Heliyon

Operating room staff and surgeon documentation curriculum improves wound classification accuracy

Joseph W. Gorvetzian [a], Katharine E. Epler [a], Samuel Schrader [a], Joshua M. Romero [a], Ronald Schrader [b], Alissa Greenbaum [c], Rohini McKee [c,*]

[a] *University of New Mexico, School of Medicine, Albuquerque, NM, USA*

[b] *RMS Biostatistics Services, Albuquerque, NM, USA*

[c] *University of New Mexico, Department of Surgery, Albuquerque, NM, USA*

* Corresponding author.
E-mail address: RMcKee@salud.unm.edu (R. McKee).

Abstract

Background: Misclassification of wounds in the operating room (OR) can adversely affect surgical site infection (SSI) reporting and reimbursement. This study aimed to measure the effects of a curriculum on documentation of surgical wound classification (SWC) for operating room staff and surgeons.

Methods: Accuracy of SWC was determined by comparing SWC documented by OR staff during the original operation to SWC determined by in-depth chart review. Patients 18 years or older undergoing inpatient surgical procedures were included. Two plan-do-act-study (PDSA) cycles were implemented over the course of 9 months. A total of 747 charts were reviewed. Accuracy of SWC documentation was retrospectively assessed across 248 randomly selected surgeries during a 5-week period prior to interventions and compared to 244 cases and 255 cases of post-intervention data from PDSA1 and PDSA2, respectively. Changes in SWC accuracy were assessed pre- and post-intervention using the kappa coefficient. A p-value for change in agreement was computed by comparing pre- and post-intervention kappa.

Results: Inaccurate documentation of surgical wound class decreased significantly following curriculum implementation (kappa improved from 0.553 to 0.739 and 0.757; p = 0.001). Classification accuracy improved across all wound classes; however, class III and IV wounds were more frequently misclassified than class I and II wounds, both before and after the intervention.

Conclusion: Implementation of a multidisciplinary documentation curriculum resulted in a significant decrease in SWC documentation error. Improved accuracy of SWC reporting may facilitate a better assessment of SSI risk in a complex patient population.

Keywords: Surgery, Education

1. Introduction

Surgical site infections (SSIs) are a significant cause of morbidity and mortality in health care. Of the 15 million procedures performed each year in the United States, 300,000 to 500,000 are associated with SSIs postoperatively [1]. The consequences of SSIs are manifold: they contribute to longer hospital stays, increase mortality risks by 2−11 fold, and cost the US healthcare system approximately $10 billion each year [1, 2, 3]. Hospital care costs have been estimated to be 1.43 to 1.93 times greater for postsurgical patients who develop a SSI [4]. The pressure for performance-based reimbursement highlights the importance of factors that are relevant to SSIs, with significant clinical, administrative, and economic implications [3, 5, 6]. Due to their profound impact on the US healthcare system, interventions aimed at reducing the risk of SSI occurrence have been implemented using perioperative protocols that are often stratified by risk.

Surgical wound classification (SWC) is an important predictor of postoperative surgical site infections and can be used to guide protocol development and surgical decision-making [7]. The SWC system has been used in the US for over half a century [5, 8, 9, 10]. Wounds are classified into one of the four categories: (I) clean, (II) clean-contaminated, (III) contaminated, and (IV) dirty-infected. For each designation the postoperative risk of SSIs has been shown to be 1−5%, 3−11%, 10−17%, and over 27%, respectively, though some studies question these percentages and reported rates of infection vary across hospitals [7, 11, 12, 13].

Inaccurate classification of wounds is a pervasive problem that may contribute to these observed inconsistencies. Two primary factors implicated in misclassification are a lack of understanding of appropriate wound classification by the operating room nurse and lack of participation on the part of the surgeon. Significant variation exists in assigning classification both among providers and between physicians and

nursing staff [5, 14, 15, 16, 17, 18]. Some studies have demonstrated discordance on up to 92% of cases [17]. While the true utility of the SWC system is the subject of some debate [5, 8, 9, 10, 13, 19, 20, 21], it does provide a relatively simple tool that enables analysis of clinical and economic outcomes that drive quality improvement initiatives and thus improved patient care [5, 16, 22]. However, if misclassification of surgical wounds is indeed widespread, models incorporating SWC as a factor may be rendered inaccurate.
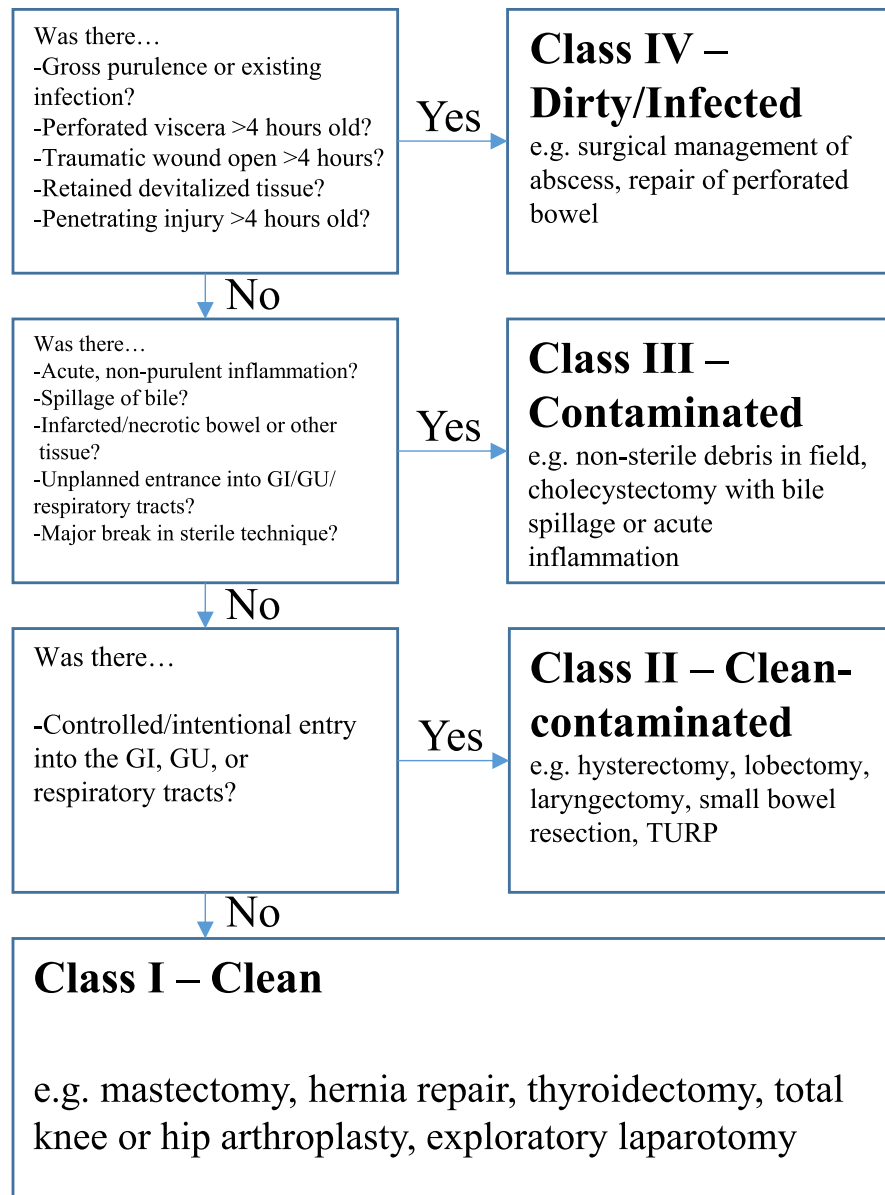
Given that surgical wound misclassification is well-documented in the literature and that SWC is a tenet of risk stratification models, the present study sought to establish a baseline for SWC accuracy at our institution. We then compared this pre-intervention data to SWC accuracy following implementation of targeted curricular interventions designed to increase education and communication with regard to SWC documentation over the course of multiple plan-do-act-study (PDSA) cycles.

2. Materials & methods

2.1. Data collection

IRB approval for medical record review was obtained for performing this study, which was HIPAA compliant. A REDCap (https://www.project-redcap.org/) survey form was created based on well-established wound classification algorithms (Fig. 1) for all data entry. For each time period (pre-intervention, PDSA1, and PDSA2), charts were gathered, dental and endoscopic cases were excluded, and a random selection of charts were chosen for retrospective analysis. Randomization was performed with an online randomizing tool (https://www.randomizer.org/). A power analysis was performed to ensure adequate case volumes were reviewed to achieve statistical significance.

Operative note review yielded an algorithmically defined Correct SWC which was compared to nursing-documented OR SWC. Pre-intervention data for establishing baseline SWC discordance was based on 1,783 cases performed at the UNM Hospital Main Operating Room (UNMH Main OR) from between May 1st, 2016−June 30th, 2016, with 248 cases randomly selected for review. PDSA1 data was based on 955 cases performed at the UNMH Main OR from November 29th, 2016−December 31st, 2016 with 244 of these cases randomly selected for review. Finally, PDSA2 data was based on 1,259 cases performed at UNM Main OR from between April 1st, 2017−May 3rd, 2017, with 255 of these cases randomly selected for review. In cases where multiple procedures were performed in a single operation, the highest assigned OR SWC was used to compare to the algorithmically-defined Correct SWC for the case.

Was there…
-Gross purulence or existing
infection?
-Perforated viscera >4 hours old?
-Traumatic wound open >4 hours?
-Retained devitalized tissue?
-Penetrating injury >4 hours old?

Yes →

**Class IV –
Dirty/Infected**
e.g. surgical management of
abscess, repair of perforated
bowel

↓ No

Was there…
-Acute, non-purulent inflammation?
-Spillage of bile?
-Infarcted/necrotic bowel or other
 tissue?
-Unplanned entrance into GI/GU/
respiratory tracts?
-Major break in sterile technique?

Yes →

**Class III –
Contaminated**
e.g. non-sterile debris in field,
cholecystectomy with bile
spillage or acute
inflammation

↓ No

Was there…

-Controlled/intentional entry
into the GI, GU, or
respiratory tracts?

Yes →

**Class II – Clean-
contaminated**
e.g. hysterectomy, lobectomy,
laryngectomy, small bowel
resection, TURP

↓ No

Class I – Clean

e.g. mastectomy, hernia repair, thyroidectomy, total
knee or hip arthroplasty, exploratory laparotomy

Fig. 1. Wound classification algorithm. Class I wounds are generally clean and without inflammation; Class II wounds involve intentional violation of tracts without infection or acute inflammation present; Class III involve acute non-purulent inflammation as seen in acute appendicitis or presence of non-sterile object in the field; Class IV entails existing clinical infection. *TURP: Trans-urethral prostatectomy.*

2.2. Interventions

Interventions began in July 2016 and continued through March 2017. PDSA cycle 1 interventions consisted of meetings with administrative and OR staff promoting education regarding classification meanings, sessions for surgeons emphasizing the importance of communication with nursing colleagues, dispersion of portable SWC algorithm cards to OR nursing staff, and the hanging of SWC algorithm

posters along the hallways of the all operating rooms. PDSA cycle 2 consisted of a change to the electronic medical record (EMR) such that the previously auto-populated wound classification field would be left blank, thereby requiring OR nursing staff to manually enter SWC for the case.

2.3. Statistical methods

Agreement between Correct SWC and OR SWC was assessed separately for each of the intervention time periods (Pre-intervention, PDSA1, and PDSA2) using Cohen's weighted kappa coefficient; weighted kappa is generally preferred with ordered classes as in the present study, but results were almost identical using a simple kappa coefficient. The three pairwise comparisons of weighted kappa between time periods was constructed as a simple z-statistic, for instance comparing Pre-intervention to PDSA1 using $z = (\kappa_{PRE} - \kappa_{PDSA1})/\sqrt{SE_1^2 + SE_2^2}$, with a Bonferroni correction applied for the multiple comparisons. For each time period and Correct classification (I − IV), the percent classified in the OR in each of I−IV was calculated and graphed (Fig. 3). Calculations were performed using SAS version 9.4.

3. Results

Correct wound classification assignments for the Pre-intervention data consisting of 248 cases included 93 class I wounds (37.5%), 64 class II wounds (25.8%), 19 class III wounds (7.7%), and 72 class IV wounds (29%). For the 244 PDSA1 cases, Correct SWCs included 103 class I (42.2%), 64 class II (26.2%), 20 class III (8.2%), and 57 class IV (23.4%). For the 255 PDSA2 cases, Correct SWCs identified 99 class I wounds (38.8%), 57 class II wounds (22.4%), 32 class III wounds (12.5%), and 67 class IV wounds (26.3%) (Table 1).

Average overall percentage misclassification between nursing-assigned OR SWC and Correct SWC in the pre-intervention period was 33%, 20% in PDSA1, and

Table 1. Number of cases and corresponding algorithmically-derived correct wound classification in pre- and post-intervention data.

Sample Size (%)

SWC	Pre-Intervention	PDSA 1	PDSA 2
I	93 (37.5)	103 (42.2)	99 (38.8)
II	64 (25.8)	64 (26.2)	57 (22.4)
III	19 (7.7)	20 (8.2)	32 (12.5)
IV	72 (29.0)	57 (23.4)	67 (26.3)
Total	248	244	255

24% in PDSA2 (Fig. 2). Weighted kappa (SE) for pre-intervention, PDSA1, and PDSA2 were 0.553 (.043), 0.739 (0.036), and 0.757 (0.030) respectively, so that both post-intervention periods show greater SWC agreement than the pre-intervention. P-values for comparison of pre-intervention data to both PDSA1 and PDSA2 were less than 0.001, while p-value for comparison of PDSA1 to PDSA2 was 0.73, so that after applying the Bonferroni correction weighted kappa in both post-intervention periods were significantly different from that for pre-intervention, but weighted kappa in the two post-intervention periods did not differ significantly, all at a 0.05 significance level. For class I wounds, the nursing-assigned OR SWC was generally good but was improved post-intervention; for class II wounds there was an improvement in PDSA1 but in PDSA2 several were misclassified as I; for class III wounds very few were correctly classified in the OR pre-intervention, but there was a substantial improvement by PDSA2 (although most still were misclassified); and for class IV wounds there was substantial improvement post-intervention though there remained a large misclassification as was present in the pre-intervention OR SWC (Fig. 3).

Delineation of the 747 total cases by surgical service demonstrated the majority of cases analyzed were performed by the Orthopedics (23.4%), General Elective surgery (which included general oncology cases; 18.3%), and Neurosurgery (11.5%) departments (Table 2). Though cases were randomized prior to analysis, there was almost a two-fold increase in the number of Emergency General surgery cases and half the number of Plastic surgery cases reviewed in PDSA2 than in the prior two periods.
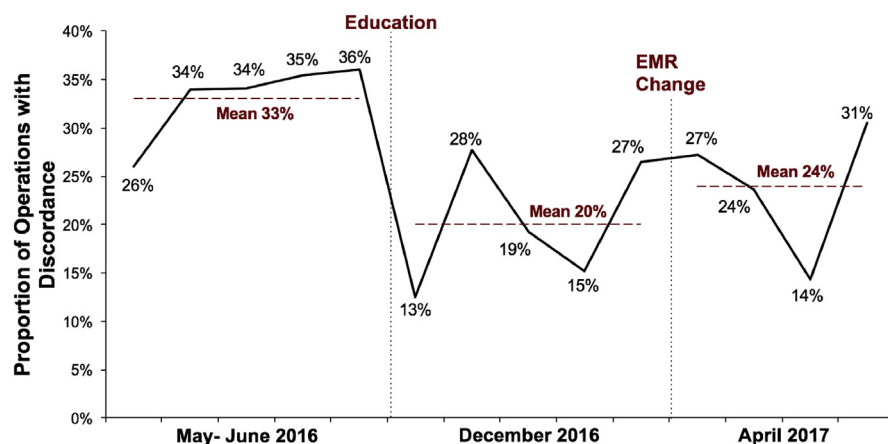
Fig. 2. Average discordance rates between OR wound classification documented in nursing record versus algorithmically-defined Correct surgical wound class. Average discordance was 33% in baseline data from May 2016−June 2016 compared to 20% in PDSA1 data and 24% in PDSA2 data. Kappa improved from 0.553 in pre-intervention to 0.739 and 0.757 in PDSA1 and PDSA2, respectively (p = 0.001). Data from both post-intervention periods demonstrated significantly increased concordance between OR SWC and Correct SWC compared to pre-intervention data but were not significantly different from each other (p = 0.05).

Surgical Wound Classification Summary
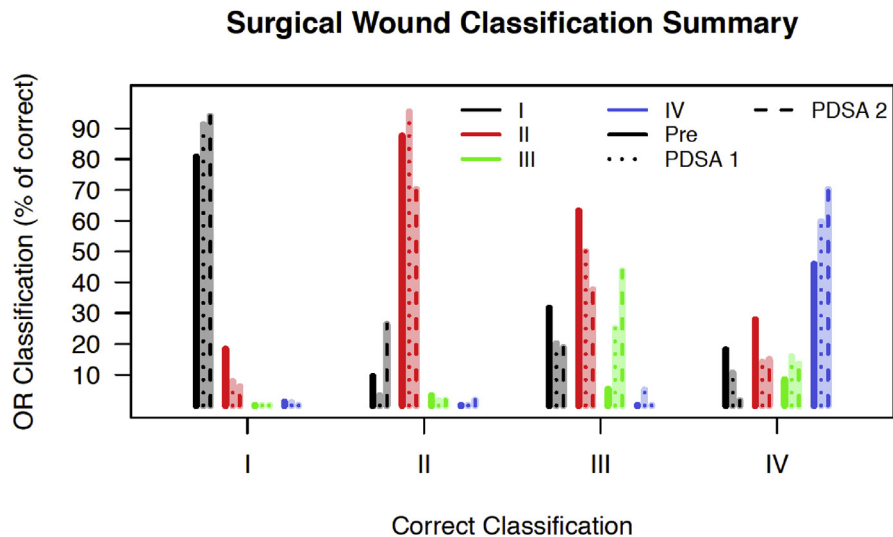
Fig. 3. Distribution of correct OR SWC frequencies for specific wound class in pre- and post-intervention data. Solid lines indicate pre-intervention data; small dotted lines indicate PDSA1 data; thick dotted lines indicate PDSA2 data, and each color represents a specific wound class. For all wound classifications, frequency of correct classification increased in both PDSA1 and PDSA2 data with exception of class II wounds (which were more frequently misclassified as class I wounds in PDSA2) and class IV wounds (which were more frequently misclassified as class III in PDSA2). Class III wounds exhibit the highest overall heterogeneity both pre- and post-intervention.

4. Discussion

Inaccurate and inconsistent classification of surgical wounds has proven to be a widely reported problem in institutions across the country [5, 14, 16, 17, 18, 23].

Table 2. Number of cases by surgical service per period of data collection.

Number of Cases (%)

Service	Pre-intervention	PDSA1	PDSA2	Total
General Elective	47 (18.6)	42 (17.2)	48 (18.9)	137 (18.3)
Emergency General	19 (7.7)	18 (7.4)	40 (15.7)	77 (10.3)
Neurosurgery	20 (8.1)	29 (11.9)	37 (14.5)	86 (11.5)
Orthopedic	51 (20.6)	66 (27.1)	58 (22.7)	175 (23.4)
Obstetrics-Gynecology	18 (7.3)	17 (7.0)	17 (6.7)	52 (7.0)
ENT	33 (13.3)	28 (11.5)	22 (8.6)	83 (11.1)
Plastics	10 (4.0)	9 (3.7)	4 (1.6)	23 (3.1)
Urology	21 (8.5)	11 (4.5)	11 (4.3)	43 (5.8)
Vascular	17 (6.9)	15 (6.2)	10 (3.9)	42 (5.6)
Cardiothoracic	8 (3.2)	5 (2.1)	5 (2.0)	18 (2.4)
Trauma	4 (1.6)	4 (1.6)	3 (1.2)	11 (3.9)
Total	248	244	255	747

Currently, the continued use of wound classification in performance risk-adjustment models for hospitals is under evaluation [9], but given the current use of SWC in these models, it is important to ensure accurate SWC classification. With this study, we found that an average of one out of every three surgical wounds was being misclassified at our institution. However, following implementation of a multifaceted education curriculum incorporating both operating room nurses and surgeons, a statistically significant 13% decrease in SWC inaccuracy rates was realized following a single PDSA cycle (Fig. 1), and this increase in SWC concordance was maintained through the second PDSA cycle. The efficacy of PDSA2 is particularly salient as the effects of the educational interventions of PDSA1 may be more ephemeral compared to the hard-coded change in the EMR. These findings contribute to the growing body of evidence demonstrating how even relatively small interventions aimed at increasing education and communication regarding SWC can have meaningful results [5, 10, 16, 22, 23, 24, 25, 26].

The data presented here also demonstrates some interesting trends with regard to which particular wound classes tend to be most often misclassified. It is rare for class I or class II wounds to be mistakenly assigned a higher classification (i.e. overclassification of low-class wounds is uncommon). It is much less rare for higher class wounds to be mistakenly attributed a lower class (i.e. underclassification of high class wounds is not uncommon). Overall, there appears to be much more heterogeneity, and therefore uncertainty, in the correct classification of class III and class IV wounds. However, by PDSA2 there was substantial improvement in SWC concordance in both of these difficult classes. These finding suggest a phenomenon of persistent underclassification of higher class wounds, which may have implications for patient care and for risk stratification models that incorporate SWC as a factor such as the ACS National Surgical Quality Improvement Program. Given the potentially significant differences in postoperative SSI risk among the different classes, it may be particularly beneficial to emphasize the proper delineation of what constitutes a class III or class IV wound in education-oriented curricular interventions.

There are many options for other interventions that may help improve SWC classification. While our institution realized a significant improvement in SWC discordance rates, there is certainly room for further improvement. This fact is perhaps best suggested by the increased discordance rate in PDSA2 versus PDSA1 (24% compared to 20%, respectively; Fig. 2). Though there was no statistical difference between these intervention periods, the slight uptrend in discordance rate hints at the difficulty of enacting sustainable change. A multifaceted and multitemporal approach would likely render the most benefit in this regard. Future interventions may include regular 'check-ins' (e.g. quizzes or audits) or incentives for surgical divisions with lowest discordance rates. Additionally, numerous studies have shown that adding wound classification to an existing checklist or implementing an SWC-oriented debriefing tool can promote accurate classification [10, 22, 23, 24, 25, 26]. Importantly,

this intervention appears particularly effective for improving accuracy of contaminated and dirty wound classification. Perhaps with increased SWC accuracy, efforts aimed at reducing SSIs can be better stratified for at-risk patients, ultimately facilitating reduced morbidity and mortality in the postoperative period.

There are some limitations to this study. There were relatively long time periods between our baseline data and each of our subsequent post-intervention data, during which some other factor besides our intervention may have concurrently influenced the decrease in SWC discordance. Additionally, this was a single-institution study and the overall time-frame of this study is relatively short. Future studies should have more regular intervals of data collection and more longitudinal follow up to assess the longevity of interventional effectiveness. Finally, future studies may wish to individually assess rates of SWC discordance among different surgical specialties for more targeted quality improvement interventions.

5. Conclusions

Misclassification of surgical wounds is ubiquitous. Though the utility of the current surgical wound classification scheme is a matter of debate, its continued use in quality-related metrics dictates a need for reliable SWC documentation. We found particular curricular interventions at our institution effected a significant and sustained increase in SWC accuracy across all wound classes; however, higher wound classes tended to be more heterogeneously classified and underclassified, presenting an area in which targeted interventions should be considered.

Declarations

Author contribution statement

Joseph W. Gorvetzian, Katharine E. Epler, Samuel Schrader, Joshua M. Romero: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Ronald Schrader: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Alissa Greenbaum: Analyzed and interpreted the data; Wrote the paper.

Rohini McKee: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

[1] D.J. Anderson, Surgical site infections, Infect. Dis. Clin. N. Am. 25 (2011) 135−153.

[2] R.G. Gaston, M.A. Kuremsky, Postoperative infections: prevention and management, Hand Clin. 26 (2) (2010) 265−280.

[3] G. de Lissovoy, K. Raeman, V. Hutchins, et al., Surgical site infection: incidence and impact on hospital utilization and treatment costs, Am. J. Infect. Control 37 (5) (2009) 387−397.

[4] M.L. Schweizer, J.J. Cullen, E.N. Perencevich, et al., Costs associated with surgical site infections in Veterans Affairs hospitals, JAMA Surg. 149 (6) (2014) 575−581.

[5] L.R. Putnam, S.M. Levy, M.L. Blakely, et al., A multicenter, pediatric quality improvement initiative improves surgical wound class assignment, but is it enough? J. Pediatr. Surg. 51 (4) (2016) 639−644.

[6] L.S. Kao, A.A. Ghaferi, C.Y. Ko, J.B. Dimick, Reliability of superficial surgical site infections as a hospital quality measure, J. Am. Coll. Surg. 213 (2) (2011) 231−235.

[7] G. Ortega, D.S. Rhee, D.J. Papandria, et al., An evaluation of surgical site infections by wound class system using the ACS-NSQIP, J. Surg. Res. 174 (1) (2012) 33−38.

[8] L.M. Mioton, S.W. Jordan, P.J. Hanwright, et al., The relationship between preoperative wound classification and postoperative infection: a multi-institutional analysis of 15,289 patients, Arch. Plast. Surg. 40 (5) (2013) 522−529.

[9] M.H. Ju, M.E. Cohen, K.Y. Bilimoria, et al., Effect of wound classification on risk adjustment in American College of Surgeons NSQIP, J. Am. Coll. Surg. 219 (3) (2014) 371−381.e5.

[10] L.R. Putnam, S.M. Levy, G. Holzmann-Pazgal, et al., Surgical wound classification for pediatric appendicitis remains poorly documented despite targeted interventions, J. Pediatr. Surg. 50 (6) (2015) 915−918.

[11] E.M. Ferraz, T.S. Bacelar, J.L.A. Aguiar, et al., Wound infection rates in clean surgery: a potentially misleading risk classification, Infect. Control Hosp. Epidemiol. 13 (8) (1992) 457−462.

[12] D. Eisenberg, Surgical site infections: time to modify the wound classification system? J. Surg. Res. 175 (1) (2012 Jun 1) 54−55.

[13] I. Onyekwlu, R. Yakkanti, L. Protzer, et al., Surgical wound classification and surgical site infections in the orthopaedic patient, JAAOS Global Res. Rev. 1 (3) (2017) e022.

[14] R.A. Snyder, L. Johnson, J. Tice, et al., Wound classification in pediatric general surgery: significant variation exists among providers, J. Am. Coll. Surg. 217 (5) (2013) 819−826.

[15] D.M. Cardo, P.S. Falk, G. Mayhall, Validation of surgical wound classification in the operating room, Infect. Control Hosp. Epidemiol. 14 (5) (1993) 255−259.

[16] L. Devaney, K.S. Rowell, Improving surgical wound classification- why it matters, AORN J. 80 (2) (2004), 208−9, 212−223.

[17] S.M. Levy, G. Holzmann-Pazgal, K.P. Lally, et al., Quality check of a quality measure: surgical wound classification discrepancies impact risk-stratified surgical site infection rates in pediatric appendicitis, J. Am. Coll. Surg. 217 (6) (2013) 969−973.

[18] S.M. Levy, K.P. Lally, M.L. Blakely, et al., Surgical wound misclassification: a multicenter evaluation, J. Am. Coll. Surg. 220 (3) (2015) 323−329.

[19] J.L. Zinn, V. Swofford, Quality-improvement initiative: classifying and documenting surgical wounds, Wound Care Advis. 3 (1) (2014) 32−38.

[20] T.A. Oyentunji, D.O. Gonzalez, K.W. Gonzalez, et al., Wound classification in pediatric surgical procedures: measured and found wanting, J. Pediatr. Surg. 51 (6) (2016) 1014−1016.

[21] A. Wang-Chang, C. Gingert, E. Angst, F.H. Hetzer, Clinical relevance and effect of surgical wound classification in appendicitis, J. Surg. Res. 215 (2017) 132−139.

[22] J.L. Zinn, Surgical wound classification: communication is needed for accuracy, AORN J. 95 (2) (2012) 274−278.

[23] D.L. Wyrick, S.D. Smith, M.S. Dassinger, Implementation of the World Health Organization checklist and debriefing improves accuracy of surgical wound class documentation, Am. J. Surg. 210 (6) (2015) 1051−1054.

[24] S.L. Zuckerman, D.J. France, C. Green, et al., Surgical debriefing: a reliable roadmap to completing the patient safety cycle, Neurosurg. Focus 33 (5) (2012) E4.

[25] R. Cima, E. Dankbar, J. Lovely, et al., Colorectal surgery surgical site infection reduction program: a national surgical quality improvement program-driven multidisciplinary single-institution experience, J. Am. Coll. Surg. 216 (1) (2013) 23−33.

[26] T.J. Zens, D.A. Rusy, A. Gosain, Pediatric surgeon-directed wound classification improves accuracy, J. Surg. Res. 201 (2) (2016) 432−439.