Published in final edited form as:

J Am Coll Surg. 2019 March ; 228(3): 255–264. doi:10.1016/j.jamcollsurg.2018.12.025.

Using Machine Learning to Identify Change in Surgical Decision Making in Current Use of Damage Control Laparotomy

John A Harvin, MD MS FACS^{1,2}, Charles Green, PhD^{3,4}, Claudia Pedroza, PhD^{3,4}, Jon E Tyson, MD MPH^{3,4}, Laura J Moore, MD FACS^{1,2}, Charles E Wade, PhD^{1,2}, John B Holcomb, MD FACS^{1,2}, and Lillian S Kao, MD MS FACS^{1,4}

¹Department of Surgery, University of Texas McGovern Medical School, Houston, TX

²Center for Translational Injury Research, University of Texas McGovern Medical School, Houston, TX

³Department of Pediatrics, University of Texas McGovern Medical School, Houston, TX

⁴Center for Clinical Research and Evidence Based Medicine, University of Texas McGovern Medical School, Houston, TX

Abstract

Background—In a prior study, we reported the successful reduction in the use of damage control laparotomy (DCL), however, no change in the relative frequencies of specific indications was observed. In this study, we aimed to use machine learning to help identify the changes in surgical decision making that occurred.

Study Design—Adult patients undergoing emergent trauma laparotomy were included: 1) PreQI: 1/1/2011–10/31/2013 and 2) PostQI: 11/1/2013–06/30/2016. Using 72 variables before or during emergent laparotomy, random forest algorithms predicting DCL before and after a QI intervention were created. The main outcome of the algorithms was the strength of individual factor significance in predicting the use of DCL, calculated by determining the mean decrease in accuracy (MDA) in the model if that variable was removed

Results—In the PreQI group, 24 of 72 factors significantly predicted DCL, the strongest being bowel resection (mean MDA 16) and operating room red blood cell transfusions (mean MDA 15). The remaining variables were spread along the continuum of care from injury to emergent laparotomy end. In the PostQI group, 12 of 72 factors significantly predicted DCL, the strongest being last operating room lactate (mean MDA 12) and operating room red blood cell transfusions

Corresponding Author and Requests for Reprints: John A Harvin, MD, 6431 Fannin Street, MSB 4.264, Houston, Texas 77030, Telephone – 713 500 6362, Fax – 713 500 0651, john.harvin@uth.tmc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures Information: Nothing to disclose. Disclosures outside the scope of this work: Dr Wade is a paid consultant to Haemonetics, receives grants money from Masimo Grifols, and receives royalties and stock options from Decisio, LLC. Dr Holcomb is Chief Medical Officer for Paytime Medical.

Publisher's Disclaimer: Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Advancing Translational Sciences or the National Institutes of Health.

(mean MDA 14). In addition to having 12 fewer significant factors predictive of DCL, the predictive factors in the PostQI group were mainly intraoperative factors.

Conclusion—A machine learning analysis provided novel insights into the changes in decisionmaking achieved by a successful QI intervention and should be considered an adjunct to understanding successful pre- and post-intervention QI studies. The analysis suggested a shift towards using mostly intraoperative factors to determine the use of DCL.

Precis

A successful quality intervention to de-implement damage control laparotomy appeared to change surgical decision making by delaying timing of decision until near the end of the index operation.

Introduction

In a learning trauma system, the use of an innovation is continuously and iteratively refined based on evolving care needs and circumstances.¹ Rapid learning occurs with real-time access to data, efficient data analytics, and a culture of learning. Machine learning, which is increasingly being utilized in healthcare, is a data analytic strategy that is ideally suited to a learning trauma system. Machine learning is a software-based strategy by which factors involved in surgical decision making can be quantified and future decisions can be predicted.² Supervised machine learning uses historical data to train models to predict outcomes in new cases.

Damage control laparotomy (DCL) is an innovation from the 1990s in which an abbreviated laparotomy is performed to achieve source control of hemorrhage and contamination after trauma. The patient is left with an open abdomen, resuscitated to normal physiology in the intensive care unit (ICU), and returned to the operating room (OR) for definitive reconstruction once clinically stable. Initially described in a small subgroup of coagulopathic patients requiring packing and massively transfused patients with a major abdominal vascular injury, DCL significantly improved survival.^{3,4} However, DCL comes at the price of significant morbidity such as enterocutaneous fistulae, complex ventral hernias, and infectious complications.^{5,6} DCL also commits patients to at least one re-operation and is associated with increased resource utilization and costs.⁷ Despite these downsides, the indications for DCL have increased exponentially over time without high quality evidence demonstrating benefit.^{8,9}

In a prior study, we reported the success of a single-center, multi-faceted quality improvement (QI) intervention in reducing the rate of DCL without increasing morbidity or mortality.¹⁰ The intervention included consensus building regarding indications for definitive laparotomy and DCL as well as audit and feedback.¹¹ The QI intervention resulted in an immediate and sustained reduction in DCL from 38% to 23%. Although the number of agreed upon indications for DCL decreased, there was no change in the relative frequencies of indications for DCLs performed before and after the QI intervention. Thus, explicit knowledge of the critical factors and conditions that resulted in the decreased use of DCL was lacking; this is a common shortcoming of QI research.¹²

The purpose of this study was to use supervised machine learning to provide insight regarding how a QI intervention changed surgical decision making. We hypothesized that machine learning could be used to identify a change in the factors predicting DCL as a result of the QI intervention. An explicit understanding of how surgical decision making changed is necessary to inform future evidence-based care.

Methods

The study was an exploratory analysis of a pre/post intervention QI project performed at the Red Duke Trauma Institute at Memorial Hermann Hospital-Texas Medical Center, an American College of Surgeons verified Level 1 trauma center that is the primary teaching hospital for the McGovern Medical School at UTHealth. The Red Duke Trauma Institute is one of the only two adult Level 1 trauma centers in Houston, Texas. The Institutional Review Board deemed the project to be QI and exempt from review.

Participants of this study were surgeons performing emergent laparotomies on adult trauma patients >15 years of age. Emergent laparotomy was defined as direct admission to the OR from the emergency department (ED). Patients who died intraoperatively were excluded.

The cohort was divided into patients undergoing emergent trauma laparotomy during two time periods: 1) PreQI – January 1, 2011 and October 31, 2013 and 2) PostQI – November 1, 2013 and June 30, 2016. Patient data for the PreQI group was retrospectively collected for two peer-reviewed studies and was same cohort as the Control group of the previously published QI study;^{5,10} the data for the PostQI group was collected prospectively using the same definitions and methods and included the QI cohort of the previously published QI study plus an additional 8 months of accrued patients.

A total of 72 variables were included in the analysis and were obtained throughout the continuum of care -- from admission (patient demographics and injury information) through the ED (vital signs, labs, and resuscitation volumes) and through the OR (first and last vital signs and labs, OR resuscitation, estimated blood loss, procedures performed, and surgeon). Surgeons with <10 laparotomies in the individual time periods were grouped together for the analysis. The outcome predicted was performance of a DCL.

Random Forest Algorithm: a Machine Learning Tool

Random forest is a computer learning algorithm based upon de-correlated decision trees.¹³ This methodology was chosen over other quantitative methods because it is robust to correlated predictors and potentially non-linear relationships between the predictors and the criterion variables. First, a portion of the data is randomly selected to be used as a training set. The Random Forest algorithm is developed on the training set and then applied to the remainder of the data, which is referred to as the test set.

To develop the forest, the individual trees must be grown. To grow the tree, each variable at each decision node along the decision tree is selected randomly from a pre-specified number of candidate variables. From these candidate variables, the algorithm randomly identifies the best variable and cut-point on that variable. This continues iteratively over and over using

the training set to minimize misclassification in the resulting branches until the tree is fully developed. The entire process is then repeated for n number of trees specified in the forest.

The most important parameter governing tree performance is the number of candidate variables selected at decision node.¹⁴ This number is empirically determined using 5 fold cross validation in the training set.¹⁵ Based upon the misclassification performance of multiple random forests on the training set, the number of candidate variables available at each decision node is selected to use in the random forest for the test set.

The optimized random forest of n trees is then applied to the test set. Individuals in the test set are given predicted categories based upon the results of the n trees. The distribution of predicted outcome values is then compared to known outcome values and the model accuracy and sensitivity can be calculated.

Variable significance is calculated by determining the mean decrease in accuracy (MDA) in the model if that variable was removed. As the MDA of a variable increases, the relative importance of that variable to the model increases.

Model Creation

For both the PreQI and PostQI groups, two models predicting DCL were created using random forest methodology (R Studio Version 1.0.143) for a total of four models.¹⁶ The two models in each time period were: 1) a model using only complete data (patients with missing data excluded) and 2) a model using imputation to account for missing data. The two models acted as a sensitivity analysis to address variable missingness.¹⁷

Imputation was performed using random forest methodology in R. First, all missing continuous variables were replaced with the variable median value and all factor variables were replaced with the most variable frequent value. Next, a random forest was run on the completed data set and a proximity matrix created and used to update the imputed values. For continuous variables, the imputed value was then updated using the scaled proximity score of the nonmissing observations as a weighing factor. For factor variables, the imputed value was the category with the largest average proximity. This process was iterated 5 times.¹⁶

The four models were created using a similar methodology. First, a training set was created using a random sample of two-thirds of the data. The most important parameter to tune in the random forest algorithm is the number of randomly selected variables evaluated for classification at each branch in the tree. Tuning this parameter used $k = 5$ fold cross-validation. The model that optimized the misclassification rate using k -fold cross-validation on the training set was then evaluated on a test set (random one-third hold-out sample).

A confusion matrix and receiver operator curve (ROC) were then created to describe the accuracy of the model performance on the test data. Finally, the MDA, which is a measure of how many standard deviations that a raw score is above or below the population mean, for each variable was calculated. As the MDA of a variable increases, the relative importance of that variable to the model increases. In this analysis, the MDA was scaled to create a z-score to allow for comparison across variables.

Determination of Significant Variables

To fulfill the aim of the project – to describe how surgeon judgement changed after institution of the successful QI project – two separate analyses were performed.

Variable Significance in Both Periods.—First, the MDA of variables in both models in both time periods were analyzed. Variables were considered to be significant if both models concurred that the variable MDA was either ≥ 2 or ≤ -2 (in the same direction). A cut off of 2 standard deviations from the mean without correction for multiplicity was chosen as we did not want to prematurely exclude any potentially modifiable factors in surgical decision making that could be used to further improve and generalize this QI effort.

Variables with a MDA meeting the threshold of 2 or ≤ -2 were then listed in a side-by-side table comparing the pre-QI and post-QI periods. This method of analysis allowed the comparison of significant variables in both time periods as a way to understand what variables changed in the determination to perform DCL.

Net Mean Decrease in Accuracy.—Second, the difference in the MDA in complete cases in the pre QI and post QI groups was calculated for each variable. This method of analysis was performed to attempt to express both the vector (increase or decrease in importance) and the magnitude of the change in variable MDA over the two time periods. A net difference in MDA was similarly derived from the two imputed models for all variables regardless of individual MDA. If the net MDA in both models was ≥ 2 or ≤ -2 , the variable was considered to have had a significant change between the preQI and postQI periods.

Results

In the PreQI Group, 12,922 patients were admitted to the trauma center; 582 (5%) patients underwent emergent laparotomy with a DCL rate of 38% (Table 1). In the PostQI Group, 13,061 patients were admitted to the trauma center; 611 (5%) patients underwent emergent laparotomy with a DCL rate of 21%. The groups had similar baseline demographics and injury severity with the exception of a higher percentage of patients suffering penetrating trauma in the PostQI group (Table 1). The relative frequencies of the stated indications for DCL were not statistically significantly different between the two time periods ($p=0.28$). Intraoperative deaths occurred in and were excluded from the analysis in 23 (4%) and 29 (5%) of the PreQI and PostQI groups, respectively.

Model Accuracy

For both the PreQI and PostQI groups, the accuracy of all models was good (range 0.772 – 0.857, Table 2). The complete case models had less than 50% of the entire cohort of each time period reflecting significant missingness of clinical variables (PreQI complete case model contained 236 of 559 patients; PostQI complete case model contained 209 of 582 patients). The variables with the most missingness in the PreQI Group were ED base excess (17%), last OR pH (21%), last OR base excess (21%), and last OR lactic acid (21%). The variables with the most missingness in the PostQI Group were last OR pH (21%), last OR base excess (21%), and last OR lactic acid (21%).

Variable Significance in Both Periods

In the PreQI Group, there were 24 variables predicting DCL. Although the majority were intraoperative variables, demographics, injury characteristics, ED variables, and procedures performed also influenced the decision for DCL (Table 3). The variables with the largest MDA in both PreQI models indicating a significant association with DCL were the performance of a bowel resection, ED systolic blood pressure, and OR transfusions (red blood cell and fresh frozen plasma transfusions).

In the PostQI Group, there were half as many variables predicting DCL. All but 3 of the variables were intraoperative variables. The most influential variables within the PostQI models were last OR lactic acid and OR transfusions (red blood cell and fresh frozen plasma transfusions).

While there were some – albeit noticeably fewer than in the PreQI Group – variables related to injury, ED factors, and OR procedures, the strongest predictors of DCL in the PostQI group were OR transfusions and physiology at the end of the case. In the PreQI Group, variables predictive of DCL were spread across the entire continuum of care from admission to the end of the OR. In the PostQI Group, variables most predictive of DCL were OR factors (Figure 1).

Net Mean Decrease in Accuracy

Based on net MDAs (Table 4), the variables with the greatest decrease in importance from the PreQI to the PostQI period were pelvic fracture, liver injury, ED temperature, ED systolic blood pressure, ED red blood cell transfusions, large bowel resection, and last OR base excess (Figure 2). The variables with the greatest increase in importance were surgeon, kidney injury, and nephrectomy.

Discussion

Supervised machine learning identified and quantified factors leading to the changes in surgical decision making for DCL before and after implementation of a successful QI intervention. Whereas the factors that contributed most to the effectiveness of the intervention were previously unknown, machine learning identified shifts in the timing of the decision for DCL and in the emphasis placed on specific physiologic variables as well as associated injuries. Although the frequency of use of the broad indications for DCL did not change, the parameters for defining those indications became more specific.

The decision of whether or not to perform DCL can be challenging for multiple reasons. First, although there are observational studies supporting its use,^{3,4} there are no completed randomized trials demonstrating benefit. In fact, a matched analysis of patients undergoing DCL versus definitive laparotomy at our trauma center suggested that DCL is associated with increased major abdominal complications, although the study was underpowered to identify a statistically significant difference.¹⁸ A study at another major trauma center reported that delay of additional interventions such as vascular procedures associated with DCL resulted in increased mortality, unplanned re-explorations, and re-exploration for hemorrhage control.¹⁹ Second, there is significant uncertainty in the literature and amongst

surgeons regarding the indications for DCL. A scoping review of the literature identified 1,099 indications for damage control surgery.⁹ A subsequent content analysis by international experts in trauma surgery cited 101 appropriate indications for DCL.⁸ Lastly, DCL is associated with increased resource utilization and costs. Higa et al demonstrated that a 78% reduction in DCLs resulted in a projected savings of \$2.2 million in costs without an increase in mortality.⁷

Although this study was conducted at a single center, our pre-intervention decision making closely mirrored that of the trauma community. In an international cross-sectional survey of 366 surgeons, 15 preoperative and 23 intraoperative indications were deemed appropriate to justify damage control surgery.²⁰ Prior to the QI intervention at our institution, 9 preoperative and 16 intraoperative factors were associated with the decision to perform DCL, including demographic, injury, ED, OR, and procedural variables. Several local influencing factors were similar to those cited in the international survey such as preoperative temperature, systolic blood pressure, intraoperative pH and estimated blood loss. After the QI intervention, our consensusbased indications were very similar to the highest-rated indications identified from the literature: injury pattern identified during operation, inability to control bleeding by conventional methods, amount of resuscitation provided, degree of physiologic insult, and need for staged abdominal or thoracic wall reconstruction (i.e., avoidance or treatment of abdominal compartment syndrome).^{8,10}

The QI intervention resulted in a de-emphasis on preoperative factors and an increased reliance on the patient's status at the end of the operation for determination of the degree of physiologic insult and the adequacy of resuscitation. In particular, there was a significant decrease in significance attributed to specific physiologic parameters such as last OR base excess. This shift correlates with a change in indication for DCL in our group from persistent acidosis alone to that associated with ongoing transfusion and continuous vasopressor infusion for hemodynamic instability. There was also a change in use of DCL for specific injury patterns and attempted bleeding control; liver injury and hepatorrhaphy were significantly associated with DCL before but not after the QI intervention. Although there was no change in the rate of packing (PreQI 61% versus PostQI 60%), the rate of hepatorrhaphy increased (PreQI 21% to PostQI 26%, $p=0.070$), and the OR duration increased (median PreQI 109 [IQR 73, 162] minutes versus PostQI 135 [IQR 94, 192], $p<0.001$). These data may reflect more aggressive efforts to obtain hemostasis or increased stability of patients in the OR with adoption of damage control resuscitation.^{21,22}

This study demonstrates how QI, research, and clinical practice can be aligned to generate evidence-based practice to improve patient care. Such agility in a trauma center is representative of the continuous improvement seen in a learning health care system.¹ This study serves as a model for future analyses of pre/post intervention studies which are commonly employed in QI. Although machine learning has been previously used to identify factors involved in decision making,²³ the use of machine learning to explain the effects of a QI intervention is novel. Our study utilized supervised machine learning to identify and weight factors contributing to a successful reduction in DCL. This information can be used to guide future decision making.

There are several limitations to this study. First, as with any single institution study, the results may not be generalizable. Specifically, the safety of avoiding DCL based on early physiologic parameters or injury patterns alone should be further tested in a larger multi-center trial. Surgeons have historically been taught to make the decision to perform damage control as early as possible, making trials of DCL difficult to perform. If surgeons can be convinced that the decision to perform damage control can be delayed, individual indications for DCL may be included in future trials. However, the utilization of machine learning to complement pre- and post-intervention studies is generalizable. Machine learning has been increasingly used to predict outcomes in trauma such as survival, morbidity, length of stay, and hospital admission.²⁴ This study demonstrates a novel use of machine learning to identify factors associated with surgical decision-making. Second, as often seen in clinical studies, the level of missingness was high for some variables. We attempted to address this by requiring agreement between both a complete case model and an imputed model for both time periods before designating a variable MDA or net MDA as significant. Lastly, as this was a study of a pre/post intervention, other variables not accounted for in this analysis may have also affected the decision to perform DCL.

Conclusion

A successful QI intervention to de-implement inappropriate DCL utilization in trauma patients changed surgical decision making by delaying the timing of the decision until near the end of the index operation and by clarifying the criteria for DCL. Supervised machine learning can be used in a learning trauma system to complement QI and research studies to elucidate the effects of interventions on providers' decision making.

Acknowledgments

Support: This work was supported by the Center for Clinical and Translational Sciences, which is funded by National Institutes of Health Clinical and Translational Award UL1 TR000371 and KL2 TR000370 from the National Center for Advancing Translational Sciences.

References

1. National Academies of Sciences E, and Medicine. A national trauma care system: Integrating military and civilian trauma systems to achieve zero preventable deaths after injury. Washington, DC: The National Academies Press; 2016.
2. Deo RC. Machine Learning in Medicine. *Circulation*. 2015;132(20):1920–1930. [PubMed: 26572668]
3. Rotondo MF, Schwab CW, McGonigal MD, et al. 'Damage control': an approach for improved survival in exsanguinating penetrating abdominal injury. *J Trauma*. 1993;35(3):375–382; discussion 382–373. [PubMed: 8371295]
4. Stone HH, Strom PR, Mullins RJ. Management of the major coagulopathy with onset during laparotomy. *Annals of surgery*. 1983;197(5):532–535. [PubMed: 6847272]
5. Harvin JA, Wray CJ, Steward J, et al. Control the damage: morbidity and mortality after emergent trauma laparotomy. *Am J Surg*. 2016;212(1):34–39. [PubMed: 26754456]
6. Miller RS, Morris JA, Jr., Diaz JJ, Jr., Herring MB, May AK. Complications after 344 damage-control open celiotomies. *J Trauma*. 2005;59(6):1365–1371; discussion 1371–1364. [PubMed: 16394910]
7. Higa G, Friese R, O'Keeffe T, et al. Damage control laparotomy: a vital tool once overused. *J Trauma*. 2010;69(1):53–59. [PubMed: 20622578]

8. Roberts DJ, Bobrovitz N, Zygun DA, et al. Indications for Use of Damage Control Surgery in Civilian Trauma Patients: A Content Analysis and Expert Appropriateness Rating Study. *Annals of surgery*. 2016;263(5):1018–1027. [PubMed: 26445471]
9. Roberts DJ, Bobrovitz N, Zygun DA, et al. Indications for use of damage control surgery and damage control interventions in civilian trauma patients: A scoping review. *The journal of trauma and acute care surgery*. 2015;78(6):1187–1196. [PubMed: 26151522]
10. Harvin JA, Kao LS, Liang MK, et al. Decreasing the Use of Damage Control Laparotomy in Trauma: A Quality Improvement Project. *J Am Coll Surg*. 2017;225(2):200–209. [PubMed: 28445796]
11. Harvin JA, Wootton SH, Miller CC. Using Quality Improvement to Promote Clinical Trials of Emergency Trauma Therapies. *JAMA*. 2018.
12. Ovretveit J, Gustafson D. Evaluation of quality improvement programmes. *Qual Saf Health Care*. 2002;11(3):270–275. [PubMed: 12486994]
13. Hartshorn S *Machine Learning With Random Forests And Decision Trees: A Visual Guide For Beginners*. Amazon Digital Services LL; 2016.
14. Kuhn M *Applied Predictive Modeling*. New York: Springer; 2013.
15. Cook D *Practical machine learning with H2O : powerful, scalable techniques for deep learning and AI*. 2017.
16. Package ‘randomForest’.
17. Trickey AW, Fox EE, del Junco DJ, et al. The impact of missing trauma data on predicting massive transfusion. *The journal of trauma and acute care surgery*. 2013;75(1 Suppl 1):S68–74. [PubMed: 23778514]
18. George MJ, Adams SD, McNutt MK, et al. The effect of damage control laparotomy on major abdominal complications: A matched analysis. *Am J Surg*. 2018;216(1):56–59. [PubMed: 29157889]
19. Lauerma MH, Dubose J, Cunningham K, et al. Delayed interventions and mortality in trauma damage control laparotomy. *Surgery*. 2016;160(6):1568–1575. [PubMed: 27499147]
20. Roberts DJ, Zygun DA, Faris PD, et al. Opinions of Practicing Surgeons on the Appropriateness of Published Indications for Use of Damage Control Surgery in Trauma Patients: An International Cross-Sectional Survey. *J Am Coll Surg*. 2016;223(3):515–529. [PubMed: 27321388]
21. Cotton BA, Reddy N, Hatch QM, et al. Damage control resuscitation is associated with a reduction in resuscitation volumes and improvement in survival in 390 damage control laparotomy patients. *Annals of surgery*. 2011;254(4):598–605. [PubMed: 21918426]
22. Holcomb JB, Tilley BC, Baraniuk S, et al. Transfusion of plasma, platelets, and red blood cells in a 1:1:1 vs a 1:1:2 ratio and mortality in patients with severe trauma: the PROPPR randomized clinical trial. *Jama*. 2015;313(5):471–482. [PubMed: 25647203]
23. Lin FP, Pokorny A, Teng C, Dear R, Epstein RJ. Computational prediction of multidisciplinary team decision-making for adjuvant breast cancer drug therapies: a machine learning approach. *BMC Cancer*. 2016;16(1):929. [PubMed: 27905893]
24. Liu NT, Salinas J. Machine Learning for Predicting Outcomes in Trauma. *Shock*. 2017;48(5):504–510. [PubMed: 28498299]

Figure 2.

The net change in variables before and after the quality improvement intervention. There was a general de-emphasis on emergency department variables and bowel resections. The greatest increase in importance was the individual surgeon, likely indicating that the surgeonspecific rates of damage control laparotomy became more disparate after the quality improvement project. AIS, Abbreviated Injury Scale; ED, emergency department; OR, operating room; RBC, red blood cell

Table 1:

Group Characteristics in the Pre-Quality Improvement and Post-Quality Improvement Periods

Characteristic	Pre-quality improvement group 1/1/11–10/31/13 (n=559 [*])	Post-quality improvement group 11/1/13–06/30/16 (n=582 [*])	p Value
All emergent laparotomy, N	582	611	<0.001
Definitive laparotomy, n (%)	339 (58)	455 (74)	
Damage control laparotomy, n (%)	220 (38)	127 (21)	
Intraoperative death, n (%)	23 (4)	29 (5)	
Relative frequency of indication			
Damage control laparotomy, N	220	127	
Packing	134 (61)	76 (60)	0.28
Second look	29 (13)	20 (16)	
Hemodynamic instability	30 (14)	17 (13)	
ACS prophylaxis	3 (1)	6 (5)	
Contamination	7 (3)	1 (1)	
Expedite CT/ICU	14 (6)	5 (4)	
Expedite IR	1 (0)	2 (2)	
Unclear	2 (1)	0 (0)	
Demographics			
Age, y, median (IQR)	34 (24, 48)	32 (23, 46)	0.16
Male sex, n (%)	423(76)	446 (77)	0.70
Penetrating mechanism, n (%)	210 (38)	288 (50)	<0.001
BMI, kg/m ² , median (IQR)	26 (24, 30)	27 (24, 31)	0.14
Emergency department vital signs and labs			
Temperature, Fahrenheit, median (IQR)	97.6 (96.7, 98.3)	97.7 (97.0, 98.3)	0.047
Systolic blood pressure, mmHg, median (IQR)	111 (90, 130)	115 (93, 130)	0.09
Heart rate, beats/minute, median (IQR)	99 (82, 115)	100 (83, 117)	0.59
Glasgow Coma Scale, median (IQR)	15 (6, 15)	15 (12, 15)	0.007
Base excess, mmol/L, median (IQR)	−3 (−7, −1)	−4 (−7, −1)	0.21
Injury severity, median (IQR)			
Head AIS	0 (0, 2)	0 (0, 0)	0.063
Chest AIS	2 (0, 3)	2 (0, 3)	0.51
Abdominal AIS	3 (2, 4)	3 (2, 4)	0.043
Injury Severity Score	19 (11, 34)	18 (10, 29)	0.035
Outcome, n (%)			
Mortality	52 (9%)	65 (11%)	0.30

^{*} N for each column is total emergent laparotomies minus intraoperative deaths, which were excluded.

AIS, abbreviated injury scale; ACS, abdominal compartment syndrome; CT, computed tomography; ICU, intensive care unit; IR, interventional radiology; QI, quality improvement; IQR, 25th and 75th interquartile ranges

Table 2:

Complete Case and Imputed Model Characteristics of the Pre-Quality Improvement and Post-Quality Improvement Groups

Characteristic	PreQI group 1/1/11–10/31/13		PostQI group 11/1/13–06/30/16	
	Complete case model (n=236)	Imputed model (n=559)	Complete case model (n=209)	Imputed model (n=582)
Accuracy	0.772	0.857	0.822	0.873
Sensitivity	0.800	0.850	0.859	0.873
Specificity	0.684	0.873	0.600	0.875
Positive predictive value	0.889	0.930	0.929	0.981
Negative predictive value	0.520	0.743	0.409	0.488
AUC of ROC	0.812	0.921	0.804	0.896

QI, quality improvement; AUC, area under the curve; ROC, receiver operating curve

Table 3:

Mean Decrease in Accuracy (MDA) of Significant Variables Predicting Damage Control Laparotomy (DCL) in the PreQI and PostQI Periods

Variable	Pre QI Group 1/1/11–10/31/13			Pos QI Group 11/1/13–06/30/16		
	Complete case model	Imputed model	Average	Complete case model	Imputed model	Average
MDA of demographic and injury variable						
Abdominal AIS	2	6	4	-	-	-
Injury Severity Score	2	8	5	4	6	5
Large bowel injury	3	4	3.5	-	-	-
Liver injury	5	4	4.5	-	-	-
Kidney injury	-	-	-	4	2	3
Pelvic fracture	4	4	4			
MDA of emergency department variable						
Temperature	6	9	7.5	-	-	-
Systolic blood pressure	6	9	7.5	4	6	5
Base excess	3	9	6	-	-	-
RBC transfusion	5	7	6	-	-	-
MDA of operating room variables						
First pH	6	7	6.5	-	-	-
First base excess	4	4	4	-	-	-
First lactic acid	7	8	7.5	7	8	7.5
RBC transfusion	9	21	15	14	15	14.5
FFP transfusion	7	15	11	15	15	15
Platelet transfusion	4	7	5.5	6	8	7
Estimated blood loss	4	13	8.5	6	9	7.5
Last systolic blood pressure	2	5	3.5	4	6	5
Last temperature	5	3	4	7	5	6
Last pH	6	3	4.5	-	-	-
Last base excess	6	5	5.5	-	-	-
Last lactic acid	6	13	9.5	11	14	12.5
MDA of procedure performed						
Any enteric resection	2	31	16.5	2	3	2.5
Large bowel resection	6	9	7.5	-	-	-
Hepatorrhaphy	7	5	6	-	-	-

Values represent the MDA of each variable scaled to a z-score. Variables were considered significant if the MDA was greater than 2. A larger the MDA represents a larger decrease in the accuracy of the model if that variable were removed, meaning the variable is highly predictive of the outcome (DCL).

AIS, abbreviated injury scale; RBC, red blood cell; FFP, fresh frozen plasma

Table 4:

Net Difference in Mean Decrease in Accuracy (MDA) Between the PreQI and PostQI Groups

Variable	Net complete case model	Net imputed model	Average net difference	Interpretation
Demographic and injury variable				
Penetrating injury	3	3	3	Decrease in importance
Body mass index	4	3	3.5	Decrease in importance
External AIS	2	5	3.5	Decrease in importance
vLiver injury	6	4	5	Decrease in importance
Kidney injury	-6	-2	-4	Increase in importance
Large bowel injury	2	3	2.5	Decrease in importance
Mesentery injury	-4	-2	-3	Increase in importance
Pelvis injury	4	5	4.5	Decrease in importance
Vascular injury	-2	-2	-2	Increase in importance
Emergency department variable				
Temperature	7	2	4.5	Decrease in importance
Systolic blood pressure	7	2	4.5	Decrease in importance
RBC transfusions	5	5	5	Decrease in importance
Operating room variable				
Surgeon	-9	-2	-5.5	Increase in importance
Last temperature	-2	-2	-2	Increase in importance
Last base excess	6	5	5.5	Decrease in importance
Procedure performed				
Hepatorrhaphy	6	3	4.5	Decrease in importance
Large bowel resection	4	9	6.5	Decrease in importance
Nephrectomy	-6	-2	-4	Increase in importance

The net difference in MDA between the PreQI and PostQI groups provided an estimate of the degree of change in surgical decision making and the vector of that change. For example, large bowel resection had the largest decrease in the importance of predicting the use of DCL while surgeon had the largest increase in the importance of predicting the use of DCL.

AIS – abbreviated injury scale; RBC – red blood cell