Received: 17 January 2023 | Revised: 21 April 2023 | Accepted: 8 May 2023

DOI: 10.1002/ags3.12704

AGSurg Annals of Gastroenterological Surgery Open Access WILEY

ORIGINAL ARTICLE

Identification of complications requiring interventions after gastrointestinal cancer surgery from real-world data: An external validation study

Hiromitsu Kinoshita[1] | Tatsuto Nishigori[1,2] | Susumu Kunisawa[3] | Koya Hida[1] | Hisahiro Hosogi[4] | Susumu Inamoto[4] | Hiroaki Hata[5] | Ryo Matsusue[5,6] | Yuichi Imanaka[3] | Kazutaka Obama[1] | Yumi Matsumura[2]

[1]Department of Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan

[2]Department of Patient Safety, Kyoto University Hospital, Kyoto, Japan

[3]Department of Healthcare Economics and Quality Management, Graduate School of Medicine, Kyoto University, Kyoto, Japan

[4]Department of Surgery, Japanese Red Cross Osaka Hospital, Osaka, Japan

[5]Department of Surgery, National Hospital Organization, Kyoto Medical Center, Kyoto, Japan

[6]Department of Gastrointestinal Surgery, Tenri Hospital, Nara, Japan

Correspondence
Tatsuto Nishigori, Department of Patient Safety, Kyoto University Hospital, 54 Shogoin-Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan
Email: nsgr@kuhp.kyoto-u.ac.jp

Abstract

Background: Recently, real-world data have been recognized to have a significant role for research and quality improvement worldwide. The decision on the existence or nonexistence of postoperative complications is complex in clinical practice. This multicenter validation study aimed to evaluate the accuracy of identification of patients who underwent gastrointestinal (GI) cancer surgery and extraction of postoperative complications from Japanese administrative claims data.

Methods: We compared data extracted from both the Diagnosis Procedure Combination (DPC) and chart review of patients who underwent GI cancer surgery from April 2016 to March 2019. Using data of 658 patients at Kyoto University Hospital, we developed algorithms for the extraction of patients and postoperative complications requiring interventions, which included an invasive procedure, reoperation, mechanical ventilation, hemodialysis, intensive care unit management, and in-hospital mortality. The accuracy of the algorithms was externally validated using the data of 1708 patients at two other hospitals.

Results: In the overall validation set, 1694 of 1708 eligible patients were correctly extracted by DPC (sensitivity 0.992 and positive predictive value 0.992). All postoperative complications requiring interventions had a sensitivity of >0.798 and a specificity of almost 1.000. The overall sensitivity and specificity of Clavien–Dindo ≥grade IIIb complications was 1.000 and 0.995, respectively.

Conclusion: Patients undergoing GI cancer surgery and postoperative complications requiring interventions can be accurately identified using the real-world data. This multicenter external validation study may contribute to future research on hospital quality improvement or to a large-scale comparison study among nationwide hospitals using real-world data.

KEYWORDS
administrative claims, gastrointestinal neoplasms, patient outcome assessment, postoperative complications, validation study

1 | INTRODUCTION

To improve the quality of medical care, identification of patients who underwent a medical intervention and evaluation of the outcomes following the intervention are essential. In addition, large-scale clinical research among multiple institutions is needed to evaluate the effectiveness of such interventions.

Administrative claims databases have played a significant role as real-world data for research and quality improvement worldwide.[1,2] Claims data contain records related to medical procedures; therefore, it should be possible to gain information on medical interventions and outcomes. Unlike manual development of a clinical database, utilization of electronic claims data can save data collection time or cost.[3] In the field of gastrointestinal (GI) surgery, utilization of claims databases has been increasing.[4–6]

However, the decision on the existence or nonexistence of postoperative complications is complex in clinical practice; hence, it cannot be mechanically handled. The accuracy of extraction of postoperative complications from real-world data should be considered carefully but has been rarely validated. The use of incorrectly extracted data for evaluation of surgical intervention and quality management could lead to inappropriate conclusions or countermeasures.[7,8] The present study aimed to develop and validate algorithms for identifying patients who underwent GI cancer surgery and for extracting the corresponding outcomes from claims data in Japan.

2 | METHODS

2.1 | Study design and participants

This was a retrospective cohort study conducted across three hospitals. First, algorithms for extracting cases of GI cancer surgery from claims data were developed using data at Kyoto University Hospital. In addition, algorithms for extracting postoperative complications from claims data were developed. Next, using data at two other hospitals (i.e. Kyoto Medical Center and Osaka Red Cross Hospital), we evaluated the accuracy of the developed algorithms for extracting the cases and complications.

The eligibility criteria were consecutive patients who underwent esophagectomy, gastrectomy, colectomy, or proctectomy for GI cancer between April 2016 and March 2019 at each participating hospital. Melanoma, lymphoma, sarcoma, gastrointestinal stromal tumor, and other nonepithelial tumors were excluded. Those who underwent concurrent or consecutive surgeries for multiple GI cancers; underwent surgery that was not covered by the national health insurance; and paid their own expenses were excluded. Because a mix of both insured and out-of-pocket treatment has been prohibited in Japan, except for special circumstances, almost all patients with resectable GI cancer during the study period were presumed to have undergone surgery using the national health insurance.

2.2 | Clinical data as the reference standard

Each hospital has prospectively maintained its own clinical database including data of postoperative complications. However, to create accurate and reliable reference standard information, the chart review was performed by H.K. First, original lists of patients who were recorded in the electronic medical record as entering the operation rooms were used to identify patients who underwent any surgeries between April 2016 and March 2019 at each hospital. Thereafter, a chart review was performed to extract those who met the eligibility criteria. Next, postoperative complications were recorded according to the Clavien–Dindo classification and were standardized by the Japan Clinical Oncology Group criteria.[9,10] Discrepancies between a hospital-maintained clinical database and chart review were resolved by involving a third physician (i.e. T.N., K.H., H.H., R.M., H.H., or S.I.). The completed lists of patients and postoperative complications were treated as a reference standard.

2.3 | Administrative claims data

Diagnosis Procedure Combination (DPC) data were used to extract cases with GI cancer surgery and postoperative complications. Japan introduced a medical reimbursement system known as the DPC-based Per-Diem Payment System (DPC/PDPS) in 2003.[11] DPC data were originally designed to accompany the DPC/PDPS and was previously required to be submitted to the Ministry of Health, Labour and Welfare, only by hospitals adopting this system. However, recently many (over 5000) hospitals have prepared and submitted DPC data to the MHLW.[12] In particular, submission of data are now mandatory in acute-care hospitals,[13] and using DPC data enables standardized data analysis related to surgery at hospitals throughout Japan. DPC data are recorded electronically in a standardized format and include a brief record of medical information, such as patient characteristics, some clinical information, admission and discharge statuses, and diagnoses. Diagnoses are recorded with the International Classification of Diseases, 10th revision codes, as well as Japanese original codes. Based on the diagnosis codes, the following five diagnoses were selected to represent hospitalization: most resource-consuming; second most resource-consuming; main; trigger for admission; and comorbidity at admission. All medical treatment procedures, examinations, or prescriptions during hospitalization are recorded using Japanese original codes.[14]

2.4 | Development of algorithms for extracting GI cancer surgery cases

Algorithms for extracting GI cancer surgery cases from the DPC data were developed using the data of 658 patients at Kyoto University Hospital. Both the codes of the operative method and the corresponding diagnosis were used to achieve a highly accurate extraction.[15] The extracted data from the DPC data and clinical data created by chart review were compared, and algorithms were modified several times in order to improve the accuracy.

The final inclusion criteria to extract GI cancer surgery cases from the DPC data were esophagectomy or gastrectomy for upper GI cancer and colectomy or proctectomy for lower GI cancer. The exclusion criteria to extract GI cancer surgery cases from the DPC data were cases with multiple codes of esophagectomy, gastrectomy, colectomy, or proctectomy during the same hospitalization, except additional upper GI surgeries after esophagectomy or gastrectomy and additional lower GI surgeries after colectomy or proctectomy. Cancer diagnosis codes were extracted from the most resource-consuming, second most resource-consuming, main, or trigger diagnosis for admission. This combination was decided based on analysis of the identification of patients among eight potential patterns of a combination of diagnosis in the DPC data. The results are listed in Table S1. The codes of each operative method and cancer diagnosis are shown in Appendices S1 and S2.

2.5 | Development of algorithms for extracting postoperative complications

Algorithms for extracting postoperative complications requiring interventions from the DPC data were developed using the data of patients at Kyoto University Hospital. Postoperative complications were defined in accordance with the general concepts of the Clavien–Dindo classification ≥grade III; invasive procedure, that is surgical, endoscopic, or radiological intervention not under general anesthesia; reoperation under general anesthesia; mechanical ventilation; hemodialysis; intensive care unit (ICU) management; and in-hospital mortality.

The final algorithms for extracting postoperative complications are shown in Table 1. In terms of mechanical ventilation, two algorithms were developed. Extubation is sometimes performed on postoperative day (POD) 1 or thereafter even without complications, especially in patients undergoing esophagectomy. Therefore, in one algorithm (i.e. mechanical ventilation algorithm A), patients who had the codes of mechanical ventilation for consecutive 1, 2, or 3 d after surgery were regarded as having no complication. In another algorithm (i.e. mechanical ventilation algorithm B), patients who had the codes for 1 or 2 consecutive d after surgery were regarded as having no complication.

Similarly, two algorithms were developed for extracting ICU management. When ICU admission was scheduled after surgery, the ICU stay can continue after POD 1, even if the patient did not develop complications. Therefore, patients who had the codes of ICU stay for 1 or 2 consecutive d after surgery were regarded as having no complication in one algorithm (i.e. ICU management algorithm A). In another algorithm (i.e. ICU management algorithm B), patients who had the codes of ICU stay for 1, 2, 3, 4, 5, or 6 consecutive d after surgery were regarded as having no complication.

The analyses of the accuracy of the final algorithms at Kyoto University Hospital are summarized in Table S2.

2.6 | Validation of the developed algorithms

We evaluated the accuracy of the developed final algorithms at Kyoto Medical Center (Hospital A) and Osaka Red Cross Hospital (Hospital B) in extracting cases of GI cancer surgery and postoperative complications requiring interventions. Data were extracted from the DPC, while being blinded to the clinical data. The accuracy of following extracted data was evaluated: patients who underwent GI cancer surgery; patient characteristics, such as age and sex; operative method; surgical approach; each postoperative complication; overall postoperative complications ≥grade IIIa and ≥grade IIIb; and postoperative course, such as length of postoperative hospital stay, discharge destination, and 30-d unplanned hospital readmissions. Overall postoperative complication was counted as the use of mechanical ventilation algorithm and ICU management algorithm A.

2.7 | Statistical analysis

The accuracy of identifying patients who underwent GI cancer surgery was calculated as the sensitivity and positive predictive value (PPV). Sensitivity was defined as the proportion of patients extracted from the DPC data among eligible patients by chart review. PPV was defined as the proportion of eligible patients by chart review among patients extracted from the DPC data. Moreover, we examined the reasons for incorrect identification of patients. Since patients who visited the hospital for diseases other than GI cancer or pharmaceutical treatment were out of the scope of this study, the number of patients who did not undergo GI cancer surgery was not tallied. Therefore, specificity and negative predictive value (NPV) were not calculated in the analyses, which evaluated the accuracy of identifying patients who underwent GI cancer surgery.

Patient characteristics, postoperative complications, and course were analyzed in cases that could be identified in both DPC and chart review. The sensitivity, specificity, PPV, and NPV were examined. Sensitivity was defined as the proportion of patients extracted as positive from the DPC data among those identified as positive by chart review. Specificity was defined as the proportion of patients

TABLE 1 Definitions of postoperative complications based on the administrative data

Invasive procedure	The codes of the invasive procedures (Appendix S3) were assigned on POD 1 or thereafter. If the codes of general anesthesia (Appendix S4) and specific interventions (Appendix S5) were assigned on the same day, the events were regarded as reoperation and excluded from this outcome.
Reoperation	The codes of general anesthesia (Appendix S4) or tracheostomy (Appendix S6) were assigned on POD 1 or thereafter. Pulmonary resection for malignant tumors; liver resection, including microwave ablation or radiofrequency ablation, for malignant tumors; stoma closure; or two-stage reconstruction after esophagectomy or proctectomy was not regarded as reoperation.
Mechanical ventilation algorithm A	1. The code of postoperative intubation as a life-saving procedure (Appendix S7) was assigned on POD 0 or thereafter; OR 2. The code of tracheotomy (Appendix S6) was assigned, followed by that of mechanical ventilation (Appendix S8) on POD 1 or thereafter; OR 3. [The codes of mechanical ventilation, excluding nasal-related mechanical ventilation, were assigned on POD 4 or thereafter; OR the codes of mechanical ventilation were assigned on POD 3 but not on POD 2; OR the codes of mechanical ventilation were assigned on POD 2, but not on POD 1.] AND [In cases of reoperation, consecutive codes, beginning from the day of reoperation until day 3 after, were not regarded as the outcome. In cases with respiratory-related codes (Appendix S9) before surgery, postoperative mechanical ventilation was not regarded as the outcome.]
Mechanical ventilation algorithm B	1. The code of postoperative intubation as a life-saving procedure was assigned on POD 0 or thereafter; OR 2. The code of tracheotomy was assigned, followed by that of mechanical ventilation on POD 1 or thereafter; OR 3. [The codes of mechanical ventilation, excluding nasal-related mechanical ventilation, were assigned on POD 3 or thereafter; OR the codes of mechanical ventilation were assigned on POD 2, but not on POD 1] AND [In cases of reoperation, consecutive codes, beginning from the day of reoperation until day 2 after, were not regarded as the outcome. In cases with respiratory-related codes before surgery, postoperative mechanical ventilation was not regarded as the outcome.]
Hemodialysis	The codes of acute hemodialysis (Appendix S10) were assigned on POD 1 or thereafter. AND In cases with codes of hemodialysis (Appendix S11) before surgery or those with the diagnosis of chronic or end-stage renal failure as comorbidities, postoperative hemodialysis was not regarded as the outcome.
ICU management algorithm A	The codes of ICU management (Appendix S12) were assigned on POD 3 or thereafter; OR The codes of ICU management were assigned on POD 2 but not on POD 1.
ICU management algorithm B	The codes of ICU management were assigned on POD 7 or thereafter; OR The codes of ICU management were assigned intermittently between POD 1 and 6.
In-hospital mortality	Death was registered as the discharge outcome.

Note: If the codes of scheduled surgery (Appendix S13), such as pulmonary resection, liver resection, stoma closure, or two-stage reconstruction after esophagectomy, following GI cancer surgery in practice were assigned during the hospitalization, extraction of any outcomes after the scheduled surgery was not counted.

Abbreviations: ICU, intensive care unit; POD, postoperative day.

identified as negative in the DPC data among those identified as negative by chart review. PPV indicated the proportion of patients who were diagnosed as positive by chart review among those extracted as positive from the DPC data. NPV indicated the proportion of patients who were diagnosed as negative by chart review among those identified as negative in the DPC data. In addition, we reviewed the reasons for incorrect identification of these outcomes.

There is no clear standard on the accuracy of data extracted from health administrative data.[16] However, because a PPV of >0.80 is considered to gain a homogeneous population, this value was used to interpret the accuracy of identifying patients who underwent GI cancer surgery.[17] Specificity has been emphasized more than sensitivity when interpreting the incidence of events in a comparative study. Especially when the outcome is a rare event, a decrease in specificity can cause a significant risk of bias.[18] If specificity values would have been approximately 1.00 or ≥0.90 at worst, we expected that the accuracy of extracting postoperative complications would be interpreted.

Although a specific number of patients was difficult to set, a sample size was estimated, as described in this paragraph. A sample size was set in consideration of estimated sensitivity, specificity of postoperative outcomes, and their 95% confidence intervals (CIs).[19] The sensitivity and specificity were assumed to be 90%, with 10% of the maximum marginal error of estimate under 95% CIs. Assuming a prevalence of 5%–10% for postoperative invasive procedure, which might be the most difficult to extract,[20–23] a sample size of 350–700 cases at each hospital was considered desirable.[19] Considering the annual number of surgical cases in Hospitals A and B, patients seen within a 3-y period were included in the present study.

For all diagnostic accuracy statistics, 95% CIs were calculated using a normal approximation of binomial distribution. All statistical analyses were conducted using R version 3.6.2 software (Vienna, Austria).

3 | RESULTS

3.1 | Identification of patients who underwent GI cancer surgery

After chart review at Hospitals A and B, a total of 1708 patients (608 at Hospital A and 1100 at Hospital B) who underwent GI cancer surgery were identified. From the DPC data of the two hospitals, 1694 patients were extracted. As shown in Table 2, the sensitivity was 0.992 (95% CI, 0.986–0.996) and the PPV was 0.992 (95% CI, 0.987–0.996). Analysis by each hospital showed similar favorable accuracy (Table S3). The details of the reasons for incorrect identification of patients are shown in Table S4. For example, cancer diagnosis codes were not included in eight patients and were incorrectly assigned for noncancer diseases in three patients.

3.2 | Accuracy of the extracted patient characteristics

Validation analyses of the matched 1694 patients (603 at Hospital A and 1091 at Hospital B) were performed. Table 3 shows the accuracy of the patient characteristics extracted from the DPC data. There were no incorrect age and sex data. In terms of operative methods, the sensitivity and specificity were above 0.985, except for proximal gastrectomy. Although 29 patients underwent proximal gastrectomy in practice, data extraction of operative methods was correct in 15 patients. On the other hand, in 14 patients there were miscoding errors, such as distal gastrectomy, probably caused by the infrequent use of proximal gastrectomy codes. For the surgical approach, the sensitivity, specificity, PPV, and NPV were all ≥0.974.

3.3 | Accuracy of the extracted postoperative complications

Table 4 shows the results of the validation analyses of postoperative complications. For all outcomes, the sensitivity was >0.798, and the specificity was almost 1.000. The reasons for incorrect identification of postoperative complications are listed in Table S5.

The invasive procedure, which was surgical, endoscopic, or radiological intervention not under general anesthesia, was actually performed in practice on 89 patients. However, the codes of the complications were not recorded in the DPC data in 18 patients (sensitivity 0.798; 95% CI, 0.699–0.876).

All reoperations that were actually performed in practice on 35 patients were correctly extracted from the DPC data (sensitivity 1.000; 95% CI, 0.900–1.000). Scheduled consecutive surgeries for other diseases following GI cancer surgery were incorrectly extracted as reoperation in four patients (specificity 0.998; 95% CI, 0.994–0.999).

Mechanical ventilation was extracted according to two algorithms. All cases of mechanical ventilation that was actually performed in practice ($n = 21$) were correctly extracted by both algorithms (sensitivity 1.000; 95% CI, 0.839–1.000). The specificity of mechanical ventilation extracted by the two algorithms was 0.998 and 0.997, respectively. Because NPPV was recorded as mechanical ventilation in the DPC data, three and four patients who underwent NPPV were incorrectly regarded as receiving mechanical ventilation in each algorithm, respectively.

Moreover, ICU management was extracted according to two algorithms. In practice, 25 patients were admitted to the ICU for life-threatening complications. Extraction was correct in 23 patients using algorithm A (sensitivity 0.920; 95% CI, 0.740–0.990) and in 21 patients using algorithm B (sensitivity 0.840; 95% CI, 0.639–0.955). Two and four patients were discharged from the ICU within POD 3 and 7, respectively, after requiring immediate postoperative ICU management because of critical postoperative complications. Those patients were regarded as having no ICU management in each algorithm. In addition, the specificity of algorithm A and B was 0.998 and 0.999, respectively.

For hemodialysis and in-hospital mortality, all diagnostic accuracy statistics were 1.000.

TABLE 2 Accuracy of identification of patients who underwent gastrointestinal cancer surgery

TP	TN	FN	FP	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
1694	NA	14	13	0.992 (0.986–0.996)	NA	0.992 (0.987–0.996)	NA

Note: TP was defined as a case that was identified in both DPC and chart review. FN was defined as a case that was identified by chart review but not in the DPC. FP was defined as a case that was identified in the DPC but not by chart review. Sensitivity was defined as the proportion of patients extracted from the DPC data among eligible patients in the chart review. PPV was defined as the proportion of eligible patients in the chart review among patients extracted from the DPC data.

Abbreviations: CI, confidence interval; FN, false negative; FP, false positive; NA, not applicable; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

TABLE 3 Accuracy of identification of patient characteristics ($n = 1694$)

	TP	TN	FN	FP	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Age[a]	NA	NA	NA	NA	NA	NA	NA	NA
Sex, male	986	708	0	0	1.000 (0.996–1.000)	1.000 (0.995–1.000)	1.000 (0.996–1.000)	1.000 (0.995–1.000)
Operative method								
Esophagectomy	105	1589	0	0	1.000 (0.965–1.000)	1.000 (0.998–1.000)	1.000 (0.965–1.000)	1.000 (0.998–1.000)
Total gastrectomy	133	1559	2	0	0.985 (0.948–0.998)	1.000 (0.998–1.000)	1.000 (0.973–1.000)	0.999 (0.995–1.000)
Proximal gastrectomy	15	1664	14	1	0.517 (0.325–0.706)	0.999 (0.997–1.000)	0.938 (0.698–0.998)	0.992 (0.986–0.995)
Distal gastrectomy	352	1325	1	16	0.997 (0.984–1.000)	0.988 (0.981–0.993)	0.957 (0.930–0.975)	0.999 (0.996–1.000)
Colectomy	679	1012	1	2	0.999 (0.992–1.000)	0.998 (0.993–1.000)	0.997 (0.989–1.000)	0.999 (0.995–1.000)
Proctectomy	390	1301	2	1	0.995 (0.982–0.999)	0.999 (0.996–1.000)	0.997 (0.986–1.000)	0.998 (0.994–1.000)
Surgical approach								
Open	147	1542	1	4	0.993 (0.963–1.00)	0.997 (0.993–0.999)	0.974 (0.934–0.993)	0.999 (0.996–1.000)
Laparoscopic/ Thoracoscopic	1475	213	4	2	0.997 (0.993–0.999)	0.991 (0.967–0.999)	0.999 (0.995–1.000)	0.982 (0.953–0.995)
Robotic	66	1627	1	0	0.985 (0.920–1.000)	1.000 (0.998–1.000)	1.000 (0.946–1.000)	0.999 (0.997–1.000)
Mediastinoscopy	0	1694	0	0	NA	1.000 (0.998–1.000)	NA	1.000 (0.998–1.000)

Note: TP was defined as a case that was identified in both DPC and chart review. TN was defined as a case that was not identified in both DPC and chart review. FN was defined as a case that was identified by chart review but not in the DPC. FP was defined as a case that was identified in the DPC but not by chart review sensitivity = TP/(TP + FN); specificity = TN/(FP + TN); PPV = TP/(TP + FP); NPV = TN/(FN + TN).

Abbreviations: CI, confidence interval; FN, false negative; FP, false positive; NA, not applicable; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

[a]Extracted age that fully coincided with that in the clinical data.

Overall complications ≥grade IIIa occurred in 111 patients. and correct extraction was achieved in 93 patients (sensitivity 0.838; 95% CI, 0.756–0.901). There were 15 negative cases incorrectly extracted as positive from the DPC data (specificity 0.991; 95% CI, 0.984–0.995). Overall complications ≥grade IIIb occurred in 47 patients; all were extracted correctly (sensitivity 1.000). There were nine negative cases incorrectly extracted as positive from the DPC data (specificity 0.995; 95% CI, 0.990–0.997). The results for each hospital and operative method are shown in Tables S6 and S7.

3.4 | Accuracy of the extracted postoperative course

Table 5 shows the results of the validation analyses of the postoperative course. All extracted data on length of postoperative hospital stay coincided with the clinical data by chart review. In practice, 47 patients were transferred to another hospital postoperatively, and this outcome was correctly extracted in 43 (sensitivity 0.915; 95% CI, 0.796–0.976). Two patients were incorrectly regarded to have this outcome (specificity 0.999; 95% CI, 0.996–1.000). Unplanned readmissions within POD 30 occurred in 55 patients in practice and was correctly extracted in 51 patients (sensitivity 0.927; 95% CI, 0.824–0.980). No patient was misidentified as experiencing readmission (specificity 1.000).

4 | DISCUSSION

Most validation studies using administrative data were internally performed at a single center.[24,25] On the other hand, in the present study algorithms for extracting GI cancer surgery cases and postoperative complications requiring interventions were developed using data from one hospital and were externally validated using data from two other hospitals. The extraction of GI cancer surgery cases had a high PPV, and the specificity of extraction of postoperative complications was almost 1.000. In studies that aim to prove hypotheses, a high PPV is required to obtain the target population, and a high specificity is required to obtain the correct relative risk.[18] Our results indicated that patients who underwent GI cancer surgery and postoperative outcomes requiring interventions can be accurately identified from the administrative claims database. As far as we know, this was the first multicenter study to validate the accuracy of extraction of postoperative complications, according to the severity of complications, from claims data.

The importance of validation studies has been emphasized to avoid misclassification bias of results.[2] Nevertheless, a past study reported that 88% of studies that used an administrative claims database employed unvalidated algorithms to extract cohorts or outcomes.[7] This present multicenter external validation study may contribute to future research on hospital quality improvement. For instance, feedback on the annual trends of overall complications ≥grade IIIb following GI cancer surgery in a hospital may be done with

TABLE 4 Accuracy of the extracted postoperative complications ($n = 1694$)

	TP	TN	FN	FP	Prevalence (%)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Invasive procedure[a]	71	1594	18	11	5.3	0.798 (0.699–0.876)	0.993 (0.988–0.997)	0.866 (0.773–0.931)	0.989 (0.982–0.993)
Reoperation	35	1655	0	4	2.1	1.000 (0.900–1.000)	0.998 (0.994–0.999)	0.897 (0.758–0.971)	1.000 (0.998–1.000)
Mechanical ventilation									
Algorithm A	21	1670	0	3	1.2	1.000 (0.839–1.000)	0.998 (0.995–1.000)	0.875 (0.676–0.973)	1.000 (0.998–1.000)
Algorithm B	21	1668	0	5	1.2	1.000 (0.839–1.000)	0.997 (0.993–0.999)	0.808 (0.606–0.934)	1.000 (0.998–1.000)
Hemodialysis	1	1693	0	0	0.1	1.000 (0.025–1.000)	1.000 (0.998–1.000)	1.000 (0.025–1.000)	1.000 (0.998–1.000)
ICU management									
Algorithm A	23	1666	2	3	1.5	0.920 (0.740–0.990)	0.998 (0.995–1.000)	0.885 (0.698–0.976)	0.999 (0.996–1.000)
Algorithm B	21	1668	4	1	1.5	0.840 (0.639–0.955)	0.999 (0.997–1.000)	0.955 (0.772–0.999)	0.998 (0.994–0.999)
In-hospital mortality	5	1689	0	0	0.3	1.000 (0.478–1.000)	1.000 (0.998–1.000)	1.000 (0.478–1.000)	1.000 (0.998–1.000)
Overall complication ≥ grade IIIa	93	1568	18	15	6.6	0.838 (0.756–0.901)	0.991 (0.984–0.995)	0.861 (0.781–0.920)	0.989 (0.982–0.993)
Overall complication ≥ grade IIIb	47	1638	0	9	2.8	1.000 (0.925–1.000)	0.995 (0.990–0.997)	0.839 (0.717–0.924)	1.000 (0.998–1.000)

Note: TP was defined as a case that was identified in both DPC and chart review. TN was defined as a case that was not identified in both DPC and chart review. FN was defined as a case that was identified by chart review but not in the DPC. FP was defined as a case that was identified in the DPC but not by chart review.Prevalence = (TP + FN)/(TP + TN + FN + FP); sensitivity = TP/(TP + FN); specificity = TN/(FP + TN); PPV = TP/(TP + FP); NPV = TN/(FN + TN).

Abbreviations: CI, confidence interval; FN, false negative; FP, false positive; ICU, intensive care unit; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

[a]Surgical, endoscopic, or radiological intervention not under general anesthesia.

benchmark data. In addition, large-scale comparisons between interventions (e.g. robotic surgery vs. thoracoscopic/laparoscopic surgery) can be performed using DPC data in nationwide hospitals in Japan.

In terms of identification of patients who underwent GI cancer surgery, we combined both DPC diagnosis and surgical procedures, which led to a highly accurate extraction. This was consistent with previous findings that utilization of a combination of diagnosis and surgery allowed a more accurate extraction of the target cohort.[15] Unlike most administrative databases, the DPC database has several diagnoses that represent hospitalizations; therefore, the combination of those diagnoses can be used.[14] This might also contribute to a higher extraction accuracy.

Information on outcomes, such as postoperative complications, has been considered difficult to collect using claims data.[26] Therefore, we did not choose a study design to validate the diagnostic codes themselves, indicating complications. However, we considered that we could extract postoperative complications requiring invasive treatments, such as reoperation and intensive care, because the administrative data routinely hold all procedure-related information with the aim of submitting claims for payment. Although we used a Japanese claims database, we believe that this fundamental principle could be applied to other countries and fields. Nevertheless, similar validation studies are warranted.

In terms of invasive procedure not under general anesthesia, we expected a lower accuracy, compared with that of the other outcomes, because of miscoding or noncoding of the procedures. Indeed, previous studies have shown that minor procedures, such as those generally performed at hospital wards or radiology departments, tended to be coded inaccurately, whereas major procedures, such as surgery, tended to be coded accurately in the administrative data.[27,28] In the present study, although the sensitivity for this outcome was lower (0.798), compared with that of the other outcomes, the specificity was high (0.993). This tendency of decreased sensitivity was observed among all operative methods: esophagectomy, gastrectomy, and colorectomy. It might be attributed to the fact that coding professionals were less likely to notice minor procedures performed in the ward, such as percutaneous drainage with nelaton catheters. We believed that the use of this outcome in future studies may be acceptable if its limitations are fully understood.

With regard to reoperation, we extracted cases of general anesthesia use postoperatively. The accuracy was considerably favorable by preliminarily excluding scheduled surgeries, such as stoma closure, or second operation of two-stage esophagectomy, which could be performed during hospitalization rather than as treatment of complications.

TABLE 5 Accuracy of the extracted postoperative course

	TP	TN	FN	FP	Prevalence (%)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Length of postoperative hospital stay[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA
Transfer to a different hospital	43	1645	4	2	2.8	0.915 (0.796–0.976)	0.999 (0.996–1.000)	0.956 (0.849–0.995)	0.998 (0.994–0.999)
30-day unplanned hospital readmission	51	1639	4	0	3.2	0.927 (0.824–0.980)	1.000 (0.998–1.000)	1.000 (0.930–1.000)	0.998 (0.994–0.999)

Note: TP was defined as a case that was identified in both DPC and chart review. TN was defined as a case that was not identified in both DPC and chart review. FN was defined as a case that was identified by chart review but not in the DPC. FP was defined as a case that was identified in the DPC but not by chart review.Prevalence = (TP + FN)/(TP + TN + FN + FP); sensitivity = TP/(TP + FN); specificity = TN/(FP + TN); PPV = TP/(TP + FP); NPV = TN/(FN + TN).

Abbreviations: CI, confidence interval; FN, false negative; FP, false positive; NA, not applicable; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

[a]Extracted length of postoperative hospital stay fully coincided with the clinical data.

For mechanical ventilation, two patterns of validation were performed, because the duration of postoperative mechanical ventilation might vary among facilities. As a result, the accuracy was better for algorithm A than for algorithm B. Nevertheless, algorithm B may be valuable, as there had been several past reports that set 48 h as the cutoff value for the postoperative ventilation period when extracting from administrative data.[29,30] In a future project or study, either of the algorithms can be selected, depending on which will suit the purpose. In both algorithms, NPPV was included as mechanical ventilation, because the Japanese coding guidelines allow NPPV to be registered as mechanical ventilation if certain conditions that define when acute respiratory failure are met. Notably, interpretation of the results needs caution, because NPPV is usually not defined as a Clavien–Dindo grade IV equivalent.[30,31]

The accuracy of two algorithms for extraction of postoperative ICU management was examined, because the length of ICU stay after a scheduled admission varied among hospitals.[32] Although the balance of sensitivity and specificity for ICU management seemed to be better with algorithm A than with algorithm B, the use of both algorithms might be acceptable in future projects and research.

With regard to hemodialysis or in-hospital mortality, no inaccurate extraction was found in the present cohort. In-hospital mortality had been previously used as an outcome in many reports using DPC data.[4,5] It will serve as a basis for ensuring the quality of prior and future researches.

The present study had some limitations. First, the inclusion of relatively large-scale affiliated hospitals in this study may limit the generalizability of our findings. Second, coding can be changed or added in the future, because revision of medical fees is conducted principally every 2y. Third, the information validated in this study is only part of the DPC data. Although some studies report a high validity of diagnoses in the DPC data, further validation studies are warranted to avoid information bias regarding patients' comorbidities and surgical characteristics.[33,34]

In conclusion, our developed algorithms to identify patients who underwent GI cancer surgery and to extract postoperative complications requiring interventions from the real-world data had high accuracy. This external validation study could provide an important basis for ensuring the quality of future studies that will use real-world data in the surgical field.

AUTHOR CONTRIBUTIONS

HK, TN, and SK contributed to the study conception and design. HK, HH, SI, HH, and RM collected the data. HK, TN, SK, KH, YI, KO, and YM analyzed and interpreted the data. HK and TN drafted the article. All authors contributed to the critical revision of the article and approved the final article.

ACKNOWLEDGMENTS

The authors thank the medical staff and the claims data managers in the hospitals participating in this project. The authors thank Enago (www.enago.jp) for the English language review.

FUNDING INFORMATION

No financial support was received for this study.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest for this article.

ETHICAL STATEMENT

Approval of the research protocol: The protocol for this research project was approved by a suitably constituted Ethics Committee of the institution and conformed to the provisions of the Declaration of Helsinki; the Committee of Kyoto University, Approval No. R2435; the committee of Kyoto Medical Center, No. 20-015 and the committee of Osaka Red Cross Hospital, No. J-0134.

Informed consent: Opt-out consent was obtained because of the retrospective nature of the study.

Registry and the registration no. of the study/trial: N/A.

Animal studies: N/A.

ORCID

Hiromitsu Kinoshita https://orcid.org/0000-0002-1647-5137
Tatsuto Nishigori https://orcid.org/0000-0002-4969-5590
Susumu Kunisawa https://orcid.org/0000-0002-3657-8686
Koya Hida https://orcid.org/0000-0001-7210-7075
Kazutaka Obama https://orcid.org/0000-0003-2924-6701

REFERENCES

1. Kreis K, Neubauer S, Klora M, Lange A, Zeidler J. Status and perspectives of claims data analyses in Germany-a systematic review. Health Policy. 2016;120(2):213–26.
2. Koram N, Delgado M, Stark JH, Setoguchi S, de Luise C. Validation studies of claims data in the Asia-Pacific region: a comprehensive review. Pharmacoepidemiol Drug Saf. 2019;28(2):156–70.
3. Enomoto LM, Hollenbeak CS, Bhayani NH, Dillon PW, Gusani NJ. Measuring surgical quality: a national clinical registry versus administrative claims data. J Gastrointest Surg. 2014;18(8):1416–22.
4. Yasunaga H, Horiguchi H, Kuwabara K, Matsuda S, Fushimi K, Hashimoto H, et al. Outcomes after laparoscopic or open distal gastrectomy for early-stage gastric cancer: a propensity-matched analysis. Ann Surg. 2013;257(4):640–6.
5. Sakamoto T, Fujiogi M, Matsui H, Fushimi K, Yasunaga H. Comparing perioperative mortality and morbidity of minimally invasive esophagectomy versus open esophagectomy for esophageal cancer: a Nationwide retrospective analysis. Ann Surg. 2021;274(2):324–30.
6. Yoneyama T, Nakashima M, Takeuchi M, Kawakami K. Comparison of laparoscopic and open inguinal hernia repair in adults: a retrospective cohort study using a medical claims database. Asian J Endosc Surg. 2022;15(3):513–23.
7. van Walraven C, Bennett C, Forster AJ. Administrative database research infrequently used validated diagnostic or procedural codes. J Clin Epidemiol. 2011;64(10):1054–9.
8. van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. J Clin Epidemiol. 2012;65(2):126–31.
9. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg. 2004;240(2):205–13.
10. Katayama H, Kurokawa Y, Nakamura K, Ito H, Kanemitsu Y, Masuda N, et al. Extended Clavien-Dindo classification of surgical complications: Japan clinical oncology group postoperative complications criteria. Surg Today. 2016;46(6):668–85.
11. Hayashida K, Murakami G, Matsuda S, Fushimi K. History and profile of diagnosis procedure combination (DPC): development of a real data collection system for acute inpatient care in Japan. J Epidemiol. 2021;31(1):1–11.
12. Ministry of Health, Labor and Welfare. Survey on Impact Assessment of DPC Introduction: Summary of Statistics (in Japanese) [cited 2 March 2023]. Available from: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000049343.html
13. Ministry of Health, Labor and Welfare. Materials on the Additional Data Submission (in Japanese) [cited 2 March 2023]. Available from: https://www.mhlw.go.jp/content/12404000/000938427.pdf
14. Yasunaga H, Matsui H, Horiguchi H, Fushimi K, Matsuda S. Clinical epidemiology and health services research using the diagnosis procedure combination database in Japan. Asian Pac J Dis Manag. 2015;7(1–2):19–24.
15. Goldsbury D, Weber M, Yap S, Banks E, O'Connell DL, Canfell K. Identifying incident colorectal and lung cancer cases in health service utilisation databases in Australia: a validation study. BMC Med Inform Decis Mak. 2017;17(1):23.
16. Benchimol EI, Manuel DG, To T, Griffiths AM, Rabeneck L, Guttmann A. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. J Clin Epidemiol. 2011;64(8):821–9.
17. Butt DA, Tu K, Young J, Green D, Wang M, Ivers N, et al. A validation study of administrative data algorithms to identify patients with parkinsonism with prevalence and incidence trends. Neuroepidemiology. 2014;43(1):28–37.
18. Chubak J, Pocobelli G, Weiss NS. Tradeoffs between accuracy measures for electronic health care data algorithms. J Clin Epidemiol. 2012;65(3):343–9.e2.
19. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inform. 2014;48:193–204.
20. Matsubara N, Miyata H, Gotoh M, Tomita N, Baba H, Kimura W, et al. Mortality after common rectal surgery in Japan: a study on low anterior resection from a newly established nationwide large-scale clinical database. Dis Colon Rectum. 2014;57(9):1075–81.
21. Kobayashi H, Miyata H, Gotoh M, Baba H, Kimura W, Kitagawa Y, et al. Risk model for right hemicolectomy based on 19,070 Japanese patients in the National Clinical Database. J Gastroenterol. 2014;49(6):1047–55.
22. Tsunoda S, Obama K, Hisamori S, Nishigori T, Okamura R, Maekawa H, et al. Lower incidence of postoperative pulmonary complications following robot-assisted minimally invasive esophagectomy for esophageal cancer: propensity score-matched comparison to conventional minimally invasive esophagectomy. Ann Surg Oncol. 2021;28(2):639–47.
23. Tokunaga M, Kurokawa Y, Machida R, Sato Y, Takiguchi S, Doki Y, et al. Impact of postoperative complications on survival outcomes in patients with gastric cancer: exploratory analysis of a randomized controlled JCOG1001 trial. Gastric Cancer. 2021;24(1):214–23.
24. Iwamoto M, Higashi T, Miura H, Kawaguchi T, Tanaka S, Yamashita I, et al. Accuracy of using diagnosis procedure combination administrative claims data for estimating the amount of opioid consumption among cancer patients in Japan. Jpn J Clin Oncol. 2015;45(11):1036–41.
25. Kanehara R, Goto A, Goto M, Takahashi T, Iwasaki M, Noda M, et al. Validation study on diabetes definitions using Japanese diagnosis procedure combination data among hospitalized patients. J Epidemiol. 2021;33(4):165–169.
26. Lawson EH, Zingmond DS, Hall BL, Louie R, Brook RH, Ko CY. Comparison between clinical registry and medicare claims data on the classification of hospital quality of surgical care. Ann Surg. 2015;261(2):290–6.

27. Quan H, Parsons GA, Ghali WA. Validity of procedure codes in international classification of diseases, 9th revision, clinical modification administrative data. Med Care. 2004;42(8):801–9.
28. Lloyd SS, Rissing JP. Physician and coding errors in patient records. JAMA. 1985;254(10):1330–6.
29. Webb S, Rubinfeld I, Velanovich V, Horst HM, Reickert C. Using National Surgical Quality Improvement Program (NSQIP) data for risk adjustment to compare Clavien 4 and 5 complications in open and laparoscopic colectomy. Surg Endosc. 2012;26(3):732–7.
30. Abt NB, Richmon JD, Koch WM, Eisele DW, Agrawal N. Assessment of the predictive value of the modified frailty index for Clavien-Dindo grade iv critical care complications in major head and neck cancer operations. JAMA Otolaryngol Head Neck Surg. 2016;142(7):658–64.
31. Ishii M, Mizuguchi T, Harada K, Ota S, Meguro M, Ueki T, et al. Comprehensive review of post-liver resection surgical complications and a new universal classification and grading system. World J Hepatol. 2014;6(10):745–51.
32. Voeten DM, van der Werf LR, Gisbertz SS, Ruurda JP, van Berge Henegouwen MI, van Hillegersberg R. Postoperative intensive care unit stay after minimally invasive esophagectomy shows large hospital variation. Results from the Dutch upper gastrointestinal cancer audit. Eur J Surg Oncol. 2021;47(8):1961–8.
33. Yamana H, Moriwaki M, Horiguchi H, Kodan M, Fushimi K, Yasunaga H. Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. J Epidemiol. 2017;27(10):476–82.
34. Shigemi D, Morishima T, Yamana H, Yasunaga H, Miyashiro I. Validity of initial cancer diagnoses in the diagnosis procedure combination data in Japan. Cancer Epidemiol. 2021;74:102016.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kinoshita H, Nishigori T, Kunisawa S, Hida K, Hosogi H, Inamoto S, et al. Identification of complications requiring interventions after gastrointestinal cancer surgery from real-world data: An external validation study. Ann Gastroenterol Surg. 2023;7:1032–1041. https://doi.org/10.1002/ags3.12704