Surgical Endoscopy (2020) 34:973–980
https://doi.org/10.1007/s00464-019-06884-x

2019 SAGES ORAL

# Quality of MBSAQIP data: bad luck, or lack of QA plan?

K. Noyes[1,2] · A. A. Myneni[2] · S. D. Schwaitzberg[2] · A. B. Hoffman[2]

Received: 10 April 2019 / Accepted: 31 May 2019 / Published online: 12 June 2019



## Abstract

**Background** National clinical registries are commonly used in clinical research, quality improvement, and health policy. However, little is known about methodological challenges associated with these registry analyses that could limit their impact and compromise patient safety. This study examined the quality of Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program (MSBASQIP) data to assess its usability potential and improve data collection methodologies.

**Methods** We developed a single flat file ($n = 168{,}093$) using five subsets (*Main, BMI, Readmission, Reoperation, and Intervention*) of the 2015 MBSAQIP Participant User Data File (PUF). Logic and validity tests included (1) individual profiles of patient's body mass index (BMI) changes over time, (2) individual patient care pathways, and (3) correlation analysis between variable pairs associated with the same clinical encounters.

**Results** 8888 (5.3%) patients did not have postoperative weight/BMI data; 20% of patients had different units for preoperative and postoperative weights. Postoperative weight measurements ranged between −71 and 132% of preoperative weight. There were 325 (3.7%) hospital readmissions reported on the day of or day after MBS. The self-reporting of "emergency" vs. "planned" interventions did not correlate with the type of procedure and its indication. Up to 20% of data could potentially be unused for analysis due to data quality issues.

**Conclusions** Our analysis revealed various data quality issues in the 2015 MBSAQIP PUF related to completeness, accuracy, and consistency. Since information on where the surgery was performed is lacking, it is not possible to conclude whether these issues represent data errors, patient outliers, or inappropriate care. Including automated data checks and biomedical informatics oversight, standardized coding for complications, additional de-identified facility and provider information, and training/mentorship opportunities in data informatics for all researchers who get access to the data have been shown to be effective in improving data quality and minimizing patient safety concerns.


**Keywords** MBSAQIP PUF · Data quality · Bariatric surgery · Surgical outcomes

> *"Quality is never an accident; it is always the result of intelligent effort."*
>
> *John Raskin*

The term "data-driven" has long left the academic boundaries and became a key household term used to describe decision making across a wide range of areas and industries including marketing, weather reports, financing, human resources, communication, and of course healthcare. It is widely recognized that clinicians, researchers, and policymakers depend on the use of large administrative, clinical, and registry datasets to deliver high-quality medical and surgical care, support research and innovation, and improve population health. Based on the programs that were originally created at the Department of Veterans Affairs, the American College of Surgeons (ACS) introduced and supports several highly used surgical datasets starting with the National Surgical Quality Improvement Program (NSQIP) made available in 2004 [1]. In 2017, the ACS in collaboration with the American Society for Metabolic and Bariatric Surgery (ASMBS) released the 2015 Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program (MBSAQIP) Participant Use Data File (PUF) [2].

✉ K. Noyes
enoyes@buffalo.edu

1 Division of Health Services Policy and Practice, Department of Epidemiology and Environmental Health, School of Public Health and Health Professions, University at Buffalo, Buffalo, NY, USA

2 Department of Surgery, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, NY, USA

Given the growing importance of clinical quality as a driver for innovation as well as cost containment efforts, it is anticipated that the MBSAQIP PUF will allow researchers to evaluate quality of care and outcomes among patients undergoing bariatric surgery [2, 3]. In just 2 years after its release, 28 papers have been published utilizing MBSAQIP data with a potential for improving quality of care and outcomes for patients with metabolic disorders. The methodology for using large datasets has been well established in other fields (economics, sociology, and epidemiology). However, only a few studies outlined general recommendations and checklists for analyzing large surgical datasets [4, 5], and most of the clinical studies utilizing MBSAQIP data do not report their approaches for data validation, imputing missing values and ensuring data accuracy, completeness, and consistency [6–33].

In this paper, we describe specific approaches that investigators can use before analyzing the MBSAQIP data regardless of the purpose of the analysis. This framework could be used to facilitate resident research training, strengthen institution quality improvement programs, or improve the quality of the national MBSAQIP registry.

## Methods

### MBSAQIP data

The MBSAQIP PUF is the largest bariatric-specific clinical dataset in the US and includes data on all bariatric surgery procedures performed at MBSAQIP-participating centers and was first released in 2017. It is a Health Insurance Portability and Accountability Act (HIPAA)-compliant data file containing cases submitted to the MBSAQIP Data Registry. The PUF is comprised of patient-level data and does not identify hospitals, healthcare providers or patients [34]. Since the MBSAQIP data do not contain any personal health information and are publicly available in an anonymous manner, it does not constitute Human Subjects Research as per federal regulations [35]. Therefore, the study was exempt from the University at Buffalo Institutional Review Board (IRB) review.

The 2015 PUF includes cases with operation dates between January 1, 2015 and December 31, 2015 [34]. The data collection processes and variable definitions are described elsewhere [34, 36]. Telem and Dimick summarized the contents, strengths, and limitations of the dataset and provided statistical recommendations for new investigators [2]. In short, the *Main* file is a flat data file in which each row represents a case (surgery), and contains data on 168,093 patients from 742 participating centers. There are 154 associated variables (e.g., patient demographics, preoperative patient characteristics, operative information, outcomes, and other case descriptors). The additional four files (*BMI, Reoperation, Readmission, and Intervention)* may have multiple rows per case. The *BMI* file has 13 variable and provides details on postoperative weight and BMI (recorded once or multiple times) and the days from the operation when they were recorded. The *Reoperation* file has 15 variables on patients who underwent any operations within 30 days of index MBS, including the type, whether or not it was an emergency procedure, whether or not it was unplanned (yes/no), whether or not it was related to the index MBS procedure, the reason for reoperation, and the number of days from the index surgery to the reoperation. The *Readmission* and *Intervention* data files have nine variables each and provide information on readmissions and interventions, respectively, performed within 30 days following the index MBS surgery, similar to the *Reoperation* file. Released data in the PUF are limited to 30-day outcomes and exclude the longer-term data variables collected within the MBSAQIP registry.

For the purpose of this study, we merged the five subsets of the 2015 MBSAQIP PUF (*Main, BMI, Reoperation, Readmission, and Intervention*) to develop a single flat analytic dataset (Fig. 1).

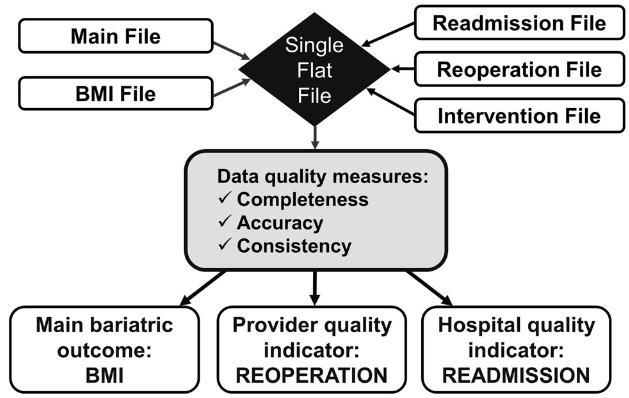


**Fig. 1** Development of a single flat file for the current study from the five component data files of the 2015 MBSAQIP PUF. Three data quality measures (completeness, accuracy, and consistency) were assessed in the analyses. Postoperative BMI is the main outcome of the bariatric surgery. Reoperations and readmissions serve as quality indicators for providers and hospitals, respectively

### Analysis

We examined data completeness, accuracy, and consistency for all registry variables in the flat analytic file. We developed several logic and validity tests including the following:

i. Individual profiles of patient BMI changes over time:
   In MBSAQIP, BMI was recorded on different dates before and after surgery. Before MBS, the high-

est BMI as well as the BMI closest to surgery were recorded. Postoperative BMI was recorded from one or more clinic visits occurring from days 0 to 150 after MBS. We examined the units of measurement [pounds (lbs.) or kilograms (kg)] used to report patient weight and calculated weight changes between the visits. If patient BMI varied more than 10 units between two subsequent appointments, we also checked the original weight and height reported at those visits and recalculated the BMI to check for computational errors. We excluded patients whose pre/postoperative weight/BMI were recorded as '0.'

ii. Individual patient care pathways (chronologic record of patient admission, discharge, and procedure history)

Using the combined flat file and the various MBSAQIP-created variables depicting the number of days from MBS to postoperative events/outcomes, we reconstructed patient chronologic care pathway starting with MBS admission, perioperative and postoperative outcomes and any (none, single or multiple) interventions, readmissions, and reoperations within 30 days of the index MBS.

iii. Clinical and face validity tests between pairs of variables describing the same clinical encounters (emergency intervention vs. procedure type, related admission with intervention vs. planned intervention)

We examined whether the entries for the admission type (urgent or elective, planned or unplanned), matched the provided categories of readmissions, interventions, and reoperations. The registry includes two follow-up questions for any reported hospital readmission. One question asks whether the admission was unplanned at the time of the principal operation. Answer "yes" implies that the readmission was unplanned and answer "no" implies that the readmission was planned. The second follow-up question requires choosing one out of 25 possible reasons for the readmission. We assessed the clinical validity of reasons for unplanned readmission using the opinion of experts in surgery and emergency medicine. Similar coding analysis was used for postoperative interventions. We excluded 11 out of 8715 readmissions and 5 out of 3855 interventions because of missing entries for reason for readmissions and interventions, respectively.

## Results

### Data completeness

The most common missing variables in the 2015 MBSAQIP PUF were BMI and ethnicity. Overall, 2442 (1.5%) patients had zero values for preoperative weight (hence, the corresponding preoperative BMI, which is calculated using patient's weight) recorded closest to surgery. Additional 8888 (5.3%) patients did not have a recorded postoperative weight (and the corresponding postoperative BMI). Hispanic ethnicity was recorded as "unknown" for 17,230 (10.3%) patients. A multivariable analysis using a standard statistical software (i.e., SAS) could potentially lose 33,868 (20%) cases resulting in an analytic dataset of 142,319 (Fig. 2).

### Data accuracy

Among 168,093 patients in the combined data file, only 125,803 (75%) patients had the same unit of measurement (lbs. or kg) for preoperative weight closest to surgery and the first recorded postoperative weight after MBS. This difference in units of measure may result in inaccurate calculations of patient BMI, which is a key indicator of MBS success. While MBS is generally indicated for patients with a BMI 35 or higher [37], 12,972 (7.7%) patients in the 2015 PUF had BMI less than 35 [BMI < 18: ($n$ = 82); 18 ≤ BMI < 25: ($n$ = 1202); 25 ≤ BMI < 30 ($n$ = 2252); 30 ≤ BMI < 35: ($n$ = 9436)]. An additional 5.5% ($n$ = 9266) of patients had a pre-surgical BMI of 60 or higher which corresponds to 350 lbs. for a 5′4″ woman or 410 lbs. for 5′9″ man [38]. The average weight loss between pre- and post-MBS was 16.1 lbs., (standard deviation = 24.2 lbs.) and ranged from −625 to 524 lbs. (−71% to 132%) of patient's preoperative weight measured closest to MBS (Fig. 3A).

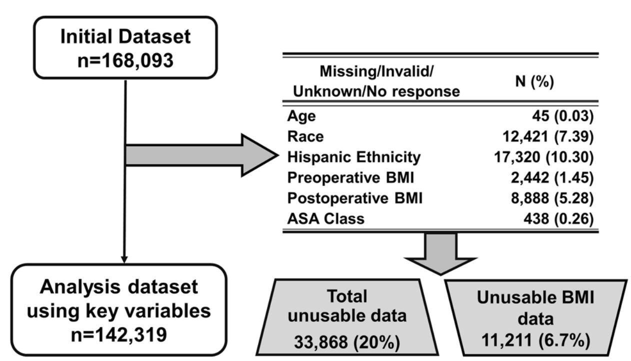


**Fig. 2** Amount of data that is potentially unusable because of missing/invalid/unknown/no response entries in the 2015 MBSAQIP PUF

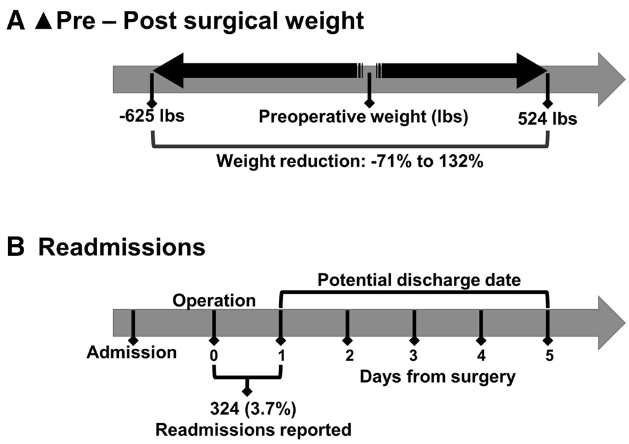


**Fig. 3** **A** Range of weight reduction comparing patients' preoperative and postoperative weights in the 2015 MBSAQIP PUF. **B** Number of readmissions reported to have occurred on the day of or the day after the index bariatric surgery procedure in the 2015 MBSAQIP PUF

Out of 8715 hospital readmissions within 30 days of MBS, 325 (3.7%) were reported on day 0 or 1 postoperatively where day 0 is the day of operation (Fig. 3B). Given that the dataset does not contain a variable describing whether the surgery was performed in an inpatient or ambulatory setting, it is impossible to conclude if this represents a data error, admission to an ICU, transfer to a higher level hospital or inpatient admission after ambulatory surgery.

### Data consistency

Among 25 provided reasons for readmission, all except for one (internal hernia, $n=49$) category represented emergency conditions according to the current bariatric and acute surgery standards (Table 1). Among the 8715 readmissions, 611 (7%) of them were coded as planned ('no' option for the 'Unplanned Readmission' variable) including 166 cases of nausea and vomiting, 67 cases of unspecified abdominal pain, and 21 cases of anastomotic leak.

Similarly, among 24 possible reasons for postoperative interventions, all except for one (internal hernia, $n=10$) category represented emergency procedures (Table 2). Out of 3850 postoperative interventions performed the 30 days following MBS, 480 (12%) were coded as planned at the time of the index MBS, including interventions for respiratory failure ($n=6$), gallstone disease ($n=4$), and pulmonary embolism ($n=4$).

## Discussion

Data cleaning and imputation of missing values are the key initial steps when working with preexisting (administrative, clinical, registry, patient-reported) datasets collected for non-research purposes. While preparing the MBSAQIP PUF for the analysis, we uncovered several problems with data completeness, accuracy, and consistency. We identified missing and out of range values for the key parameters (such as weight, BMI, date of readmission). We also encountered several variables with implausible values (600 lbs weight loss from the day of surgery to post-surgical follow-up appointment or patients with BMI as low as 15) and erroneous categories (e.g., planned readmissions for fever, nausea and respiratory failure, zero value weight/BMI). Preoperative BMI was less than 35 for 7.7% of patients. Recently, ASMBS updated its recommendation that MBS may be offered as an option for individuals with Class I obesity ($30 \leq \text{BMI} < 35$) and obesity-related comorbidities, who do not achieve substantial, durable weight loss and improvement in their comorbidities with reasonable non-surgical methods [39]. However, without additional information, it is difficult to determine whether some weight and BMI recordings of patients were real, due to errors in data entry, or whether there were unusual circumstances where these patients had to undergo MBS. Overall, data quality issues resulted in the loss of 20% of observations from the analytic dataset.

Despite these significant data quality issues, to date, none of the manuscripts using the MBSAQIP PUF reported a comprehensive plan for data cleaning, imputation of missing values, and data quality assurance. While the current MBSAQIP Program Standards Manual (version 2.0) [36] emphasizes the importance of data quality for assessing quality of surgical care, no specific steps and procedures for data quality assurance are currently required. Such lack of transparency and awareness on data quality is, at best, impractical and potentially dangerous since the data are being widely used for QI/QA, provider comparison, and payment.

The issue of data quality is not new and not unique to surgical datasets or the US. Several comprehensive reviews on methodology for validation of data quality, data standardization to facilitate data sharing, and harmonization of data collection practices and systems have been published during the last decade by leading national and European healthcare, public health, health informatics, health insurance, and clinical quality organizations [40–49]. The approaches and procedures that could reliably improve data quality include the following:

1. Automated data checks for completeness (no-skip pattern) and accuracy (e.g., flagging values outside of predefined acceptable range, BMI = 140) [43, 45, 46]
2. Data audits for consistency (e.g., use multiple variables to conduct logic checks: readmission before index admission discharge is not valid) [45]
3. In addition to providing definitions of all registry variables, the data publisher should also include standard-

**Table 1** Unplanned and planned readmissions and their reasons in the 2015 MBSAQIP PUF

| Reasons for readmission[a,b] | N (%) | | |
|---|---|---|---|
| | Unplanned[c] | | Total |
| | Yes | No | |
| Abdominal pain, not otherwise specified | 1110 (13.7) | 67 (11.0) | 1177 (13.5) |
| Anastomotic ulcer | 93 (1.1) | 5 (0.8) | 98 (1.1) |
| Anastomotic/staple line leak | 398 (4.9) | 21 (3.4) | 419 (4.8) |
| Band erosion | 4 (0.0) | 1 (0.2) | 5 (0.1) |
| Band slippage/prolapse | 4 (0.0) | 1 (0.2) | 5 (0.1) |
| Bile reflux gastritis | 8 (0.1) | 0 (0.0) | 8 (0.1) |
| Bleeding | 313 (3.9) | 20 (3.3) | 333 (3.8) |
| Gallstone disease | 117 (1.4) | 12 (2.0) | 129 (1.5) |
| Gastric distention | 10 (0.2) | 1 (0.2) | 11 (0.1) |
| Gastro-gastric fistula | 13 (0.2) | 3 (0.5) | 16 (0.2) |
| Gastrointestinal perforation | 52 (0.6) | 6 (1.0) | 58 (0.7) |
| Incisional hernia | 55 (0.7) | 6 (1.0) | 61 (0.7) |
| Infection/fever | 296 (3.6) | 22 (3.6) | 318 (3.6) |
| Internal hernia | 46 (0.6) | 3 (0.5) | 49 (0.6) |
| Intestinal obstruction | 309 (3.8) | 22 (3.6) | 331 (3.8) |
| LAGB—port, tubing, or band problem | 10 (0.1) | 1 (0.2) | 11 (0.1) |
| Nausea and vomiting, fluid, electrolyte, or nutritional depletion | 2318 (28.6) | 166 (27.2) | 2484 (28.5) |
| Other abdominal sepsis | 100 (1.2) | 2 (0.3) | 102 (1.2) |
| Other respiratory failure | 43 (0.5) | 5 (0.8) | 48 (0.6) |
| Pneumonia | 209 (2.6) | 9 (1.5) | 218 (2.5) |
| Pulmonary embolism | 191 (2.4) | 13 (2.1) | 204 (2.3) |
| Strictures/stomal obstruction | 201 (2.5) | 8 (1.3) | 209 (2.4) |
| Vein thrombosis requiring therapy | 250 (3.1) | 17 (2.8) | 267 (3.1) |
| Wound infection/evisceration | 264 (3.3) | 16 (2.6) | 280 (3.2) |
| Other[d] | 1690 (20.8) | 184 (30.1) | 1874 (21.5) |
| Total | 8104 (93.0) | 611 (7.0) | 8715 |

*LAGB* laparoscopic adjustable gastric banding

[a]Defined as a hospital readmission within 30 days of the principal MBS procedure

[b]Presents most likely reason for readmission based on medical records review

[c]Defined as whether or not the readmission was unplanned at the time of the principal operative procedure

[d]Readmission was due to a reason other than those specifically listed

ized algorithms for abstracting these values based on commonly used clinical data systems (e.g., which values to use if multiple inconsistent values for patient weight are reported in the EMR, whether to assign reasons for readmission based on the DRG, CPT, or clinical notes) [40, 41, 43].

4. Annual conferences, continuous education courses, and workshops for clinical data reviewers and abstractors to share successful practices and enhance homogeneity of national registries and databases [42, 44, 48].
5. Reporting commonly used diagnostic codes, procedures, and interventions to facilitate linkage with insurance reimbursement schedules [47].
6. For any new or recoded variables created by the data registry and released with PUF, a detailed methodology should be included to help researchers understand the definition and limitations of the new parameter (was the readmission unplanned, yes/no) [50].
7. Regular workshops to learn how to use MBSAQIP appropriately and avoid common data and analytic pitfalls, similar to the approach used by the CMS/ResDAC [51]

Depending on the type of data standards and the aspects of data quality considered, data standards may or may not impact quality of data analysis. Because of the substantial differences in normative language and jargon used by different professionals, significant expectation differences between suppliers (data reviewers and abstractors) and customers (researchers and data analysts) can result. For

**Table 2** Unplanned and planned interventions and their reasons in the 2015 MBSAQIP PUF

| Reason for intervention[b] | N (%) | | |
|---|---|---|---|
| | Unplanned[c] | | Total |
| | Yes | No | |
| Abdominal pain, not otherwise specified | 207 (6.1) | 15 (3.1) | 222 (5.8) |
| Anastomotic ulcer | 161 (4.8) | 5 (1.0) | 166 (4.3) |
| Anastomotic/staple line leak | 435 (12.9) | 40 (8.4) | 475 (12.3) |
| Band erosion | 5 (0.1) | 3 (0.6) | 8 (0.2) |
| Bile reflux gastritis | 17 (0.5) | 1 (0.2) | 18 (0.5) |
| Bleeding | 218 (6.5) | 15 (3.1) | 233 (6.1) |
| Gallstone disease | 20 (0.6) | 4 (0.8) | 24 (0.6) |
| Gastric distention | 14 (0.4) | 1 (0.2) | 15 (0.4) |
| Gastro-gastric fistula | 29 (0.9) | 1 (0.2) | 30 (0.8) |
| Gastrointestinal perforation | 39 (1.2) | 6 (1.3) | 45 (1.2) |
| Incisional hernia | 4 (0.1) | 0 (0.0) | 4 (0.1) |
| Infection/fever | 80 (2.4) | 7 (1.5) | 87 (2.3) |
| Internal hernia | 9 (0.3) | 1 (0.2) | 10 (0.3) |
| Intestinal obstruction | 46 (1.3) | 9 (1.9) | 55 (1.4) |
| LAGB—port, tubing, or band problem | 4 (0.1) | 0 (0.0) | 4 (0.1) |
| Nausea and vomiting, fluid, electrolyte, or nutritional depletion | 582 (17.3) | 69 (14.3) | 651 (16.9) |
| Other abdominal sepsis | 83 (2.5) | 9 (1.9) | 92 (2.4) |
| Other respiratory failure | 51 (1.5) | 6 (1.2) | 57 (1.5) |
| Pneumonia | 17 (0.5) | 4 (0.8) | 21 (0.5) |
| Pulmonary embolism | 30 (0.9) | 4 (0.8) | 34 (0.9) |
| Strictures/stomal obstruction | 530 (15.7) | 70 (14.5) | 600 (15.6) |
| Vein thrombosis requiring therapy | 49 (1.5) | 10 (2.1) | 59 (1.5) |
| Wound infection/evisceration | 88 (2.6) | 12 (2.5) | 100 (2.6) |
| Other[d] | 649 (19.3) | 191 (39.5) | 840 (21.8) |
| Total | 3367 (87.5) | 483 (12.5) | 3850 |

*LAGB* laparoscopic adjustable gastric banding

[a]Defined as an intervention within 30 days of the principal MBS procedure

[b]Presents most likely reason for intervention after MBS procedure

[c]Defined as whether or not the intervention was unplanned at the time of the principal operative procedure

[d]Intervention was due to a reason other than those specifically listed

example, some view data quality as those things pertaining to the data values themselves (e.g., when was the patient discharged) while others also include contextual features such as discharge disposition, discharge plan, who performed the discharge education, did patient feel ready to be discharged, etc. Lack of specificity about data standards and data quality enables diverging expectations of the value of data standards and impedes their adoption and progress. We hope instead that clarity in language and intent will hasten the promise of standardization that has benefitted so many other industries.

While not immune from its own flaws, one of the major advantages of the ACS NSQIP dataset is its dependence on inter-rater reliability between the data collector and audits by participating institutions. In this system, variables with greater than 5% disagreement are flagged for future data collection education programs [1]. NSQIP has seen tremendous benefits with this system. A similar and successful system is demonstrated by The Society for Thoracic Surgeons (STS). However, their unique distinction is in categorizing three levels of quality assurances—at the internal, regional, and national levels, which has led to the STS having correctly verified data with an average of 96.4% [52, 53].

## Conclusion

The importance and value of large datasets in this new era of research cannot be overstated. Moreover, as we continue to focus on value-based, patient-centered care, quality assessments and its subsequent consequences have implications over many areas within the healthcare system. While always keeping patient safety and evidence-based medicine at the forefront, it is imperative that we work together within the medical community. There is need for collaboration with other allied healthcare personnel, as well as partnering with other experts in the public health sector, including epidemiologists, statisticians, and policy makers, in order to properly utilize and interpret these complex datasets. Finally, we must learn from other successful databases, and utilize some of their techniques in our pursuit for excellence in data quality and patient care.

**Acknowledgements** The authors would like to thank Monami Majumder and Evan Kessler for their assistance with the preparation of this manuscript.

**Funding** This manuscript was in part funded by Surgical Outcomes and Research (UB SOAR) program at the Department of Surgery, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, New York.

## Compliance with ethical standards

**Disclosures** The earlier version of the study was presented at the 2019 SAGES Annual Meeting. Dr. Hoffman is a paid consultant for Ethicon US, LLC (not related to the study). Dr. Steven Schwaitzberg is a paid consultant for Nu View Surgical, Acuity Bio, Activ Surgical, Human Extensions, Levitra Magnetics and Arch Therapeutics (not related to the study). K. Noyes and A. A. Myneni declare that they have no conflict of interest or financial ties to disclose. The American College of Surgeons Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program and the centers participating in the ACS MBSAQIP are the source of the data used herein; they have not verified and are not responsible for the statistical validity of the data analysis or the conclusions derived by the authors.

## References

1. Shiloach M, Frencher SK Jr, Steeger JE, Rowell KS, Bartzokis K, Tomeh MG, Richards KE, Ko CY, Hall BL (2010) Toward robust information: data quality and inter-rater reliability in the American College of Surgeons National Surgical Quality Improvement Program. J Am Coll Surg 210:6–16
2. Telem DA, Dimick JB (2018) Practical guide to surgical data sets: Metabolic and Bariatric Surgery Accreditation and Quality Program (MBSAQIP). JAMA Surg 153:766–767
3. Brethauer SA, Kim J, el Chaar M, Papasavas P, Eisenberg D, Rogers A, Ballem N, Kligman M, Kothari S, Committee ACI (2015) Standardized outcomes reporting in metabolic and bariatric surgery. Surg Obes Relat Dis 11:489–506
4. Ehrenstein V, Nielsen H, Pedersen AB, Johnsen SP, Pedersen L (2017) Clinical epidemiology in the era of big data: new opportunities, familiar challenges. Clin Epidemiol 9:245–250
5. Kruse CS, Goswamy R, Raval Y, Marawi S (2016) Challenges and opportunities of big data in health care: a systematic review. JMIR Med Inform 4:e38
6. Ardila-Gatas J, Sharma G, Lloyd SJ, Khorgami Z, Tu C, Schauer PR, Brethauer SA, Aminian A (2019) A nationwide safety analysis of discharge on the first postoperative day after bariatric surgery in selected patients. Obes Surg 29:15–22
7. Berger ER, Clements RH, Morton JM, Huffman KM, Wolfe BM, Nguyen NT, Ko CY, Hutter MM (2016) The impact of different surgical techniques on outcomes in laparoscopic sleeve gastrectomies: the first report from the metabolic and bariatric surgery accreditation and quality improvement program (MBSAQIP). Ann Surg 264:464–473
8. Chaar ME, Lundberg P, Stoltzfus J (2018) Thirty-day outcomes of sleeve gastrectomy versus Roux-en-Y gastric bypass: first report based on Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program database. Surg Obes Relat Dis 14:545–551
9. Clapp B, Hahn J, Dodoo C, Guerra A, de la Rosa E, Tyroch A (2018) Evaluation of the rate of marginal ulcer formation after bariatric surgery using the MBSAQIP database. Surg Endosc. https://doi.org/10.1007/s00464-018-6468-6
10. Daigle CR, Brethauer SA, Tu C, Petrick AT, Morton JM, Schauer PR, Aminian A (2018) Which postoperative complications matter most after bariatric surgery? Prioritizing quality improvement efforts to improve national outcomes. Surg Obes Relat Dis 14:652–657
11. Dang JT, Switzer N, Delisle M, Laffin M, Gill R, Birch DW, Karmali S (2019) Predicting venous thromboembolism following laparoscopic bariatric surgery: development of the BariClot tool using the MBSAQIP database. Surg Endosc 33:821–831
12. Dang JT, Switzer NJ, Sun WYL, Raghavji F, Birch DW, Karmali S (2018) Evaluating the safety of intragastric balloon: an analysis of the Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program. Surg Obes Relat Dis 14:1340–1347
13. Demeusy A, Sill A, Averbach A (2018) Current role of staple line reinforcement in 30-day outcomes of primary laparoscopic sleeve gastrectomy: an analysis of MBSAQIP data, 2015-2016 PUF. Surg Obes Relat Dis 14:1454–1461
14. El Chaar M, Stoltzfus J (2018) Assessment of sleeve gastrectomy surgical technique: first look at 30-day outcomes based on the MBSAQIP database. J Am Coll Surg 227:564–572
15. El Chaar M, Stoltzfus J, Melitics M, Claros L, Zeido A (2018) 30-Day outcomes of revisional bariatric stapling procedures: first report based on MBSAQIP data registry. Obes Surg 28:2233–2240
16. Fazl Alizadeh R, Li S, Inaba CS, Dinicu AI, Hinojosa MW, Smith BR, Stamos MJ, Nguyen NT (2019) Robotic versus laparoscopic sleeve gastrectomy: a MBSAQIP analysis. Surg Endosc 33:917–922
17. Feng X, Andalib A, Brethauer SA, Schauer PR, Aminian A (2018) How safe is bariatric surgery in patients with class I obesity (body mass index 30–35 kg/m(2))? Surg Obes Relat Dis 15:253–260
18. Gondal AB, Hsu CH, Khoubyari R, Ghaderi I (2019) Development of a bariatric surgery specific risk assessment tool for perioperative myocardial infarction. Surg Obes Relat Dis 15:462–468
19. Haskins IN, Ju T, Whitlock AE, Rivas L, Amdur RL, Lin PP, Vaziri K (2018) Older age confers a higher risk of 30-day morbidity and mortality following laparoscopic bariatric surgery: an analysis of the metabolic and bariatric surgery quality improvement program. Obes Surg 28:2745–2752
20. Inaba CS, Koh CY, Sujatha-Bhaskar S, Zhang L, Nguyen NT (2018) Same-day discharge after laparoscopic Roux-en-Y Gastric Bypass: an analysis of the metabolic and bariatric surgery accreditation and quality improvement program database. J Am Coll Surg 226:868–873
21. Janik MR, Mustafa RR, Rogula TG, Alhaj Saleh A, Abbas M, Khaitan L (2018) Safety of laparoscopic sleeve gastrectomy and Roux-en-Y gastric bypass in elderly patients—analysis of the MBSAQIP. Surg Obes Relat Dis 14:1276–1282
22. Janik MR, Mustafa RR, Rogula TG, Alhaj Saleh A, Abbas M, Khaitan L (2018) Application of HARM score to measure

surgical quality and outcomes in bariatric patients. Obes Surg 28:2815–2819
23. Janik MR, Rogula TG, Mustafa RR, Alhaj Saleh A, Khaitan L (2019) Safety of revision sleeve gastrectomy compared to Roux-Y Gastric bypass after failed gastric banding: analysis of the MBSAQIP. Ann Surg 269:299–303
24. Jurgensen JA, Reidt W, Kellogg T, Mundi M, Shah M, Collazo Clavell ML (2019) Impact of patient attrition from bariatric surgery practice on clinical outcomes. Obes Surg 29:579–584
25. Kanters AE, Shubeck SP, Varban OA, Dimick JB, Telem DA (2018) Incidence and efficacy of stent placement in leak management after bariatric surgery: an MBSAQIP analysis. Ann Surg. https://doi.org/10.1097/SLA.0000000000003023
26. Kumar SB, Hamilton BC, Wood SG, Rogers SJ, Carter JT, Lin MY (2018) Is laparoscopic sleeve gastrectomy safer than laparoscopic gastric bypass? a comparison of 30-day complications using the MBSAQIP data registry. Surg Obes Relat Dis 14:264–269
27. Lundberg PW, Stoltzfus J, El Chaar M (2019) 30-day outcomes of robot-assisted versus conventional laparoscopic sleeve gastrectomy: first analysis based on MBSAQIP. Surg Obes Relat Dis 15:1–7
28. Lundberg PW, Wolfe S, Seaone J, Stoltzfus J, Claros L, El Chaar M (2018) Robotic gastric bypass is getting better: first results from the Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program. Surg Obes Relat Dis 14:1240–1245
29. Rahman U, Docimo S, Pryor AD, Bates A, Obeid NR, Spaniolas K (2018) Routine contrast imaging after bariatric surgery and the effect on hospital length of stay. Surg Obes Relat Dis 14:517–520
30. Sebastian R, Howell MH, Chang KH, Adrales G, Magnuson T, Schweitzer M, Nguyen H (2018) Robot-assisted versus laparoscopic Roux-en-Y gastric bypass and sleeve gastrectomy: a propensity score-matched comparative analysis using the 2015–2016 MBSAQIP database. Surg Endosc 33:1600–1612
31. Sharma G, Strong AT, Tu C, Brethauer SA, Schauer PR, Aminian A (2018) Robotic platform for gastric bypass is associated with more resource utilization: an analysis of MBSAQIP dataset. Surg Obes Relat Dis 14:304–310
32. Zafar SN, Miller K, Felton J, Wise ES, Kligman M (2019) Postoperative bleeding after laparoscopic Roux en Y gastric bypass: predictors and consequences. Surg Endosc 33:272–280
33. Zafar SN, Shockcor N, Prasad N, Wise ES, Jackson HT, Kligman M (2019) Laparoscopic sleeve gastrectomy has a lower risk of postoperative bleeding than laparoscopic Roux-en-Y gastric bypass. Surg Laparosc Endosc Percutan Tech 29:53–57
34. Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program (Published January 2017) User Guide for the 2015 Participant Use Data File (PUF)
35. Department of Health and Human Services Office for Human Research Protections (2018) 45 CFR 46-Protection of Human Subjects, 2018 Requirements: 104 Exempt research. Department of Health and Human Services Office for Human Research Protections, Rockville
36. Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program (October 2016) Standards Manual version 2.0. Resources for Optimal Care of the Metabolic and Bariatric Surgery Patient. 2016
37. American Society for Metabolic and Bariatric Surgery (2018) Who is a Candidate for Bariatric Surgery?. American Society for Metabolic and Bariatric Surgery, Gainsville
38. Fryar CD, Kruszon-Moran D, Gu Q, Ogden CL (2018) Mean body weight, height, waist circumference, and body mass index among adults: United States, 1999–2000 through 2015–2016. Natl Health Stat Report 1:16
39. Aminian A, Chang J, Brethauer SA, Kim JJ, American Society for M, Bariatric Surgery Clinical Issues C (2018) ASMBS updated position statement on bariatric surgery in class I obesity (BMI 30-35 kg/m(2)). Surg Obes Relat Dis 14:1071–1087
40. AHIMA e-HIM Work Group on EHR Data Content (2006) Data standard time: data content standardization and the HIM role. J AHIMA 77:26–32
41. Amster A, Jentzsch J, Pasupuleti H, Subramanian KG (2015) Completeness, accuracy, and computability of National Quality Forum-specified eMeasures. J Am Med Inform Assoc 22:409–416
42. Hall BL, Hamilton BH, Richards K, Bilimoria KY, Cohen ME, Ko CY (2009) Does surgical quality improve in the American College of Surgeons National Surgical Quality Improvement Program: an evaluation of all participating hospitals. Ann Surg 250:363–376
43. Huser V, Kahn MG, Brown JS, Gouripeddi R (2018) Methods for examining data quality in healthcare integrated data repositories. Pac Symp Biocomput 23:628–633
44. Ko C (2012) The critical importance of good data to improving quality. Patient Safety & Quality Healthcare
45. Kupzyk KA, Cohen MZ (2015) Data validation and other strategies for data entry. West J Nurs Res 37:546–556
46. Marc DT, Beattie J, Herasevich V, Gatewood L, Zhang R (2016) Assessing metadata quality of a federally sponsored health data repository. AMIA Annu Symp Proc 2016:864–873
47. Nouraei SA, Hudovsky A, Virk JS, Saleh HA (2017) Impact of developing a multidisciplinary coded dataset standard on administrative data accuracy for septoplasty, septorhinoplasty and nasal trauma surgery. J Laryngol Otol 131:341–346
48. Raval MV, Pawlik TM (2018) Practical guide to surgical data sets: national surgical quality improvement program (NSQIP) and Pediatric NSQIP. JAMA Surg 153:764–765
49. Royal College of Surgeons of England (2018) Improving surgical data. Royal College of Surgeons of England, London
50. American College of Surgeons National Surgical Quality Improvement Program (2017) User guide for the 2016 ACS NSQIP procedure targeted participant use data file (PUF). Chicago, IL
51. Research Data Assistance Center (ResDAC) (2019) Learn about CMS data. University of Minnesota School of Public Health, Minneapolis
52. Ferguson TB Jr, Hammill BG, Peterson ED, DeLong ER, Grover FL, Committee STSND (2002) A decade of change-risk profiles and outcomes for isolated coronary artery bypass grafting procedures, 1990–1999: a report from the STS National Database Committee and the Duke Clinical Research Institute. Society of Thoracic Surgeons. Ann Thorac Surg 73:480–489 **(discussion 489–490)**
53. Welke KF, Ferguson TB Jr, Coombs LP, Dokholyan RS, Murray CJ, Schrader MA, Peterson ED (2004) Validity of the society of thoracic surgeons national adult cardiac surgery database. Ann Thorac Surg 77:1137–1139

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.