe quelle intervention chirurgicale. La plupart des systèmes n'avaient pas été largement étudiés. Sept des 8 systèmes avaient été développés sans tenir compte du point de vue des patients, et 4 permettaient de calculer un score de morbidité composite ayant des retombées concrètes limitées pour les patients.

Conclusion : Tous les systèmes retenus s'accompagnaient d'avantages, mais aucun ne pouvait servir d'outil centré sur le patient permettant de calculer un score composite pour toutes les issues postopératoires possibles (complications, séquelles et échec), score qui pourrait servir à comparer les prises en charge conservatrice et chirurgicale des maladies. La création d'un système de classification des issues postopératoires indésirables exhaustif centré sur le patient est nécessaire.

Lack of consensus in defining and measuring the severity of adverse surgical outcomes hinders reliable comparison and categoric assessment of the quality and risks of surgical procedures.^{1,2} About 80% of studies describing postoperative complications fail to indicate their severity.² After a dramatic decrease in postoperative mortality in recent years, morbidity from surgical procedures is emerging as the main parameter in defining procedural safety and quality.^{3,4} The ability to classify, grade, risk-adjust and compare adverse surgical outcomes in a standardized and reproducible manner is necessary for quality improvement.⁵

Adverse surgical events can be divided into sequelae, procedural failures and complications.⁵ Surgical sequelae (e.g., loss of a limb after surgical amputation for treating wet gangrene) are negative outcomes inherent to a given procedure. Surgical failures (e.g., tumour recurrence after resection) are events in which the purpose of the procedure is not fulfilled. Surgical complications are unexpected negative outcomes of a given procedure (e.g., postoperative pneumonia). Even though surgical sequelae are preventable only if the surgical procedure does not take place, they have a definite impact on patients' quality of life after surgery. Data on the impact of treatment sequelae or surgical failure on clinical decision-making are limited, yet such adverse events are proven to affect quality of life,^{6,7} which, in turn, should influence the treatment decisions made by patients and providers.

The Clavien–Dindo classification system is one of the first surgical complication grading systems to become widely accepted and used in high-quality trials and national databases.^{8,9} Its popularity lies in its strengths: simplicity, adaptability to all procedures and reduction in subjectivity of reporting postoperative complications (by focusing on the interventions needed to treat complications).^{8,9} Focusing on interventions also permits retrospective measurement of complications in a more objective manner, less affected by subjective reporting of intra- and postoperative complications.¹ This allows the Clavien–Dindo system to be less dependent on continuous monitoring, as it is focused on symptomatic complications requiring medical or surgical intervention.¹⁰

A key limitation of the Clavien–Dindo system is the absence of the patient's perspective, as complications are described and graded based only on the interventions required to treat them, rather than on patient-reported outcomes. This is highlighted by a study by Winslow and colleagues¹¹ showing no correlation between the complication grade derived from the Clavien–Dindo system and patient-reported severity scores for negative postoperative outcomes. Moreover, the Clavien–Dindo system does not provide an overall morbidity burden for a given procedure, as it focuses on the clinically most severe complication in any given patient.¹² This system is also not validated for evaluation of adverse outcomes after radiologic and medical interventions,

and has limited ability to grade the severity of adverse outcomes of nonoperative treatment of “surgical” disease.

There is a paucity of information in the literature regarding the grading of severity of adverse events resulting from nonoperative treatments. Operative and nonoperative treatment strategies are often compared in trials with the use of mortality and specific predefined morbidities as primary outcomes, often without a systematic approach to grading the severity of these morbidities. This becomes increasingly of concern when assessing pediatric and trauma populations, as the prevalence of nonoperative management of diseases that were previously treated surgically is increasing.^{13–15}

Given the differing inherent qualities of each complication grading system in existence, this systematic review aimed to enumerate the current systems for grading adverse surgical outcomes. Our secondary objective was to outline the properties of each grading system to better characterize their strengths and weaknesses.

METHODS

This systematic review was registered with PROSPERO (CRD42017058650) on Mar. 29, 2017, and was conducted according to the PRISMA Statement guidelines.¹⁶

Search strategy

We searched the following databases from 1992 until Mar. 2, 2017: Medline, Embase, Biosis, Global Health, Cochrane, PubMed, Africa-Wide Information, LILACS (Latin American and Caribbean Center on Health Sciences Information) and Web of Science, with no language restrictions. We chose the cut-off date of 1992 because the Clavien–Dindo system was developed in that year.⁴

The search strategy used variations in text words found in the title, abstract or keyword fields, and relevant subject headings to retrieve articles pertaining to postoperative complications and various grading scales, classifications, health indicators or surveys, with modifications to search terms as necessary. The grey literature was included in our search strategy after 2012 to minimize the selection bias for more recent studies that would have not yet been published in peer-reviewed journals. Articles in languages other than English were translated by means of Google Translate. We used the snowballing technique to extend the scope of the search, searching the reference list or citations of the papers selected for full-text review to identify additional papers. Full details of the search strategy are provided in Appendix 1 (available at canjsurg.ca/016919-a1).

Study selection and inclusion and exclusion criteria

Two reviewers (S.B. and A.T.) independently assessed titles, abstracts and selected studies for full-text review. In

case of any disagreement regarding inclusion or exclusion, a third, independent reviewer (E.S.) assessed the article in question for inclusion. The PRISMA flow diagram was used to track the number of records identified, included or excluded.¹⁶

The inclusion criteria used in the review were 1) the instrument self-identified as an instrument for grading adverse surgical outcomes (complication, failure or sequelae), 2) the aim of the study was development of a new instrument or improvement of an already existing instrument, 3) the study was published as a full-text original article and 4) the measurement tool was validated in at least 1 institution.

Exclusion criteria were 1) disease-specific or procedure-specific instruments; 2) symptom-specific instruments, 3) adaptation of an already existing instrument for a given procedure; 4) case reports, comments, news and editorials; 5) measures of pre- and intraoperative complications; and 6) animal studies.

Data extraction and analysis

The reviewers (S.B. and A.T.) independently extracted the following data from the selected full-length articles: publication data (e.g., author, year of publication, field of surgical specialty), patient population data (demographic characteristics, diagnosis) and details of the grading system (instrument framework, existence of composite score or patient's perspective in the creation of the instrument). A sample of the data-extraction worksheets is found in Appendix 2 (available at canjsurg.ca/016919-a2). Items not available were noted and reported as missing in the final report. We reported this systematic review using a narrative synthesis approach.¹⁷

RESULTS

After removal of duplicates, we identified 17 147 citations, among which 30 articles were selected for full-text review. Nine articles met the inclusion criteria for qualitative analysis, including 2 articles that were obtained after screening the references of full-text studies that were reviewed (Figure 1). The main reasons for article exclusion were focus on adaptation of an already existing instrument, and use of procedure- or disease-specific instruments within a nongeneralizable framework.

We identified 9 studies on 8 grading systems for adverse postoperative outcomes with frameworks generalizable to any surgical procedure (Table 1).^{1,8,9,11,18-23} The Clavien–Dindo system and the Common Terminology Criteria for Adverse Events version 3.0 (CTCAE v3.0)¹⁹ were the most commonly cited systems, followed by the Comprehensive Complication Index (CCI).¹² The other systems identified did not appear to be widely incorporated in the literature. All systems focus on complications only, without

consideration of other adverse postoperative outcomes (i.e., sequelae and treatment failure). All instruments except the CCI were produced without including patients' perspectives. In the CCI framework, the severity of complications is obtained based on both patient- and physician-assigned severity scores.

All instruments except the Clavien–Dindo system, CTCAE v3.0 and Plastic Surgery Complication Grading System²¹ allow the derivation of a composite morbidity score ranging from 0 to 100, 1 to 5, 1 to 4, or 1 to 3; however, these scores have limited concrete communicable significance for patients. The CTCAE v3.0 was the only instrument identified that enables grading of adverse medical and radiologic outcomes in addition to adverse surgical outcomes. Length of follow-up or surveillance needed for comprehensive evaluation of adverse surgical events for each instrument was generally not specified in the identified instruments.

Instrument frameworks

The frameworks of the instruments for grading the severity of adverse surgical outcomes are outlined in Table 2. The 1992 Clavien–Dindo framework is based on the invasiveness of the interventions required to address the complication. The 2004 Clavien–Dindo framework¹ is a modification of the 1992 framework, with higher grades of severity associated with life-threatening complications. Disability is no longer a grade on its own but, rather, is highlighted by the suffix “d.” The Plastic Surgery Complication Grading System uses the same framework as the 2004 Clavien–Dindo system but also considers the need for hospital resources (such as length of stay) and postdischarge care (home care) in its framework.

With the Surgical Complication Outcome (SCOUT) score,¹⁸ CTCAE v3.0 and Congenital Heart Disease Morbidity Score (CHDMS),²² a group of clinical experts (surgeons, anesthesiologists and physicians) identify a list of all possible complications for a given surgical procedure and associate severity grades (within a predetermined arbitrary numeric range) to each adverse outcome based on its clinical significance. With the SCOUT score, the experts grade the severity of each identified complication from 0 to 100 based on the subjective concept that if this complication were to happen to them as a patient, how would they rate it in terms of physiologic stress. The composite score is derived by linear summation of individual scores for adverse outcome severity. In the CTCAE v3.0, the severity grading (1–5) is assigned based on symptoms, treatment modality used, change in patients' functionality, and the life-threatening or disability-inducing nature of the complication. In the CHDMS, adverse outcomes are graded by severity from 1 to 4 based on their clinical severity and cost; death is not included or graded in this instrument. The composite score is derived by

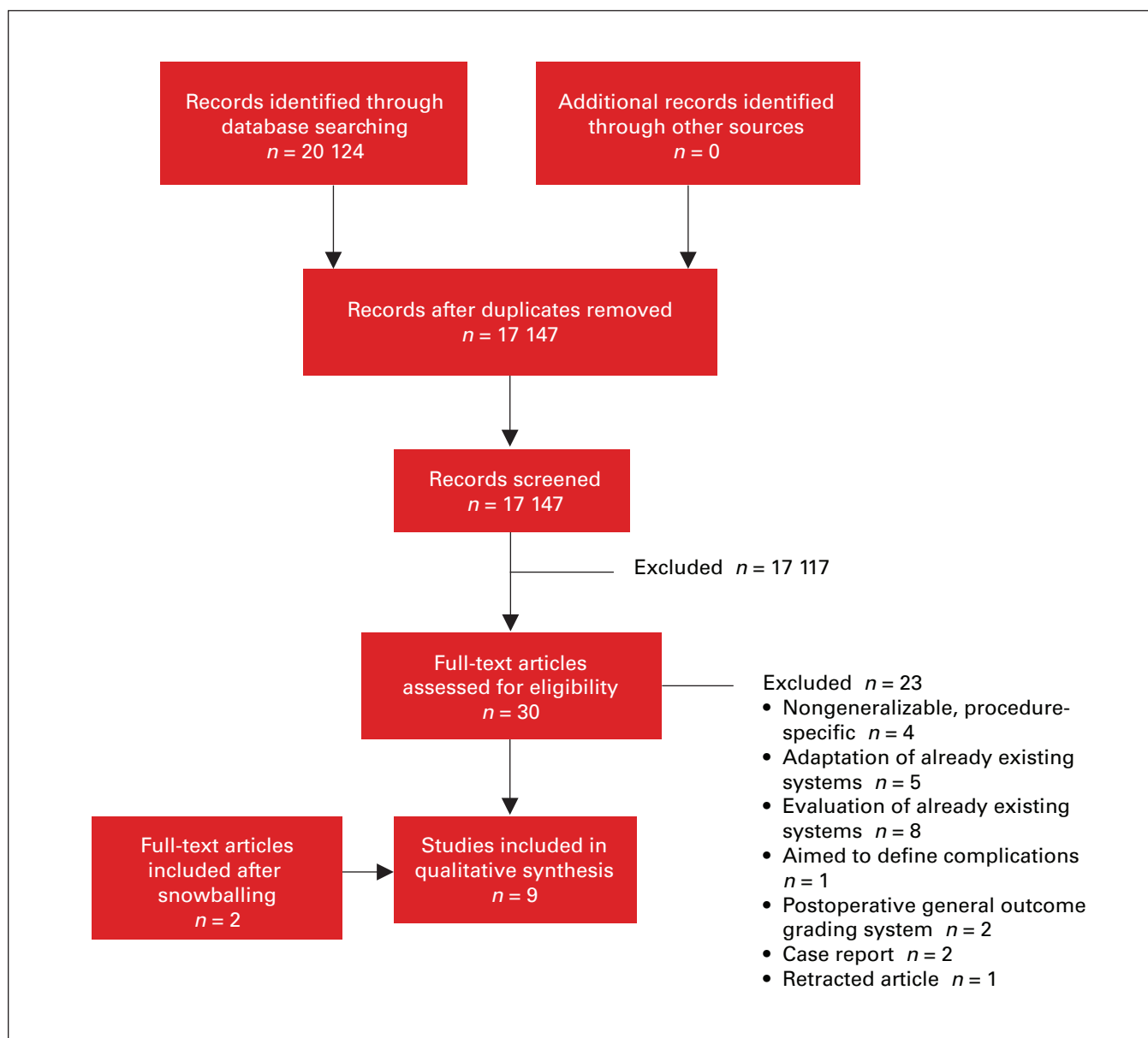


Fig. 1. Flow diagram showing selection of studies for systematic review.

linear summation of individual scores for adverse outcome severity, with a maximum possible composite score of 5.

With the Postoperative Morbidity Index (PMI),²⁰ the severity of a given complication is decided by a panel of surgical experts, who assign severity scores of 0 to 100 to different severity levels using the Accordion Severity Grading System⁹ (2004 Clavien–Dindo system with its grades renamed from I–VI to “minor,” “moderate” and “major”). Consequently, the numeric severity score of a given adverse outcome is determined from its grade in the Accordion Severity Grading System. The PMI uses severity weighting based on the concept of utility weighting, with the value of a given adverse event being based on its severity and duration. The composite severity score of a given procedure is calculated with the following formula:

$$\sum \frac{(\text{severity} \times \text{frequency})}{\text{number of patients}}$$

With the Pediatric Cardiac Surgical Complication Assessment tool,²³ the severity of adverse postoperative outcomes is graded from 0 to 100 based on clinical expert consensus regarding the permanence of the complication. The composite severity score is calculated based on the formula $\sum (\text{frequency} \times \text{severity})$ of adverse postoperative outcomes of a given patient cohort.

In the CCI, a severity rating of 0 to 100 is assigned to each grade of the 2004 Clavien–Dindo system by both patients and physicians. A given grade’s severity is then calculated by means of operation risk index analysis (a methodology from marketing research) by multiplication

Table 1. General characteristics of the included grading systems for adverse surgical outcomes

System; investigator	Characteristic							
	Frequency of citation	Severity grading	Composite score available	Modality assessed	Perspective	Disability reflected in score	Length of follow-up	Generalizable to nonsurgical treatments
Clavien–Dindo, Clavien et al., ^{8,9} 1992	1069	1–5	No	Operation	Health care	Yes	6 wk	No
Surgical Complication Outcome score, Pomposelli et al., ¹⁸ 1997	51	0–100	Yes (patient)	Operation	Health care	No	Until hospital discharge	No
Common Terminology Criteria for Adverse Events version 3.0, Trotti et al., ¹⁹ 2003	2081	1–5	No	Operation, radiation and medical management	Health care	Yes	NS	Yes
Clavien–Dindo, Dindo et al., ¹ 2004	10 649	1–5	No	Operation	Health care	Yes	NS	No
Postoperative Morbidity Index, Porembka et al., ²⁰ 2010	114	0–100	Yes (procedure)	Operation	Health care	No	NS	No
Plastic Surgery Complication Grading System, Qassemyar et al., ²¹ 2010	2	1–5	No	Operation	Health care	Yes	6 mo	No
Congenital Heart Disease Morbidity Score, Sata et al., ²² 2012	9	1–4	Yes (patient)	Operation	Health care	No	NS	No
Pediatric Cardiac Surgical Complication Assessment tool, Shanmugam et al., ²³ 2012	7	1–3	Yes (procedure)	Operation	Health care	No	1 mo	No
Comprehensive Complication Index, Slankamenac et al., ¹¹ 2013	177	0–100	Yes (patient)	Operation	Patient and health care	No	NS	No

NS = not specified.

of the median severity graded by patients and physicians ($\Sigma [median_{phys} \times median_{pat}]$). The raw composite severity score for each patient is obtained by summation of the severity ratings for the adverse outcomes for a given patient. The following formula shows the mathematical transformation of the raw composite score to a normal distribution, giving the CCI a set limit between 0 and 100:

$$\frac{\sqrt{\Sigma (MRV_{phys} \times MRV_{pat})}}{2}$$

where MRV = median reference value.

DISCUSSION

Although surgical sequelae and failure of surgical therapy are permanent causes of morbidity, disability and decreased quality of life, these negative outcomes are left out of the current systems for grading the severity of adverse postoperative outcomes identified in this systematic review.

The majority of the grading systems identified focus on objective clinical and physiologic outcomes during postoperative recovery.²⁴ All the instruments except the CCI rely on clinical experts to grade complications by assigning

severity to given adverse events (SCOUT score), grading the invasiveness of the intervention required to address the complication (Clavien–Dindo system, Plastic Surgery Complication Grading System, PMI and CHDMS) or assessing patient’s function (CTCAE v3.0). However, expert assessment of postoperative outcomes does not necessarily correlate with patient-reported outcomes¹¹ owing to the multidimensional (physiologic, social, psychological and economic) nature of postoperative recovery. This can best be understood by placing patients as the main stakeholders at the centre of weighting of severity of adverse postoperative outcomes to allow for a more comprehensive assessment of their postoperative function, disability and morbidity.^{25,26} In our systematic review, the CCI was the only instrument identified that used the patient’s perspective on postoperative complications (along with the clinical experts’ perspective) in assigning severity grades to a given adverse event.

The SCOUT score, CHDMS and CCI all allow for calculation of postoperative complication composite scores for a given patient. The PMI and Pediatric Cardiac Surgical Complication Assessment tool allow a composite score of complications to be calculated for a given procedure in a patient cohort. The composite scores in the instruments

Table 2. Framework of instruments for grading severity of adverse surgical outcomes

Instrument	Framework
Clavien–Dindo 1992	<ul style="list-style-type: none"> • Whether complication is life-threatening, or causes disability or death. • Invasiveness of intervention required to address complication. • Length of stay.
Surgical Complication Outcome score	<ul style="list-style-type: none"> • 50 postoperative complication types (specific to vascular surgery) were identified from a database and grouped into 9 categories (e.g., vascular, cardiac, pulmonary). • Complications unrelated to vascular surgery (e.g., shoulder pain from pneumoperitoneum) not included. • Each complication type (e.g., pneumothorax) was stratified into 4 severity grades (mild, moderate, severe, death) based on their physiologic significance. • Grading system was used only to help orient the expert panel, who determined the score for each complication using a range from 0 (no complication) to 100 (death); mean score was used. • Reasoning behind the classification was subjective: if this complication were to happen to the experts as a patient, how would they rate it in terms of physiologic stress? • For composite score, individual severity scores are summed.
Common Terminology Criteria for Adverse Events version 3.0	<ul style="list-style-type: none"> • All possible system-based adverse events are identified by multiple multidisciplinary committees, and their severity is graded from 1 to 5 based on <ul style="list-style-type: none"> ◊ Symptoms ◊ Treatment modality used ◊ Effects on functionality/activities of daily living ◊ Life-threatening or disability-inducing
Clavien–Dindo 2004	<ul style="list-style-type: none"> • Modifications to 1992 Clavien–Dindo system. • Length of stay no longer a criterion for grading owing to differences across centres. • Life-threatening complications have higher grades associated with them. • Disability no longer a grade on its own but is now highlighted by the suffix “d” (for disability). Thus, any grade of complication may be supplemented with this information.
Postoperative Morbidity Index	<ul style="list-style-type: none"> • Utility weighting used for calculating severity (useful for aggregating multidimensional scales into a single unified measure of overall impact; severity = utility × duration). • Panel of surgical experts employed to derive a set of severity weights for the grades of the Accordion Severity Grading system between 0 and 100. The average of these scores was then associated to each grading level as its severity weight. • Composite score for each procedure (Postoperative Morbidity Index) is then calculated with the following formula: $\sum \frac{(\text{severity} \times \text{frequency})}{\text{Number of patients}}$ <p>Example: 1857 National Surgical Quality Improvement Program patients experienced a total complication weight of 16 238 severity points (on a 100-point scale). On average, each patient had a severity rating of $16\,238/1857 = 8.7$. This can be considered analogous to a loss of utility of 8.7% (on a scale of 1 to 100).</p>
Plastic Surgery Complication Grading System	<ul style="list-style-type: none"> • Similar to Clavien–Dindo 2004, which considers interventions to treat complications. • Also incorporates length of stay, resource use (home care) and disability.
Congenital Heart Disease Morbidity Score	<ul style="list-style-type: none"> • Team consisting of pediatric anesthetist, cardiac intensivist and cardiac surgeon drew up a list of conditions or diseases that can affect patients in the course of congenital heart surgery or increase costs of hospital stay. • Conditions/diseases were assigned a severity score between 1 and 4 (need for mechanical ventilation for more than 7 days) according to estimated severity or costliness. If no complication, score of 0.5 assigned. • Complication scores are added linearly; if the sum is more than 5, score of 5 is assigned. • Developers chose to not include hospital death among complication as they believed that concepts of mortality and morbidity should not be mixed and that complications leading to death should be evaluated separately.
Pediatric Cardiac Surgical Complication Assessment tool	<ul style="list-style-type: none"> • Complications arising from congenital heart surgery at IWK Health Centre over 3 years were compiled. • Each complication was weighted by assigning a severity coefficient of 1, 2 or 3 according to expert consensus, where 1 = mild or temporary deficit predicted to affect patient’s course minimally, and 3 = severe or permanent complication predicted to impair patient’s outcome significantly and possibly precede death. • Morbidity burden or total magnitude of complication for a given procedure calculated as follows: $\sum (\text{frequency} \times \text{severity})$
Comprehensive Complication Index	<ul style="list-style-type: none"> • Clinical experts and patients assign severity score of 0–100 to complication scenarios involving all grades of Clavien–Dindo system. • Severity of a given grade is calculated by means of operation risk index analysis by multiplication of the median severity graded by patients and physicians (a summative approach is not used because it ignores the gravity of high-grade complications). • Raw composite score obtained by summation of severity ratings for complications for a given patient. • To facilitate clinical applicability of Comprehensive Complication Index, the developers tested different mathematical transformations of the raw composite score in 3 years of cohort data to find a distribution of the index that is closest to normal distribution, which would give the CCI a set limit between 0 and 100. • The following is the mathematical transformation that worked (where MRV = median reference value): $\sqrt{\frac{\sum (MRV_{phys} \times MRV_{pat})}{2}}$

identified in this systematic review are obtained through different methods. The SCOUT score and CHDMS use linear summation of severity scores, whereas the Pediatric Cardiac Surgical Complication Assessment tool enables derivation of a composite severity score for a given procedure after accounting for the frequency of occurrence of these complications. However, simply summing complication severity scores places too much weight on adverse events of minor and moderate severity (when happening concurrently), hence producing a composite score with an inappropriately high value.¹² The CCI uses operation risk index analysis to synthesize patient and physician perspectives of severity appropriately.¹² It then uses a mathematical formula that transforms the composite score into a normal distribution with lower and upper limits of 0 and 100, respectively, which accounts for the possibility of an unlimited number of adverse outcomes per patient. To facilitate the use of the CCI, given its relatively complicated mathematical formula, a user-friendly online CCI calculator has been created (www.assessurgery.com).¹²

The PMI uses severity weighting based on the concept of utility weighting, a well-established, standard method of assigning weights to multidimensional outcome states to reflect their overall impact (severity and duration).^{20,27,28} Strasberg and Hall²⁹ reported quantitative morbidity scores for several abdominal procedures using this severity score and data from the American College of Surgeons' National Surgical Quality Improvement Program.²⁰ However, unlike the CCI, the PMI composite score is calculated by linear summation of the severity scores for all individual complications. Therefore, owing to the possibility of an unlimited number of adverse outcomes per patient, one cannot define a maximum numeric value for the PMI score. Hence, one would be unable to mathematically calculate statistically significant differences between different composite scores.²⁹

Although some of the composite scores in these instruments can be used by clinicians for research and quality control, the numeric value of each score has limited concrete communicable significance for patients and does not translate into a clinically meaningful concept easily. The PMI expresses its composite score as a percentage of utility loss, which, in the absence of any indication of duration, is also difficult to interpret at the individual patient level.

The increasing use of nonoperative treatment in fields such as trauma and pediatric surgery in the past decades^{15,30,31} calls for the development of grading systems that allow evaluation of adverse outcomes other than postoperative complications. Although this issue falls outside the scope of the present study and its search strategy, we observed that the CTCAE v3.0 was the only instrument that allows the grading of adverse events associated with nonoperative treatment. Although developed by the National Cancer Institute for cancer treatment, its framework is generalizable to other fields. The framework

involves identification and grading of all possible system-based adverse events, allowing it to quantify and compare the adverse symptoms experienced by the patient after operative and nonoperative treatment. It also offers standard nomenclature and definitions to standardize the reporting process. Versions 4.0 and 5.0 have been published on the National Institutes of Health website,³² with updates on adverse outcome definitions, although the framework has remained the same.

Based on our findings, the ideal attributes of a grading system for adverse postoperative outcomes include, but are not limited to, 1) the ability to take into account all the adverse postoperative events that can affect quality of life, including sequelae, procedure failure and complications; 2) inclusion of patient-centred weighting of both the duration and the severity of adverse postoperative outcomes, allowing for a more comprehensive assessment of patients' postoperative function, disability and morbidity as experienced by them; 3) the ability to generate a composite severity score for all adverse events, enabling a better understanding of the global morbidity associated with a given procedure; and 4) the ability to grade adverse outcomes of both operative and nonoperative treatment, enabling comparison of morbidity after different treatment modalities.

Limitations

The strengths of our study are its broad literature search strategy, snowballing technique and absence of any language restrictions, which allowed us to ensure identification of most pertinent studies.

Our findings, however, were restricted by the small number of studies identified for inclusion. During title and abstract review, the majority of the articles excluded were ones that used instruments for grading the severity of negative postoperative outcomes as simple outcome measures. Furthermore, most of the identified articles were excluded after a full-text review because the severity grading instruments were symptom- or procedure-specific, without a logical framework. This finding is not a limitation but, rather, shows the paucity of relevant studies in the literature. Our search strategy did not include the grey literature before 2012. As a result, we may have overlooked severity grading systems developed by clinical societies or governmental bodies outside the traditional academic publishing routes. Most of the included studies do not focus on the psychometric properties of the instruments, and this also lies beyond the scope of our review.

CONCLUSION

Our review identified several efforts to create "ideal" systems for grading the severity of adverse postoperative outcomes. Each instrument offered its own advantages.

However, none appeared able to meet the need for a patient-centred instrument capable of generating a composite score of all possible adverse postoperative outcomes (including the morbidity caused by surgical sequelae and procedure failure), and enabling comparisons of noninterventional and surgical management of disease. The benefit of such grading systems will be in facilitating physician–patient communication. The CCI has valuable features that should be highlighted. It encompasses both provider and patient perspectives, and enables calculation of a composite score of all postoperative complications. However, the composite score is a pure numeric value, devoid of any significance to the patient. Despite centuries of treating patients surgically and decades of using complication scores to evaluate surgical treatment, the ideal of a patient-based, comprehensive score for adverse surgical outcomes remains elusive. Research efforts aimed at merging patient-reported and patient-valued outcomes with postoperative complications will facilitate the much-needed process of fostering patient-centred surgical care.

Affiliations: From the Division of Paediatric General and Thoracic Surgery, McGill University, Montréal, Que. (Balvardi, St-Louis, Yousef, Toobaie, Poenaru); the McConnell Resource Centre, McGill University Health Centre, Montréal, Que. (Guadagno); and the Division of Pediatric General Surgery, University of British Columbia, Vancouver, BC (Baird).

Competing interests: None declared.

Contributors: All the authors designed the study. S. Balvardi and E. Guadagno acquired the data, which S. Balvardi and D. Poenaru analyzed. S. Balvardi wrote the article, which all authors critically revised. All authors gave final approval of the article to be published.

Content licence: This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY-NC-ND 4.0) licence, which permits use, distribution and reproduction in any medium, provided that the original publication is properly cited, the use is noncommercial (i.e., research or educational use), and no modifications or adaptations are made. See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

Funding: No funding was received for this work.

References

- Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 2004;240:205-13.
- Martin RC 2nd, Brennan MF, Jaques DP. Quality of complication reporting in the surgical literature. *Ann Surg* 2002;235:803-13.
- Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. *N Engl J Med* 2011;364:2128-37.
- Brennan MF, Radzyner M, Rubin DM. Outcome — more than just operative mortality. *J Surg Oncol* 2009;99:470-7.
- Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility in cholecystectomy. *Surgery* 1992;111:518-26.
- Joly F, Brune D, Couette JE, et al. Health-related quality of life and sequelae in patients treated with brachytherapy and external beam irradiation for localized prostate cancer. *Ann Oncol* 1998;9:751-7.
- Cohen M, Guger S, Hamilton J. Long term sequelae of pediatric craniopharyngioma — literature review and 20 years of experience. *Front Endocrinol (Lausanne)* 2011;2:81.
- Clavien PA, Barkun J, de Oliveira ML, et al. The Clavien–Dindo classification of surgical complications: five-year experience. *Ann Surg* 2009;250:187-96.
- Strasberg SM, Linehan DC, Hawkins WG. The Accordion Severity Grading System of surgical complications. *Ann Surg* 2009;250:177-86.
- Clavien PA, Sanabria JR, Mentha G, et al. Recent results of elective open cholecystectomy in a North American and a European center. Comparison of complications and risk factors. *Ann Surg* 1992;216:618-26.
- Winslow E, Siy A, Stafford LC, et al. Examining the severity of postoperative complications from the perspective of the patient. *J Am Coll Surg* 2015;221:e27-36.
- Slankamenac K, Graf R, Barkun J, et al. The comprehensive complication index: a novel continuous scale to measure surgical morbidity. *Ann Surg* 2013;258:1-7.
- Santucci RA, Fisher MB. The literature increasingly supports expectant (conservative) management of renal trauma — a systematic review. *J Trauma* 2005;59:493-503.
- Oumar N, Dominique F, Nikola K, et al. Results of non-operative management of splenic trauma and its complications in children. *J Indian Assoc Pediatr Surg* 2014;19:147-50.
- Stawicki SP. Trends in nonoperative management of traumatic injuries — a synopsis. *Int J Crit Illn Inj Sci* 2017;7:38-57.
- Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1-34.
- Popay J, Roberts H, Sowden A, et al. *Guidance on the conduct of narrative synthesis in systematic reviews: a product from the ESRC Methods Programme. Version 1.* Swindon (UK): Economic and Social Research Council; 2006.
- Pomposelli JJ, Gupta SK, Zacharoulis DC, et al. Surgical complication outcome (SCOUT) score: a new method to evaluate quality of care in vascular surgery. *J Vasc Surg* 1997;25:1007-15.
- Trotti A, Colevas AD, Setser A, et al. CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. *Semin Radiat Oncol* 2003;13:176-81.
- Porembka MR, Hall BL, Hirbe M, et al. Quantitative weighting of postoperative complications based on the Accordion Severity Grading System: demonstration of potential impact using the American College of Surgeons National Surgical Quality Improvement Program. *J Am Coll Surg* 2010;210:286-98.
- Qassemayr Q, Sinna R. Classification of plastic surgery complications: proposal for an objective evaluation of publications and professional practices [article in French]. *Ann Chir Plast Esthet* 2010;55:561-7.
- Sata S, Haun C, Weber T, et al. A morbidity score for congenital heart surgery based on observed complications. *Eur J Cardiothorac Surg* 2012;41:898-904.
- Shanmugam G, Clark LL, Burton HJ, et al. Improving and standardizing capture of pediatric cardiac surgical complications. *J Thorac Cardiovasc Surg* 2012;144:570-6.
- Korolija D, Wood-Dauphinee S, Pointner R. Patient-reported outcomes. How important are they? *Surg Endosc* 2007;21:503-7.
- Bowyer A, Jakobsson J, Ljungqvist O, et al. A review of the scope and measurement of postoperative quality of recovery. *Anaesthesia* 2014;69:1266-78.
- Stomberg MW, Saxborn E, Gambreus S, et al. Tools for the assessment of the recovery process following discharge from day surgery: a literature review. *J Perioper Pract* 2015;25:219-24.
- Drummond MF, Sculpher MJ, Claxton K, et al. *Methods for the economic evaluation of health care programmes.* 4th ed. Oxford (UK): Oxford University Press; 2015.

28. The world health report 2000. Health systems: improving performance. Geneva: World Health Organization; 2000. Available: www.who.int/whr/2000/en/ (accessed 2018 Oct. 6).
29. Strasberg SM, Hall BL. Postoperative morbidity index: a quantitative measure of severity of postoperative complications. *J Am Coll Surg* 2011;213:616-26.
30. Taniguchi T, Matsui H, Fujita T, et al. Structure and expression of a cloned cDNA for human interleukin-2. *Nature* 1983;302:305-10.
31. Leppaniemi A. Nonoperative management of solid abdominal organ injuries: from past to present. *Scand J Surg* 2019;108:95-100.
32. Common Terminology Criteria for Adverse Events (CTCAE). Version 5.0. Bethesda (MD): National Cancer Institute, U.S. Department of Health and Human Services; 2017. Available: https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm#ctc_50 (accessed 2018 Oct. 6).