Nikhil L Chervu MD MS, Jeff Balian, Arjun Verma BS, Sara Sakowitz MS MPH, Nam Yong Cho BS, Saad Mallick MD, Tara A Russell MD MPH PhD, Peyman Benharash MD

Structured Abstract

Objective

To create a novel comorbidity score tailored for surgical database research.

Summary Background Data

Despite their use in surgical research, the Elixhauser (ECI) and Charlson Comorbidity Indices (CCI) were developed nearly four decades ago utilizing primarily non-surgical cohorts.

Methods

Adults undergoing 62 operations across 14 specialties were queried from the 2019 National Inpatient Sample (NIS) using International Classification of Diseases, 10th Revision (ICD-10) codes. ICD-10 codes for chronic diseases were sorted i

0.90 vs 0.84, 95%CI:0.84-0.84), SID (0.91, 95%CI:0.90-0.91 vs 0.86, 95%CI:0.86-0.87), and institutional (0.88, 95%CI:0.87-0.89 vs 0.84, 95%CI:0.83-0.85) databases (all $p<0.001$). Likewise, it outperformed the CCI for the NIS (0.76, 95%CI:0.76-0.76), SID (0.78, 95%CI:0.77-0.78), and institutional (0.62, 95%CI:0.60-0.64) cohorts (all $p<0.001$).

Conclusions

The CORE score may better predict in-hospital mortality after surgery due to comorbid diseases

Introduction

Since their introduction nearly four decades ago, the Charlson Comorbidity Index (CCI) and Elixhauser Comorbidity Index (ECI) have been widely used in healthcare research to quantify the burden of pre-existing conditions.[1–3] Although the superiority of CCI and ECI in specific populations continues to be debated, neither was developed in the context of surgical hospitalizations.[4–6]

In recent years, increasing computational capabilities have enabled healthcare researchers to exploit large databases for highly powered retrospective studie

In the present work, we aimed to create a heuristic tool to assess the association between pre-existing conditions and the risk of in-hospital mortality after major operation. We will validate this metric, the Comorbid Operative Risk Evaluation (CORE) score, using nationwide, state-level, and institutional data to improve and validate its performance. This scoring system may represent an improved discriminatory instrument for future risk models and benchmarking across surgical specialties.

Methods

Data Source and Study Population

The CORE score was developed using the 2019 National Inpatient Sample (NIS). Maintained as part of the Healthcare Costs and Utilization Project (HCUP), NIS is the largest, all-payer inpatient database entailing weighted subsets of individual State Inpatient Databases (SID).[12] Data collected by SID contain approximately 97% of all inpatient discharges within a given state.[13] Each record in the NIS and SID can be associated with 40 diagnoses, which are recorded using *International Classification of Disease, 10th Revision* (ICD-10) codes. Relevant codes are captured by medical billing and coding specialists following each hospitalization from physician notes, operative reports, and radiologic or other diagnostic studies. These codes are further grouped by HCUP into over 530 clinical categories from 22 body systems named "Clinical Classifications Software Refined" (CCSR).[14] CCSR have been used in both clinical research and healthcare utilization analyses to objectively define and classify both acute and chronic conditions in administrative data.[15–17] The *elixhauser* and *charlson* Stata commands were used to calculate the ECI and CCI, respectively.

All hospitalization records for adults (≥18 years) undergoing major neurosurgical, otolaryngologic (ENT), endocrine, cardiac, thoracic, acute care surgery (ACS), foregut/bariatric, hepatopancreatobiliary (HPB), colorectal, urologic, gynecologic, plastic, orthopedic, and further evaluated, whereas those without an associated CCSR were excluded. CCSR groups that were most likely due to non-chronic conditions were then excluded (i.e., CCSR groups with the words “symptom of,” “postoperative,” “postprocedural,” “acute,” or “complication of” in their title). Finally, ICD-10 codes with “acute” in their descriptions were not considered chronic conditions (Supplemental Figure 1, Supplemental Digital Content 2, http://links.lww.com/SLA/F314). In total, 9,811 codes across 325 CCSR were subject to analysis (Supplemental Table 2, Supplemental Digital Content 1, http://links.lww.com/SLA/F313). The primary outcome of the study was in-hospital mortality during the same admission as the operative intervention. This was selected because of its ubiquity as an adverse event across all surgical specialties. Other complications, such as atrial fibrillation, prolonged ventilation, and postoperative transfusion, may be considered less severe or necessary for routine postoperative management after certain cardiac, transplant, and trauma operations. Metrics such as unplanned reoperation and prolonged length of stay are also variable depending on the specialty and operation of interest, thereby limiting the broad applicability of our score.

Covariates included age, sex, elective case status, race, income quartile, primary payer status, bed size, and hospital location/teaching status, in addition to CCSR codes. Model discrimination was assessed using the area under the receiver operating characteristics (AUROC) and precision recall curves (AUPRC) with 95% confidence intervals (CI) generated by 5-fold cross-validation. Probabilistic estimation accuracy was assessed using Brier scores with 95% CI (Supplemental Figure 2, Supplemental Digital Content 2, http://links.lww.com/SLA/F314).[34] True (TPR) and false positive rates (FPR) as well as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), balanced accuracy, and reliability scores were also obtained for each model (Supplemental Table 3, Supplemental Digital Content 1, http://links.lww.com/SLA/F313).

Score Development

After ML-assisted feature down-selection, final score development was conducted using parameter estimates derived from logistic regression. This methodology was initially developed by Sullivan et al., who established a mathematical approach to risk score development for multivariable clinical data as part of the Framingham Heart Study.[35] Notably, this algorithm was used by van Walraven to establish a numerical Elixhauser Comorbidity Index.[2] In summary, points are assigned to each CCSR by obtaining parameter estimates from logistic regression. The final point values are calculated by dividing each logistic coefficient by the lowest CCSR coefficient corresponding to the “weakest” (i.e., lowest absolute value) association with in-hospital mortality (Table 1). A final CORE score is then calculated for each patient using the following equation (where b is the absolute value of the weakest estimate):

$$100 \quad \frac{1}{1 \quad \text{[illegible]}}$$

and non-deceased patients as compared to the ECI (6 [4 - 7] vs 2 [1 - 4], $p<0.001$, Cohen's d = -1.52) and CCI (3 [2 - 5] vs 1 [0 - 2], $p<0.001$, Cohen's d = -1.06). Improved discrimination

context of operative admissions; studies examining their comparative effectiveness in surgical populations are thus mixed.[4–6]

In the present study, we developed the CORE score to incorporate pre-existing conditions more accurately in surgical database research. It represents the first contemporary comorbidity score specifically designed for multispecialty surgical research using retrospective databases. Prior modifications to the CCI and ECI have yet to account for the baseline differences between surgical and non-surgical patients.[40–42] Patients requiring surgical admission often present electively, with preoperative risk stratification and medical optimization prior to surgery.[11,43] Compared to their non-surgical counterparts, surgical patients are younger, less often frail, and have reduced lengths of stay, the diagnosis that restrict its use to purpose-built or institutional databases. However, by incorporating similar techniques with administrative data, we believe that this instrument holds merit as a tool for retrospective health data.

Furthermore, the CORE score provides greater discrimination between deceased and non-deceased patients. In our experience, ECI and CCI medians and confidence intervals between groups can overlap despite the observed statistical significance when dealing with large sample sizes. This is likely due to the inclusion of comorbidities not frequently encountered in surgical populations – congestive heart failure, paralysis, chronic pulmonary disease, renal failure, and liver disease.[2,3] These patients are often deemed too high of a surgical risk to undergo operation. Therefore, the range of possible ECI and CCI values are reduced for surgical patients. When increasing the sample size to tens and hundreds of thousands, however, clinically irrelevant differences can be deemed statistically different. The CORE score increases discrimination by only including comorbidities present in surgical populations and by providing a larger 100-point scale.

Limitations

The present study has several important limitations. As an administrative database, the NIS relies on accurate coding by billing specialists, and may be subject to some error. Furthermore, ICD-10 codes are recorded primarily for financial, and not clinical, purposes. The score is built using the average risk over many people for each condition. We therefore cannot reliably calculate the actual risk of in-hospital mortality for a specific condition for each patient. Finally, rare diagnosis codes that are associated with extremes of risk in mortality may skew the overall score. However, by grouping diagnoses by CCSR and using a large dataset, we attempted to mitigate these risks.

Conclusions

In this large contemporary study, we have established the first comorbidity score for use in administrative database research specifically designed using a surgical cohort. We hope that incorporation of this score in future analyses will allow for more robust adjustment of pre-existing conditions to enhance statistical discrimination. In addition, the increased discriminatory power afforded by a 100-point scale may make subjective analysis of mortality risk easier to determine. Future work applying this score to other studies will allow for continued validation.

Conflict of Interest/Disclosure: PB received fees from AtriCure as a surgical proctor. This manuscript does not discuss any related products or services. Other authors report no conflicts.

Funding/Financial Support: No direct or indirect financial support by extramural sources was received for this research.

Data Sharing Statement: The NIS and SID are publicly available for purchase via the HCUP website. Institutional data is available upon reasonable request with appropriate completion of cross-institutional data sharing agreements. The authors have made a Stata package publicly available within the SSC database. This can be installed using the command *ssc install core*.

Acknowledgements: The authors would like to thank the UCLA Department of Surgery and the Division of Cardiac Surgery, specifically Drs. O. Joe Hines, Timothy Donahue, and Richard Shemin, for the Tpẹ aụsur ! otb D G€ ð 0 @ p °

References

1. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity Measures for Use with Administrative Data. *Med Care*. 1998;36(1):8-27.

2. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A Modification of the Elixhauser Comorbidity Measures into a Point System for Hospital Death Using Administrative Data. *Med Care*. 2009;47(6):626-633.

3. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chronic Dis*. 1987;40(5):373-383. doi:10.1016/0021-9681(87)90171-8

4. Maron SZ, Neifert SN, Ranson WA, et al. Elixhauser Comorbidity Measure is Superior to Charlson Comorbidity Index In-Predicting Hospital Complications Following Elective Posterior Cervical Decompression and Fusion.

12. HCUP-US NIS Overview. Accessed April 11, 2023. https://hcup-us.ahrq.gov/nisoverview.jsp

13. HCUP-US SID Overview. Accessed May 17, 2024. https://hcup-us.ahrq.gov/sidoverview.jsp

14. Clinical Classifications Software Refined (CCSR). Accessed May 17, 2024. https://hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp

15. Johnston S, Jha A, Roy S, Pollack E. Surgical Complication Risk Factor Identification Using High-Dimensional Hospital Data: An Illustrative Example in Hemostasis-Related Complications.

24. Till BM, Mack SJ, Whitehorn G, et al. The Epidemiology of Surgically Managed Hiatal Hernia: A Nine Year Review of National Trends. *Foregut*. 2023;3(1):20-28. doi:10.1177/26345161221137089

25. Ng AP, Sanaiha Y, Bakhtiyar SS, Ebrahimian S, Branche C, Benharash P. National analysis of cost disparities in robotic-assisted versus laparoscopic abdominal operations. *Surgery*. 2023;173(6):1340-1345. doi:10.1016/j.surg.2023.02.016

26. Mabeza RM, Mao Y, Maynard K, Lee C, Benharash P, Yetasook A. Bariatric surgery outcomes in geriatric patients: a contemporary, nationwide analysis. *Surg Obes Relat Dis*. 2022;18(8):1005-1011. doi:10.1016/j.soard.2022.04.014

27. Williamson CG, Ebrahimian S, Sakowitz S, et al. Race, Insurance, and Sex-Based Disparities in Access to High-Volume Centers for Pancreatectomy. *Ann Surg Oncol*. 2023;30(5):3002-3010. doi:10.1245/s10434-022-13032-8

28. Madrigal J, Hadaya J, Lee C, Tran Z, Benharash P. Association of Frailty with Perioperative Outcomes Following Hepatic Resection: A National Study. *J Am Med Dir Assoc*. 2022;23(4):684-689.e1. doi:10.1016/j.jamda.2022.02.004

29. Mazzone E, Mistretta FA, Knipper S, et al. Contemporary National Assessment of Robot-Assisted Surgery Rates and Total Hospital Charges for Major Surgical Uro-Oncological Procedures in the United States. *J Endourol*. 2019;33(6):438-447. doi:10.1089/end.2018.0840

30. Chan JK, Gardner AB, Mann AK, Kapp DS. Hospital-acquired conditions after surgery for gynecologic cancer — An analysis of 82,304 patients. *Gynecol Oncol*. 2018;150(3):515-520. doi:10.1016/j.ygyno.2018.07.009

31. Sayegh F, Warburton A, Sanati-Mehrizy P, Burish NM, Taub PJ. Determining the Factors Associated With 30- and 90-day Readmissions Following Plastic Surgery Within the National Readmissions Database. *Plast Reconstr Surg – Glob Open*. 2020;8(9S):132. doi:10.1097/01.GOX.0000721012.03765.82

32. Choi YS, Kim TW, Chang MJ, Kang SB, Chang CB. Enhanced recovery after surgery for major orthopedic surgery: a narrative review. *Knee Surg Relat Res*. 2022;34(1):8. doi:10.1186/s43019-022-00137-3

33. Beam AL, Kohanbr

36. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Routledge; 2013.

37. Guller U. Surgical Outcomes Research Based on Administrative Data: Inferior or Complementary to Prospective Randomized Clinical Trials? *World J Surg*. 2006;30(3):255-266. doi:10.1007/s00268-005-0156-0

38. Talari K, Goyal M. Retrospective Studies – Utility and Caveats. *J R Coll Physicians Edinb*. 2020;50(4):398-402. doi:10.4997/jrcpe.2020.409

39. Xiang J, Durance PW, Griffes LC, Chen Y, Bakshi RR. Measuring case severity: a novel tool for benchmarking and clinical documentation improvement. *BMC Health Serv Res*. 2022;22(1):513. doi:10.1186/s12913-022-07935-1

40. Holman CDJ, Preen DB, Baynham NJ, Finn JC, Semmens JB. A multipurpose comorbidity scoring system performed better than the Charlson index. *J Clin Epidemiol*. 2005;58(10):1006-1014. doi:10.1016/j.jclinepi.2005.01.020

41. Quan H, Li B, Couris CM, et al. Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries. *Am J Epidemiol*. 2011;173(6):676-682. doi:10.1093/aje/kwq433

42. Klabunde CN, Potosky AL, Legler JM, Warren JL. Development of a comorbidity index using physician claims data. *J Clin Epidemiol*. 2000;53(12):1258-1267. doi:10.1016/S0895-4356(00)00256-0

43. Boeker EB, Ram K, Klopotowska JE, et al. An individual patient data meta-analysis on factors associated with adverse drug events in surgical and non-surgical inpatients. *Br J Clin Pharmacol*. 2015;79(4):548-557. doi:10.1111/bcp.12504

44. Wang J, Chen K, Li X, et al. Postoperative adverse events in patients with diabetes undergoing orthopedic and general surgery. *Medicine (Baltimore)*. 2019;98(14):e15089. doi:10.1097/MD.0000000000015089

45. Bilimoria KY, Liu Y, Paruch JL, et al. Development and Evaluation of the Universal ACS NSQIP Surgical Risk Calculator: A Decision Aid and Informed Consent Tool for Patients and Surgeons. *J Am Coll Surg*. 2013;217(5):833-842.e3. doi:10.1016/j.jamcollsurg.2013.07.385

46. Raphael CE, Sabbagh AE, Rihal C. CHAPTER 7 - Transcatheter aortic valve implantation: Work up and indications for intervention. In: Rihal CS, Raphael CE, eds. *Handbook of Structural Heart Interventions*. Elsevier; 2021:81-87. doi:10.1016/B978-0-323-67278-8.00016-8

47. Moll M, Qiao D, Regan EA, et al. Machine Learning and Prediction of All-Cause Mortality in COPD. *Chest*. 2020;158(3):952-964. doi:10.1016/j.chest.2020.02.079

48. Panay B, Baloian N, Pino JA, Peñafiel S, Sanson H, Bersano N. Feature Selection for Health Care Costs Prediction Using Weighted Evidential Regression. *Sensors*. 2020;20(16):4392. doi:10.3390/s20164392

49. Sharma A, Mishra PK. Performance analysis of machine learning based optimized feature selection approaches for breast cancer diagnosis. *Int J Inf Technol*. 2022;14(4):1949-1960. doi:10.1007/s41870-021-00671-5

50. Bertsimas D, Dunn J, Velmahos GC, Kaafarani HMA. Surgical Risk Is Not Linear: Derivation and Validation of a Novel, User-friendly, and Machine-learning-based Predictive OpTimal Trees in Emergency Surgery Risk (POTTER) Calculator. *Ann Surg*. 2018;268(4):574. doi:10.1097/SLA.0000000000002956

Figure 1. CONSORT diagram of patients included from the 2019 National Inpatient Sample to build the Comorbid Operative Risk Evaluation Score

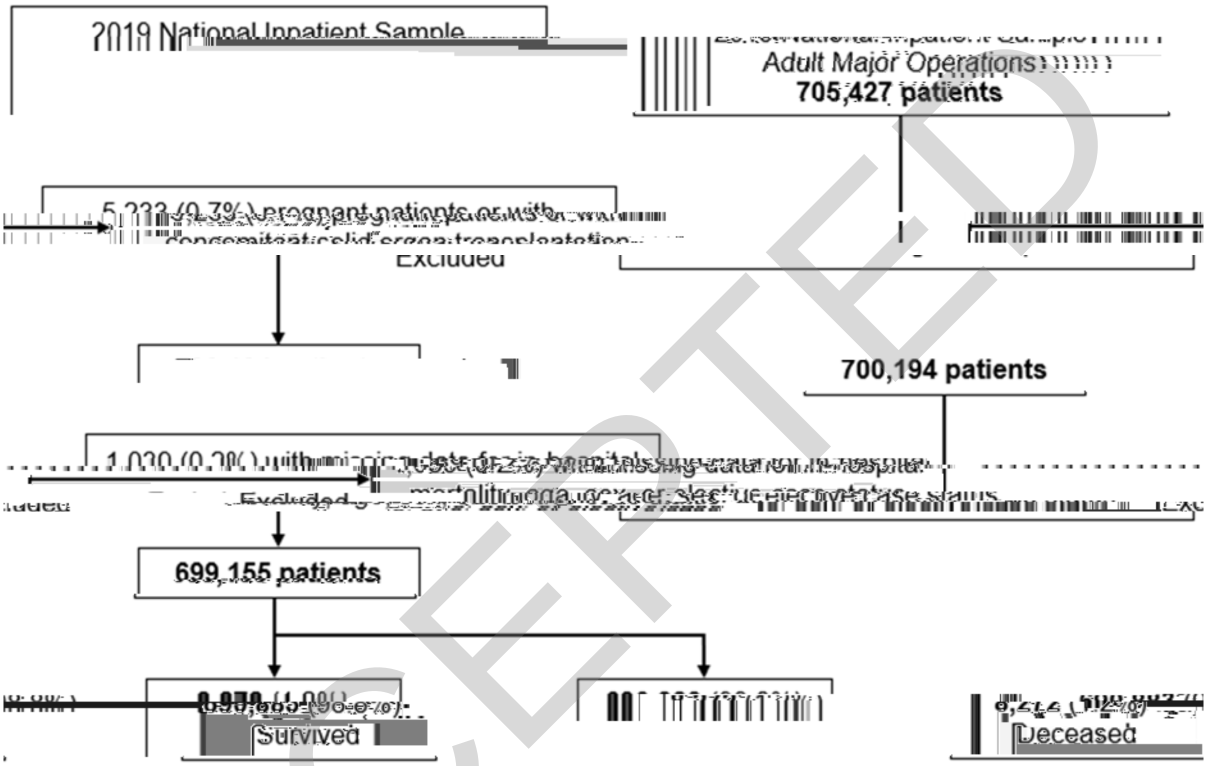

Figure 2. Area Under the Receiver Operator Characteristics (AUROC) with 95% confidence intervals between the Comorbid Operative Risk Evaluation (CORE) Score, Elixhauser Comorbidity Index (ECI), and Charlson Comorbidity Index (CCI) in predicting in-hospital mortality for the National Inpatient Sample (NIS), combined Florida and New York State Inpatient Database (SID), and institutional data

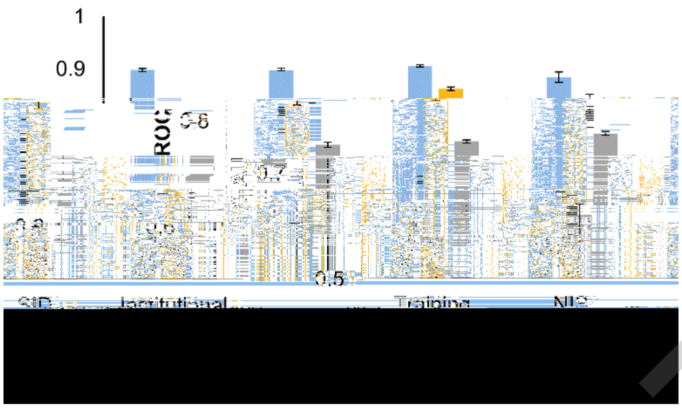

Figure 3. Comparison of Area Under the Receiver Operator Characteristics (AUROC) with 95% confidence intervals between the Comorbid Operative Risk Evaluation (CORE) Score, Elixhauser Comorbidity Index (ECI), and Charlson Comorbidity Index (CCI) by year in predicting in-hospital mortality for the a) 2016-2018 National Inpatient Sample (NIS), b) 2018-2019 combined Florida and New York State Inpatient Database (SID), and c) 2016-2022 institutional data

Figure 4. Comparison of Area Under the Receiver Operator Characteristics (AUROC) with 95% confidence intervals between the Comorbid Operative Risk Evaluation (CORE) Score, Elixhauser Comorbidity Index (ECI), and Charlson Comorbidity Index (CCI) by specialty in predicting in-hospital mortality for the a) 2016-2018 National Inpatient Sample (NIS), b) 2018-2019 combined Florida and New York State Inpatient Database (SID), and c) 2016-2022 institutional data; ENT, otolaryngology; ACS, acute care surgery; HPB, hepatopancreatobiliary

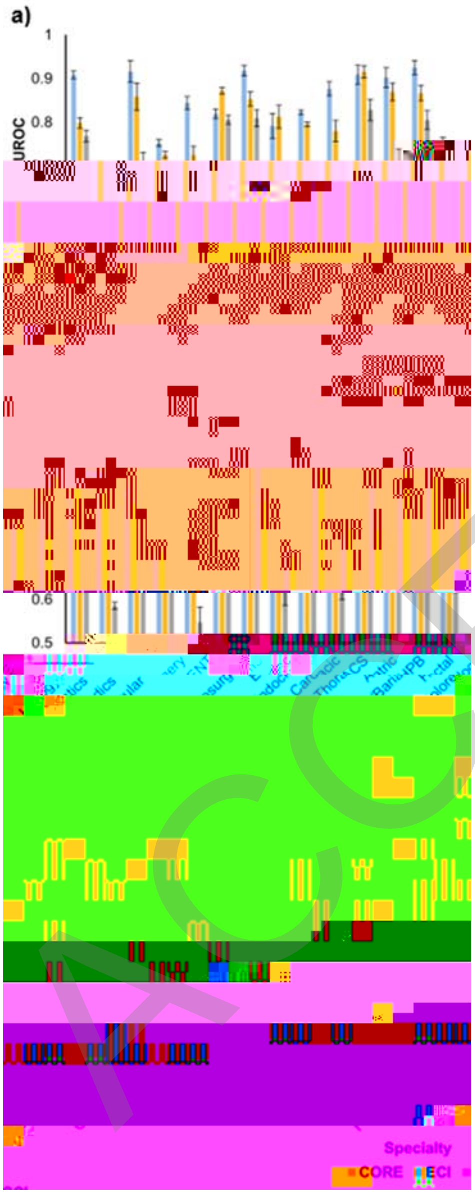

Table 1. Logistic regression coefficients and point totals of Clinical Classifications Software

Table 2. Patient, clinical, hospital characteristics testing and training cohorts derived from the 2019 National Inpatient Sample used to develop the Comorbid Operative Risk Evaluation (CORE) Score; IQR, interquartile range; ECI, Elixhauser Comorbidity Index; CCI, Charlson Comorbidity Index; ENT, ear nose and throat; ACS, acute care surgery; HPB, hepatopancreatobiliary

	Testing	**Training**	**p-value**	**Cohen's d**
Age (years, median [IQR])	64 [54 - 72]	64 [54 - 72]	0.85	< 0.01
Female (%)	75,003 (53.6)	299,887 (53.6)	0.88	< - 0.01