BMJ Open

Accuracy of Algorithms to Identify Patients With a Diagnosis of Major Cancers and Cancer Related Adverse Events Using Japanese Health Administrative Data

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-055459
Article Type:	Original research
Date Submitted by the Author:	15-Jul-2021
Complete List of Authors:	Fujiwara, Takashi; Kurashiki Central Hospital, Department of Management, Clinical Research Center; Kurashiki Central Hospital, Department of Otolaryngology/Head and Neck Surgery Kanemitsu, Takashi; Chugai Pharmaceutical Co Ltd, Medical Affairs Division Tajima, Kosei ; Chugai Pharmaceutical Co Ltd, Clinical Development Division Yuri, Akinori; Chugai Pharmaceutical Co Ltd, Drug Safety Division Iwasaku, Masahiro; Kurashiki Central Hospital, Department of Management, Clinical Research Center Okumura, Yasuyuki; Real world Data Co., Ltd. Tokumasu, Hironobu; Kurashiki Central Hospital, Department of Management, Clinical Research Center; Real World Data Co., Ltd.
Keywords:	Adult oncology < ONCOLOGY, Breast tumours < ONCOLOGY, Gynaecological oncology < GYNAECOLOGY, Respiratory tract tumours < ONCOLOGY, Urological tumours < ONCOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title page

Title

Accuracy of Algorithms to Identify Patients With a Diagnosis of Major Cancers and Cancer-Related Adverse Events Using Japanese Health Administrative Data

Authors

Takashi Fujiwara[1,2], Takashi Kanemitsu[3], Kosei Tajima[4], Akinori Yuri[5], Masahiro Iwasaku[1], Yasuyuki Okumura[6], Hironobu Tokumasu[1,6]

Affiliations

1. Department of Management, Clinical Research Center, Kurashiki Central Hospital, Okayama, Japan

2. Department of Otolaryngology/Head and Neck Surgery, Kurashiki Central Hospital, Okayama, Japan

3. Medical Affairs Division, Chugai Pharmaceutical Co., Ltd., Tokyo, Japan

4. Clinical Development Division, Chugai Pharmaceutical Co., Ltd., Tokyo, Japan

5. Drug Safety Division, Chugai Pharmaceutical Co., Ltd., Tokyo, Japan

6. Real World Data Co., Ltd., Kyoto, Japan

Corresponding author

Name: Hironobu Tokumasu

Address: Department of Management, Clinical Research Center, Kurashiki Central Hospital, 1-1-1 Miwa, Kurashiki, Okayama, 710-8602 Japan

Phone number: +81-86-422-0210

Fax number: +81- 86-421-3424

Email: tokumasu@rwdata.co.jp

Word count: 3,871

Abstract

Objectives: Validation studies in oncology are limited in Japan. This study was conducted to evaluate the accuracy of diagnosis and adverse event (AE) definitions for specific cancers in a Japanese health administrative real-world database (RWD).

Design and setting: Retrospective observational validation study to assess the diagnostic accuracy of electrical medical records (EMR) and claim coding regarding oncology diagnosis and AEs based on medical record review in the RWD.

Participants: The validation cohort included patients with lung (n=2,257), breast (n=1,121), colorectal (n=1,773), ovarian (n=216), and bladder (n=575) cancer who visited the hospital between January 2014 and December 2018, and those with prostate cancer (n=3,491) visiting between January 2009 and December 2018, who were identified using EMRs.

Outcomes: Key outcomes included primary diagnosis, deaths, and AEs.

Results: Data on International Classification of Diseases, 10th revision (ICD-10)–definitive diagnosis and death could be extracted with high accuracy. The positive predictive value (PPV; 95% confidence interval [CI]) for primary diagnosis was high (lung, 81.0 [74.9–86.2]; breast, 74.0 [67.3–79.9]; colorectal, 80.5 [74.3–85.8]; ovarian, 49.5 [39.3–59.7]; bladder, 42.0 [32.2–52.3]; prostate, 79.0 [69.7–86.5]). Sensitivity (95% CI) for death was high (lung, 97.0 [84.2–99.9]; breast, 100.0 [1.3–100.0]; colorectal, 100.0 [28.4–100.0]; ovarian, 100.0 [35.9–100.0]; bladder, 100.0 [9.4–100.0]; prostate, 75.0 [19.4–99.4]). Overall, PPV tended to be low, with the definition based on ICD-10 alone for AEs.

Conclusion: EMR data were deemed appropriate to comprehensively identify patients with specific cancers or deceased patients using RWD in Japan.

Trial registration: University hospital Medical Information Network (UMIN) Clinical Trials Registry; UMIN000039345

Strengths and limitations of this study

- To our knowledge, this is the first study in oncology in Japan that validates disease names and adverse event definitions in a health administrative real-world database (RWD) using chart review based on electronic medical records data from a hospital as the reference standard.
- Validation was performed at a single facility; therefore, there is a possibility of selection bias.
- Study results are limited by the inherent issues related to the use of an RWD, which primarily stores medical information for the purpose of insurance claims.
- The diagnosis and adverse event definitions used in this study may not be the most suitable; thus, there is an opportunity to further deepen these definitions.
- Study methods for the consolidation of true positives for events with low incidence need to be further investigated as it was challenging to investigate outcomes with extremely low incidence.

Keywords

database, electronic medical record, health administrative, real-world database, validation study

INTRODUCTION

In recent years, evidence from routine clinical practice using data from real-world databases (RWDs) has increasingly gained importance in decision-making in healthcare, research, and drug development.[1] In addition, RWD studies can help generate evidence for advancement in precision medicine and facilitation of targeted and efficient patient care.[2] In line with this trend, evidence related to several aspects, such as health technology, expenditure forecasting, survival outcomes, time to therapy, and treatment efficacy, are increasingly being collected from RWD studies in oncology.[3-6]

However, it is important to validate case-identification algorithms to evaluate the accuracy of information sourced from RWDs, which is usually collected for purposes other than research.[7] To this end, several studies have been conducted outside of Japan to evaluate the accuracy of algorithms based on health administrative data in identifying cancer diagnoses or other outcomes using databases, such as registries, population-based cohorts, chart reviews, and electronic medical records (EMRs) as reference standards.[8-17]

The implementation of the revised ordinance of Good Postmarketing Study Practice by the Pharmaceuticals and Medical Devices Agency (PMDA) in 2018 suggests that the importance of using RWDs in post-marketing surveillance to investigate the safety and efficacy of pharmaceutical products is being recognized in Japan as well.[18] To encourage validation studies, the PMDA of Japan and Japan Society for Pharmacoepidemiology established a basic concept for conducting validation studies to verify diagnosis codes and other outcome definitions in Japanese RWDs.[19, 20] However, among the validation studies conducted to date,[21-31] to our knowledge, only one claims-based study reported on outcomes in cancer, more specifically breast cancer; a cancer registry was used as the reference standard in this study.[32] Thus, there is a need to perform validation studies on a wider range of cancer types in Japan using a reliable database as a reference standard. This

study was conducted for validation of diagnosis and adverse event (AE) definitions for specific cancers in a Japanese RWD using a chart review by EMR.

PATIENTS AND METHODS

Study design

This was a validation study of diagnosis and AE definitions in the health administrative RWD of the Health, Clinic, and Education Information Evaluation Institute (HCEI) conducted by chart review from Kurashiki Central Hospital, Japan, as the reference standard.

Data collection

Data were collected retrospectively from EMRs at the Kurashiki Central Hospital, Japan (**Figure 1**), which were the primary data source. All possible cases that met the diagnosis and AE definitions and cases other than all possible cases were identified using International Classification of Diseases, 10th revision (ICD-10) codes (**Figures S1–S6**) from the EMRs. Further, these cohorts were randomly sampled to verify the diagnoses and related events. EMRs were manually reviewed to verify the diagnosis of all possible cases. This verified data set was anonymized and sent to Real World Data Co. Ltd., the vendor for HCEI. The verified data set was linked deterministically to claims data and EMRs originally derived from the hospital.

Chart review based on EMR

A chart review for all possible cases was conducted by medical professionals, including medical doctors involved in the management of cancer patients and four clinical research coordinators (CRCs) at the Kurashiki Central Hospital, Japan. At least two CRCs conducted chart reviews independently. Any disagreements were resolved by discussion between the two CRCs or by a medical doctor if the disagreement was not resolved following a discussion.

HCEI database

HCEI is an integrated RWD initiated in Japan and supported by Real World Data Co., Ltd. (Kyoto).[33] As of August 2020, HCEI was collecting information from approximately 20 million patients from 190 medical institutions in Japan, including Kurashiki Central Hospital. The HCEI database covers 1.2% of the overall Japanese population and includes data from 1.3 million outpatients and 0.21 million inpatients in 2019.[33] Medical information is extracted from EMRs, claims, and Diagnosis Procedure Combination (DPC) in the HCEI database. Patient-level data from DPC, EMRs, and claims are integrated in advance at the hospital, anonymized, linked to a unique code, and standardized (**Figure 1**). The linked data are then provided to HCEI for storage on their server. Information on procedures (such as surgery) is obtained from claims, while information on laboratory tests and treatments is obtained from EMRs. Diagnosis data are obtained from both claims and EMRs. According to HCEI's security policy, personal identifiable information (such as date of birth) is not collected during data extraction. Master lists are constructed based on the national standards of the Ministry of Health, Labour and Welfare (MHLW) of Japan.[34]

Ethics Approval

This study was approved by the Research Institute of Healthcare Data Science (https://rihds.org/ethic/) (RI2019010) and the institutional ethics committee of Kurashiki Central Hospita (KCH3301) l, and conducted under the tenets of the Declaration of Helsinki, Act on the Protection of Personal Information,[35] and Ethical Guidelines for Medical and Health Research Involving Human Subjects.[36] It was conducted under a joint research agreement between Kurashiki Central Hospital, Chugai Pharmaceutical Co., Ltd., and HCEI, and is registered at the UMIN Clinical Trials Registry (UMIN000039345). Target patients at

Kurashiki Central Hospital had the option, on the hospital's website, to refuse disclosure of their information.

Patient and public involvement in research

Patients or the public were not involved in the design or conduct, reporting or dissemination plans of our research.

Patient selection

Patients with lung, breast, colorectal, ovarian, and bladder cancer who visited Kurashiki Central Hospital between January 2014 and December 2018 (**Figures S1–S5**), and those with prostate cancer (**Figure S6**) who visited the hospital between January 2009 and December 2018 were eligible for inclusion in the study. Patients participating in clinical trials during the data extraction periods and those who were assigned the respective ICD-10 code for lung, colorectal, breast, ovarian, and bladder cancer from January 1, 2014, to January 31, 2014, and from November 1, 2018, to December 31, 2018, and that for prostate cancer from January 1, 2009, to January 31, 2009, and from November 1, 2018, to December 31, 2018, were excluded from the study. Patients diagnosed during these periods were excluded to avoid bias due to the time lag between suspected diagnosis by medical examination and confirmation of diagnosis by biopsy, when the outcome definition was potentially met.

The cohort entry date was the date when the respective cancer was diagnosed—January 2014 for lung, breast, colorectal, ovarian, and bladder cancer and January 2009 for prostate cancer—and the end date was December 31, 2018. To avoid selection of cases diagnosed before the cohort entry date, patients who were assigned the respective ICD-10 code for lung, colorectal, breast, ovarian, and bladder cancer before December 31, 2013, and that for prostate cancer before December 31, 2008, were excluded.

Eligible patients were stratified by random sampling as all possible and not possible

cases. All possible cases included patients who met the ICD-10 code for the respective support during the specified data extraction period. Patients who were never assigned an ICD-10 code for the respective cancer; those with lung, colorectal, breast, ovarian, and bladder cancer who visited the hospital between January 1, 2014, and December 31, 2018; and those with prostate cancer between January 1, 2009, and December 31, 2018, were stratified as not possible cases. Overall, 200 cases each with lung, breast, or colorectal cancer and 100 cases each with ovarian, bladder, or prostate cancer were targeted and randomly selected from all possible cases for the EMR review, and not possible cases were also randomly selected using the same proportions.

Outcomes and assessment of accuracy

Outcomes for validation included primary diagnosis, performance status (PS) ≥2,[37] first/second/third recurrence or exacerbation, death, and AEs, particularly immune-related AEs (irAEs), associated with new diagnoses for patients with lung, breast, colorectal, ovarian, bladder, and prostate cancer. AEs included interstitial pneumonia, liver dysfunction, colitis/diarrhea, type 1 diabetes mellitus (T1DM), encephalitis/meningitis, nerve disorders (excluding paresthesia), myasthenia gravis, Guillain-Barré syndrome, skin disorder, rhabdomyolysis, myocarditis, perforation of digestive tract/fistula, hypoadrenocorticism, and febrile neutropenia.

Outcomes were defined by separate algorithms (**Tables S1 and S2**) for each cancer type using one variable or a combination of two or more variables, such as diagnoses, treatments, procedures, and laboratory test results. Lung cancer was further classified as primary, non-small cell, and small cell.

Statistical analysis

The target sample size for random sampling was determined based on the feasibility of the

chart review, assuming that the 95% confidence interval (CI) for positive predictive value (PPV) and sensitivity can be estimated with an accuracy of maximum ±10% if ≥100 patients met the definition of primary diagnosis and ≥100 were true positives.[38]

In the data set submitted by HCEI, accuracy for each cancer type was evaluated using sensitivity, specificity, PPV, and negative predictive value (NPV) for primary diagnosis, first recurrence/exacerbation, and death. Other outcomes were evaluated using only PPV to determine if the cases were true for those meeting the outcome definition. AEs were validated in patients with true primary cancer who had received chemotherapy. PPV was calculated only after confirming whether the outcome occurred within (before or after) 30 days of the patient meeting the outcome definition.

All possible cases refer to the population that is assumed to include all true patients,[19, 39-41] and included patients who met the ICD-10 code for the respective cancer in the EMRs during the specified data extraction period. True positives were defined as patients in whom the outcomes occurred based on HCEI information and EMR review. In addition, patients were randomly selected from cases other than all possible cases at the same extraction rate as that for "all possible cases" to calculate the specificity and NPV for primary diagnosis, first recurrence/exacerbation, and death. The data extraction period for different cancer types was estimated based on the national survival rate survey of 2019 conducted by the National Cancer Center Council,[42] in which the survival period was 10 years for prostate cancer and 5 years for other cancer types. Likewise, a longer data extraction period was considered for prostate cancer to allow for the collection of true positives.

The frequency and 95% CIs were calculated for sensitivity, specificity, PPV, and NPV. The degree of agreement between two chart reviewers was evaluated using the kappa coefficient. Extrapolability of the Kurashiki Central Hospital database to that of other hospitals in HCEI database was assessed by comparing the distribution of patient

characteristics. Matching was performed through deterministic linkage and statistical analyses were conducted using R-4.0.2 software.

RESULTS

Patient disposition

Of the 256,418 patients who received medical treatment from 2014 to 2018, 2,257 with lung cancer (**Figure S1**), 1,121 with breast cancer (**Figure S2**), 1,773 with colorectal cancer (**Figure S3**), 216 with ovarian cancer (**Figure S4**), and 575 with bladder cancer (**Figure S5**) were included as all possible cases (**Table 1**). From 2009 to 2018, 3,491 patients with prostate cancer of 413,631 patients receiving medical treatment (**Figure S6**) were included as all possible cases (**Table 1**).

Table 1. Study cohort

Cancer type	Patients who underwent medical treatment from January 2014 to December 2018,* n	Target patients, n	All possible cases, n	True cases, n
Lung cancer	256,418	252,847	2,257	162
Breast cancer	256,418	253,358	1,121	148
Colorectal cancer	256,418	252,733	1,773	161
Ovarian cancer	256,418	254,995	216	49
Bladder cancer	256,418	254,520	575	42
Prostate cancer	413,631	410,356	3,491	79

*Period: January 2009 to December 2018 for prostate cancer

Lung cancer

The kappa value in chart reviews for diagnosis definitions was 0.982 (95% CI: 0.947–1.017) for primary lung cancer, 0.979 (95% CI: 0.950–1.008) for non-small cell lung cancer (NSCLC), 1.00 for small cell lung cancer (SCLC), and 0.982 (95% CI: 0.947–1.017) for death. There were 30 false negatives and 132 true positives for A1 using DPC diagnosis (**Table 2**). Sensitivity was 100% with A2 using related definitive diagnosis (**Table 2**). Although specificity, PPV, and NPV for NSCLC were high for B1 and B2 using cancer-related diagnosis codes, sensitivity was low (38.3%; **Table S3**). Accuracy was high for all statistical parameters for SCLC (**Table 2**). Data on death could be extracted with high accuracy using EMR definitions (E1; **Table 3**).

Table 2. Diagnosis definitions with high* accuracy

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV % (95% CI)
Lung cancer								
Primary lung cancer								
A1	132	7	22,237	30	81.5 (74.6–87.1)	100.0 (99.9–100.0)	95.0 (89.9–98.0)	99.9 (99.8–99.9)
A2	162	38	22,206	0	100.0 (96.6–100.0)	99.8 (99.8–99.9)	81.0 (74.9–86.2)	100.0 (100.0–100.0)
A4	128	7	22,237	34	79.0 (71.8–85.0)	100.0 (99.9–100.0)	94.8 (89.6–97.9)	99.8 (99.8–99.9)
Small cell lung cancer								
C1	10	0	22,395	1	90.9 (58.7–99.8)	100.0 (100.0–100.0)	100.0 (58.7–100.0)	100.0 (100.0–100.0)
Breast cancer								
Primary breast cancer								
α2	148	52	45,002	0	100.0 (96.3–100.0)	99.9 (99.8–99.9)	74.0 (67.3–79.9)	100.0 (100.0–100.0)
Colorectal cancer								
Primary colorectal cancer								
β2	161	39	28,309	0	100.0	99.9	80.5	100.0

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV % (95% CI)
					(96.6–100.0)	(99.8–99.9)	(74.3–85.8)	(100.0–100.0)
Ovarian cancer								
Primary ovarian cancer								
γ1	44	14	11,692	5	89.8 (77.8–96.6)	99.9 (99.8–99.9)	75.9 (62.8–86.1)	100.0 (99.7–100.0)
Bladder cancer								
Primary bladder cancer								
ε1	33	16	44,206	9	78.6 (63.2–89.7)	100.0 (99.9–100.0)	67.3 (52.5–80.1)	100.0 (100.0–100.0)
Prostate cancer								
Primary prostate cancer								
δ2	79	21	11,655	0	100.0 (93.2–100.0)	99.8 (99.7–99.9)	79.0 (69.7–86.5)	100.0 (100.0–100.0)

CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value

*All accuracy values included for a definition are approximately 70% or more.

Table 3. Death definitions with high* accuracy

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Lung cancer								
E1	32	0	40	1	97.0 (84.2–99.9)	100.0 (87.1–100.0)	100.0 (84.2–100.0)	97.6 (87.1–99.9)
E4	32	0	40	1	97.0 (84.2–99.9)	100.0 (87.1–100.0)	100.0 (84.2–100.0)	97.6 (87.1–99.9)
Breast cancer								
E1	1	0	104	0	100.0 (1.3–100.0)	100.0 (94.8–100.0)	100.0 (1.3–100.0)	100.0 (94.8–100.0)
E4	1	0	104	0	100.0 (1.3–100.0)	100.0 (94.8–100.0)	100.0 (1.3–100.0)	100.0 (94.8–100.0)
Colorectal cancer								
E1	4	0	53	0	100.0 (28.4–100.0)	100.0 (90.1–100.0)	100.0 (28.4–100.0)	100.0 (90.1–100.0)
E4	4	0	53	0	100.0 (28.4–100.0)	100.0 (90.1–100.0)	100.0 (28.4–100.0)	100.0 (90.1–100.0)
Ovarian cancer								
E1	5	0	16	0	100.0	100.0	100.0	100.0

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
					(35.9–100.0)	(71.3–100.0)	(35.9-100.0)	(71.3–100.0)
E4	5	0	16	0	100.0 (35.9–100.0)	100.0 (71.3–100.0)	100.0 (35.9-100.0)	100.0 (71.3–100.0)
Bladder cancer								
E1	2	0	8	0	100.0 (9.4–100.0)	100.0 (51.8–100.0)	100.0 (9.4–100.0)	100.0 (51.8–100.0)
E4	2	0	8	0	100.0 (9.4–100.0)	100.0 (51.8–100.0)	100.0 (9.4–100.0)	100.0 (51.8–100.0)
Prostate cancer								
E1	3	0	32	1	75.0 (19.4–99.4)	100.0 (94.2–100.0)	100.0 (19.4–100.0)	97.0 (84.2–99.9)
E4	3	0	32	1	75.0 (19.4–99.4)	100.0 (94.2–100.0)	100.0 (19.4–100.0)	97.0 (84.2–99.9)

CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value

*All accuracy values included for a definition are >70%.

Breast cancer

The kappa value in the chart review for diagnosis definitions was 1.000 and 0.961 (95% CI: 0.917–1.005) for death. The sensitivity was 100% for α2 using EMR diagnosis (**Table 2**) Sensitivity was as low as 62.8% and there were 55 false negatives in α1 using DPC diagnosis (**Table S3**). The accuracy of death definitions for breast cancer was challenging to calculate because outcome events were very few owing to good disease prognosis (**Table S4**).

Colorectal cancer

The kappa value in the chart review for both diagnosis definitions and death was 0.953 (95% CI: 0.900–1.006). There were 39 false positives in β2 (**Table 2**); 15 were diagnosed with colorectal cancer before 2014, two had malignancies that were excluded, and the remaining patients were diagnosed with another cancer on subsequent examination of EMR. Death occurred in four of 57 target patients, and sensitivity and specificity of E1 were 100% each (**Table 3**).

Ovarian cancer

The kappa value in the chart review for diagnosis definitions was 0.920 (95% CI: 0.843–0.997) and 0.940 (95% CI: 0.873–1.007) for death. PPV was higher with γ1 than with γ2 (75.9% vs 49.5%) (**Table S3**). Sensitivity was higher with γ2 than with γ1 (100.0% vs 89.8%) (**Table S3**). Death occurred in five of 21 target patients, and the sensitivity and specificity of E1 were 100% each (**Table 3**).

Bladder cancer

The kappa value in the chart review for diagnosis definitions was 0.898 (95% CI: 0.812–0.985) and 0.878 (95% CI: 0.784–0.973) for death. Sensitivity was 100% in ε2, but PPV was as low as 42.0% (**Table S3**). PPV was higher with ε1 than with ε2 (67.3% vs 42.0%) (**Table S3**). Death occurred in two of 10 target patients, and the sensitivity and

specificity of E1 were 100% each (**Table 3**).

Prostate cancer

The kappa value in the chart review for diagnosis definitions was 0.875 (95% CI: 0.755–0.995) and 0.9045 (95% CI: 0.798–1.011) for death. PPV was 100% in δ1 (**Table S3**), and sensitivity was 100% in δ2 (**Table 2**). Death occurred in four of 36 target patients, and the sensitivity and specificity of E1 were 75% and 100%, respectively (**Table 3**).

Adverse events

The overall PPV for all cancer types was <50%: 47.1% for interstitial pneumonia, 34.6% for liver disorders, 25.5% for colitis/diarrhea, and 13.3% for nerve disorders (excluding paresthesia) by related ICD-10 definitive diagnosis. Although PPV was 100% for encephalitis/meningitis and gastrointestinal perforation by related ICD-10 definitive diagnosis, only one case each was identified as these are rare AEs. For skin disorders, PPV was 76.4% by related ICD-10 definitive diagnosis and 70.4% when treatments were combined in the definition. A combination of related ICD-10 definitive diagnosis and treatments resulted in a PPV of 87.5% for liver disorders. By ICD-10-related definitive diagnosis and intravenous antibiotics use, PPV ranged between 76.9% and 100% for febrile neutropenia. The PPV was 0% for T1DM.

No events of myasthenia gravis, Guillain-Barré syndrome, rhabdomyolysis, adrenal hypofunction, and myocarditis were identified in this analysis.

Other outcomes

Only 1 true positive case was extracted for PS ≥2 for lung cancer using the definition of rehabilitation status. Although the PPV was high, evaluation was difficult. Similarly, the accuracy of the definition of first recurrence/exacerbation was extremely low for all cancer

types owing to very few true positives. Since the accuracy of the second and third recurrence/exacerbation was calculated based on the number of true positives during first recurrence/exacerbation, it could not be evaluated.

Extrapolability of EMR data

Sex and age of all possible cases at the Kurashiki Central hospital and all hospitals were similar (**Table 4**).

Table 4. Demographic and observation period of study population

	All possible cases, n	Male, n (%)	Age (years) at data extraction, mean (SD)	Age (years) at the time of granting ICD-10, mean (SD)	Observation period (days) mean (SD)	Observation period (days) person-years
Lung cancer						
Kurashiki Central Hospital	2,477	1,728 (69.8)	75.0 (9.9)	72.8 (10.2)	801.4 (626.7)	1,985,024
All hospitals	19,861	13,136 (66.1)	74.8 (10.2)	73.5 (10.4)	523.9 (552.4)	10,405,993
Breast cancer						
Kurashiki Central Hospital	1,166	10 (0.9)	67.0 (13.3)	64.1 (13.3)	1,023.6 (650.8)	1,192,400
All hospitals	18,289	131 (0.7)	64.7 (14.1)	62.6 (14.1)	780.5 (618.6)	14,274,791
Colorectal cancer						
Kurashiki Central Hospital	1,684	989 (58.7)	73.6 (11.3)	71.1 (11.6)	930.5 (613.5)	1,566,924
All hospitals	23,501	13,836 (58.9)	74.1 (11.3)	72.1 (11.5)	770.6 (596.2)	18,110,552
Ovarian cancer						
Kurashiki Central Hospital	265	34 (12.8)	66.4 (15.4)	63.9 (15.5)	896.2 (653.5)	237,497

	All possible cases, n	Male, n (%)	Age (years) at data extraction, mean (SD)	Age (years) at the time of granting ICD-10, mean (SD)	Observation period (days) mean (SD)	Observation period (days) person-years
All hospitals	2,592	145 (5.6)	64.1 (14.9)	62.3 (15.1)	6672 (581.1)	1,729,551
Bladder cancer						
Kurashiki Central Hospital	568	446 (78.5)	77.6 (10.0)	75.0 (10.5)	9910 (611.8)	563,042
All hospitals	7,408	5,810 (78.4)	76.9 (10.4)	74.9 (10.6)	7999 (595.8)	5,925,496
Prostate cancer						
Kurashiki Central Hospital	3,131	3,057 (97.6)	76.5 (8.4)	71.9 (8.7)	1,7031 (1,118.3)	5,332,446
All hospitals	32,136	28,690 (89.3)	77.7 (8.9)	74.2 (9.2)	1,3413 (1,041.6)	43,105,126

ICD-10, International Classification of Diseases, 10th revision; SD, standard deviation

DISCUSSION

To our knowledge, this is the first study in oncology in Japan that validates disease names and AE definitions in an RWD, using chart review based on EMR as the gold standard. The accuracy of diagnosis definitions by the ICD-10 code in EMRs was high, with a high sensitivity; therefore, diagnosis definitions by ICD-10 may be generalizable. It was expected that both PPV and NPV would increase by using diagnosis definitions with exact matches, but PPV remained stable and sensitivity decreased. Therefore, definitions including related diagnoses were deemed more appropriate. The PPV of diagnosis definition by DPC was relatively high, but sensitivity tended to be low. Although the diagnosis definition using DPC showed false negatives, it can be used for identifying patients with the respective disease. In the definitions using a definitive diagnosis from claims, PPV tended to decrease, but sensitivity tended to increase. This suggests that it is important to select the outcome definition for use according to the purpose of the study.

Lung cancer, SCLC, and NSCLC could be classified with high accuracy using diagnosis codes. However, there were very few true positives with SCLC. Since the database is used primarily for insurance purposes, precise documentation of a histological classification of lung cancer in EMR was likely not deemed important to be recorded by physicians; therefore, the numbers were low. PPV was high, but sensitivity was low for diagnostic codes for NSCLC; therefore, further studies are required to understand how false negatives can be extracted.

The sensitivity for the EMR definition of breast cancer was 100% and DPC definition was as low as 62.8%. However, specificity was high with both EMR and DPC, and PPV ranged between 74.0% and 83.8%. In a previous study,[32] high sensitivity, specificity, and PPV were observed using definitions obtained by combining diagnostic and procedure codes in a Japanese claims database, suggesting that a combination of codes may result in higher

accuracy.

The accuracy of the evaluation for death was high using the EMR definition for lung cancer. Although sensitivity was high using the EMR definition for other cancers too, further studies are needed in a greater number of cases for confirmation. In cancer types other than lung cancer, which generally have a short prognosis, high sensitivity and PPV were observed for some definitions. However, there were many true negatives because survival was longer than expected and deaths were few, which made evaluation challenging. This could be due to the longer survival of cancer patients at Kurashiki Central Hospital compared with that observed in the national cancer survival rate survey,[42] which was used as a basis for determining the extraction period. Since the survival was long for the hospital database used in this study and fatal events occurred rarely, further investigation is necessary.

Identification of cases with "recurrence/exacerbation" was extremely difficult in all cancer types by definition using items such as diagnoses with "recurrent" as a modifier, pathology-related medical practice code, or relevant surgical history. A previous validation study in breast cancer suggested that the quality of recurrence data may improve by the use of multiple recurrence algorithms in health administrative databases along with selective analysis of medical record data.[17] Another validation study that evaluated breast cancer recurrences achieved high sensitivity and PPV using definitions based on the second round of chemotherapy, diagnostic procedures, treatment, visit to oncologists, patient age, and tumor stage.[15] True positives may be identified if specific therapies are used for the first recurrence/exacerbation, but further investigation is required. Similarly, PS $\geq$ 2, an important variable for cancer, needs further investigation since it was extremely difficult to identify in this study.

For AEs, PPV tended to be low overall with the definition based on ICD-10 alone, suggesting that a combination of definitions based on specific treatment approaches for AEs

could be more appropriate. The definitions of febrile neutropenia and skin disorders had high PPVs and therefore, can be generalized. The validation of T1DM as an AE was challenging as it was difficult to differentiate whether it was an existing comorbidity or developed newly. Moreover, T1DM as a primary diagnosis is rarely found, as the treatment usually targets complications of T1DM. For a few AEs, no true positives were identified, which could be because the outcome definition was developed for irAEs. However, owing to the absence of any reference standard for irAEs in clinical practice, a chart review was instead conducted for AEs in general. For AEs with a low incidence, further studies with a greater number of cases and a more appropriate validation method are required.

Since RWDs contain a large volume of information, it is not realistic to perform validation of multiple outcomes using all cases; instead, representative samples should be used as much as possible. However, such investigations are possible only in a small number of medical facilities. A validation data set, which is a compact version of the database of the concerned medical facility and represents the entire database, should be developed to minimize bias. Further, the definition of the disease and outcomes with low incidence should allow for the collection of as many true positives as possible. An optimal validation methodology should be developed in consideration of the above requirements.

In our study, all possible cases were extracted using the related ICD-10 code from medical information available in the study institution. In order to provide health insurance, the Health Insurance Bureau of the MHLW requires that a suspected diagnosis is changed to a definitive diagnosis as soon as a diagnosis is confirmed.[43] Since the RWD used in this study is a health insurance database, patients with a definitive diagnosis identified by ICD-10 code were deemed as all possible cases. To confirm the robustness of this hypothesis, 100 cases for each cancer type were randomly sampled from cases other than all possible cases to ensure that no patients with a primary diagnosis were included. In future, when conducting a

validation study prior to a pharmacoepidemiology study using information from an RWD, a more efficient method is warranted. In randomized controlled trials (RCTs), the efficacy and safety of treatments are assessed objectively; therefore, assessments are preset. However, in daily clinical practice, treatment decisions are subjective and based on the availability and type of medical resources, capabilities, treatment cost, and patient needs. Therefore, diagnosis and outcome definitions based on efficacy and safety assessments used in RCTs may not be suitable in RWD studies and should be carefully vetted for use in daily clinical practice.

In this study, validation was performed at a single facility; therefore, there is a possibility of selection bias. Further, the results are limited by the inherent issues related to the use of an RWD, which primarily stores medical information for the purpose of insurance claims. Moreover, the diagnosis and AE definitions used in this study may not be the most suitable, and there is an opportunity to further deepen the definitions. For instance, the definition of AE in this study was developed based on treatment-associated irAEs and information on therapeutic agents such as steroids and treatments for allergy; however, definitions based on therapies used for general AE treatment could have been more appropriate. Also, it was challenging to investigate outcomes with extremely low incidence, for example, certain AEs. Therefore, study methods for the consolidation of true positives for events with low incidence need to be investigated.

CONCLUSIONS

The results from our study suggest that patient populations with various cancer types and death can be identified with high sensitivity and predictability by the diagnosis and AE definitions used in this study. DPC data could identify only a limited proportion of patients with cancer, while claims or DPC data could identify only a limited proportion of deceased patients. Since the number of cases was limited in this study, further investigation is required to validate the definitions using DPC and claims data. In view of the current claims process in

Japan, EMR data are deemed appropriate to comprehensively identify patients with cancer or deceased patients for postmarketing surveillance using RWD. Although a high PPV was observed for a few AEs, precision could have been low owing to the low incidence of AEs, and therefore, validation of AEs warrants further investigation.

Acknowledgments

The following persons from Kurashiki Central Hospital Clinical Research Center (Department of Management, Clinical Research Center, Kurashiki Central Hospital, Okayama, Japan) provided additional support: Maki Satomi coordinated at the study site for implementation of protocol procedures and Ryo Ishida, Emi Sato, Mami Yamaguchi, and Yuri Komatsubara contributed to the chart review. Takeshi Kimura of Real World Data Co., Ltd. provided support for statistical analysis and Yusuke Miyoshi of Chugai Pharmaceuticals Co., Ltd. provided administrative support. Akihiro Seki of Chugai Pharmaceuticals supported in developing the outcome definitions.

Editorial support in the form of medical writing, assembling tables, and creating high-resolution images based on the authors' detailed directions, collating author comments, copyediting, fact checking, and referencing was provided by Dr. Deepali Garg, MBBS, PGDHA, of Cactus Life Sciences (part of Cactus Communications) and funded by Chugai Pharmaceutical Co., Ltd.

Funding

This study was funded by Chugai Pharmaceutical Co., Ltd. The funding did not have Award/Grant numbe.

Competing interests

TK, KT, and AY are employees of Chugai Pharmaceutical Co., Ltd. TF reports personal fee

for statistical analysis from Real World Data Co., Ltd. during the conduct of the study; personal fee for collaborative research from Chugai Pharmaceutical Co., Ltd.; and personal fee for statistical analysis from Real World Data Co., Ltd. outside the submitted work. MI has nothing to disclose. YO is an employee of Real World Data Co., Ltd. and reports personal fees from MSD K.K., Otsuka Pharmaceutical, and Kurashiki Central Hospital, outside the submitted work. HT reports personal fees for lecture from AYUMI Pharmaceutical Corporation and Chugai Pharmaceutical Co., Ltd., outside the submitted work and is an employee of Kurashiki Central Hospital and the Director of Real World Data, Co., Ltd.

Author contributions

TF contributed to the study concept and design, and collection, analysis, and interpretation of data. TK, KT, YA and HT contributed to study concept and design, and data interpretation. MI contributed to collection and interpretation of data. YA contributed to analysis and interpretation of data. All authors provided final approval for the version to be published.

Data sharing statement

Data are available upon reasonable request.

Figure legend

Figure 1. Health, Clinic, and Education Information Evaluation Institute/real-world database

EMR: Electronic medical records; HCEI: Health, Clinic, and Education Information Evaluation Institute; KCH: Kurashiki Central Hospital; RWD: real-world database

Figure S1. Patient disposition: Lung cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 199 duplicates; #Study observation periods lasted from January 1, 2014, to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods;

100 patients were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer

Random sampling was performed based on the extraction percentage

Figure S2. Patient disposition: Breast cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 61 duplicates; #Study observation periods lasted from January 1, 2014 to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer

Random sampling was performed based on the extraction percentage

Figure S3. Patient disposition: Colorectal cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 61 duplicates; #Study observation periods lasted from January 1 to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer

Random sampling was performed based on the extraction percentage

Figure S4. Patient disposition: Ovarian cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including three duplicates; #Study observation periods lasted from January 1 to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer

Random sampling was performed based on the extraction percentage

Figure S5. Patient disposition: Bladder cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 25 duplicates; #Study observation periods lasted from January 1, 2014, to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients

were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer

Random sampling was performed based on the extraction percentage

Figure S6. Patient disposition: Prostate cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 44 duplicates; #Study observation periods lasted from January 1, 2009, to January 31, 2009, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients were sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer

Random sampling was performed based on the extraction percentage

References

1 Miksad RA, Abernethy AP. Harnessing the power of real-world evidence (RWE): A checklist to ensure regulatory-grade data quality. *Clin Pharmacol Ther* 2018;103:202–05.

2 Tsai CJ, Riaz N, Gomez SL. Big data in cancer research: Real-world resources for precision oncology to improve cancer care delivery. *Semin Radiat Oncol* 2019;29:306–10.

3 Hess LM, Cui ZL, Mytelka DS, et al. Treatment patterns and survival outcomes for patients receiving second-line treatment for metastatic colorectal cancer in the USA. *Int J Colorectal Dis* 2019;34:581–88.

4 Lin YS, Shen YC, Wu CY, et al. Danshen improves survival of patients with breast cancer and dihydroisotanshinone i induces ferroptosis and apoptosis of breast cancer cells. *Front Pharmacol* 2019;10:1226.

5 Liu JM, Lin CC, Liu KL, et al. Second-line hormonal therapy for the management of metastatic castration-resistant prostate cancer: A real-world data study using a claims database. *Sci Rep* 2020;10:4240.

6 Piccinni C, Dondi L, Ronconi G, et al. HR+/HER2- metastatic breast cancer: Epidemiology, prescription patterns, healthcare resource utilisation and costs from a large Italian real-world database. *Clin Drug Investig* 2019;39:945–51.

7 Mahajan R. Real world data: Additional source for making clinical decisions. *Int J Appl Basic Med Res* 2015;5:82.

8 Bronson MR, Kapadia NS, Austin AM, et al. Leveraging linkage of cohort studies with administrative claims data to identify individuals with cancer. *Med Care* 2018;56:e83–e89.

9 Fenton JJ, Onega T, Zhu W, et al. Validation of a medicare claims-based algorithm for identifying breast cancers detected at screening mammography. *Med Care* 2016;54:e15–22.

10 Gold HT, Do HT. Evaluation of three algorithms to identify incident breast cancer in Medicare claims data. *Health Serv Res* 2007;42:2056–69.

11 Nattinger AB, Laud PW, Bajorunaite R, et al. An algorithm for the use of Medicare claims data to identify women with incident breast cancer. *Health Serv Res* 2004;39:1733–49.

12 Smith GL, Shih YC, Giordano SH, et al. A method to predict breast cancer stage using Medicare claims. *Epidemiol Perspect Innov* 2010;7:1.

13 Yen TW, Laud PW, Sparapani RA, et al. An algorithm to identify the development of lymphedema after breast cancer treatment. *J Cancer Surviv* 2015;9:161–71.

14 Nordstrom BL, Whyte JL, Stolar M, et al. Identification of metastatic cancer in claims data. *Pharmacoepidemiol Drug Saf* 2012;21(Suppl 2):21–8.

15 Xu Y, Kong S, Cheung WY, et al. Development and validation of case-finding algorithms for recurrence of breast cancer using routinely collected administrative data. *BMC Cancer* 2019;19:210.

16 Du XL, Key CR, Dickie L, et al. External validation of Medicare claims for breast cancer chemotherapy compared with medical chart reviews. *Med Care* 2006;44:124–31.

17 Kroenke CH, Chubak J, Johnson L, et al. Enhancing breast cancer recurrence algorithms through selective use of medical record data. *J Natl Cancer Inst* 2016;108:djv336.

18 Chapter 4: Post-marketing surveillance of drugs. Pharmaceutical regulations in Japan: Japan Pharmaceutical Manufacturers Association; 2018. Available at: http://www.jpma.or.jp/english/parj/pdf/2020e_ch04.pdf. Accessed December 21, 2020.

19 Basic concept of validation of outcome definition used in post-marketing database survey: Pharmaceuticals and Medical Devices Agency, Japan; 2020 Available at: https://www.pmda.go.jp/files/000235927.pdf. Accessed October 26, 2020.

20 Task force on validation of indicators obtained from claims centered on injury and illness names in Japan: Japan Society for Pharmacoepidemiology; 2018 Available at:

http://www.jspe.jp/committee/020/0271_1/. Accessed November 10, 2020.

21 Ando T, Ooba N, Mochizuki M, et al. Positive predictive value of ICD-10 codes for acute myocardial infarction in Japan: A validation study at a single center. *BMC Health Serv Res* 2018;18:895.

22 Imai S, Yamana H, Inoue N, et al. Validity of administrative database detection of previously resolved hepatitis B virus in Japan. *J Med Virol* 2019;91:1944–48.

23 Iwamoto M, Higashi T, Miura H, 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;45:1036–41.

24 Lee J, Imanaka Y, Sekimoto M, et al. Validation of a novel method to identify healthcare-associated infections. *J Hosp Infect* 2011;77:316–20.

25 Ooba N, Setoguchi S, Ando T, et al. Claims-based definition of death in Japanese claims database: Validity and implications. *PLoS One* 2013;8:e66116.

26 Takeda T, Mihara N, Murata T, et al. Estimating the ratio of patients with a certain disease between hospitals for the allocation of patients to clinical trials using health insurance claims data in Japan. *Stud Health Technol Inform* 2016;228:537–41.

27 Tanaka S, Hagino H, Ishizuka A, et al. Validation study of claims-based definitions of suspected atypical femoral fractures using clinical information. *Jpn J Pharmacoepidemiol* 2016;21:13–19.

28 Yamana H, Moriwaki M, Horiguchi H, et al. Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. *J Epidemiol* 2017;27:476–82.

29 Koretsune Y, Yamashita T, Yasaka M, et al. Usefulness of a healthcare database for epidemiological research in atrial fibrillation. *J Cardiol* 2017;70:169–79.

30 Sakai M, Ohtera S, Iwao T, et al. Validation of claims data to identify death among aged persons utilizing enrollment data from health insurance unions. *Environ Health Prev*

Med 2019;24:63.

31 Ono Y, Taneda Y, Takeshima T, et al. Validity of Claims Diagnosis Codes for Cardiovascular Diseases in Diabetes Patients in Japanese Administrative Database. *Clin Epidemiol* 2020;12:367–75.

32 Sato I, Yagata H, Ohashi Y. The accuracy of Japanese claims data in identifying breast cancer cases. *Biol Pharm Bull* 2015;38:53–7.

33 Databases available for pharmacoepidemiology researches in Japan (information obtained from survey answers as of August 2020) Japanese Society for Pharmacoepidemiology; 2020 Available at: http://www.jspe.jp/mt-static/FileUpload/files/JSPE_DB_TF_E.pdf. Accessed October 26, 2020.

34 Kimura E, Ueno S. Trends in health information and communication standards in Japan. *J Natl Inst Public Health* 2020;69 52–62.

35 Act on the Protection of Personal Information "The Every-Three-Year Review" Outline of the System Reform 2019 Available at: https://www.ppc.go.jp/files/pdf/APPI_The_Every_Three_Year_Review_Outline_of_the_System_Reform.pdf. Accessed October 26, 2020.

36 Ethical Guidelines for Medical and Health Research Involving Human Subjects: Ministry of Health, Labour and Welfare, Japan; Available at: https://www.mhlw.go.jp/file/06-Seisakujouhou-10600000-Daijinkanboukouseikagakuka/0000080278.pdf. Accessed October 26, 2000.

37 Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982;5:649–55.

38 Cutrona SL, Toh S, Iyer A, et al. Design for validation of acute myocardial infarction cases in Mini-Sentinel. *Pharmacoepidemiol Drug Saf* 2012;21(Suppl 1):274–81.

39 Krysko KM, Ivers NM, Young J, et al. Identifying individuals with multiple sclerosis

in an electronic medical record. *Mult Scler* 2015;21:217–24.

40 Widdifield J, Ivers NM, Young J, et al. Development and validation of an administrative data algorithm to estimate the disease burden and epidemiology of multiple sclerosis in Ontario, Canada. *Mult Scler* 2015;21:1045–54.

41 Iwagami M, Aoki S, Akazawa M, et al. Task force related to validation of indicators obtained from receipt information focusing on disease names in Japan. *Pharmacoepidemiology* 2018;23:95–123.

42 National Cancer Center Council. Survival rate survey Japanese Association of Clinical Cancer Centers; 2019 Available at: http://www.zengankyo.ncc.go.jp/etc/index.html. Accessed October 26, 2020.

43 For the understanding of health insurance treatment [medical department] Guidance and Audit Office, Medical Economics Division, Health Insurance Bureau of the MHLW; 2018 Available at: https://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryou/iryouhoken/dl/shidou_kansa_01.pdf. Accessed December 4, 2020.

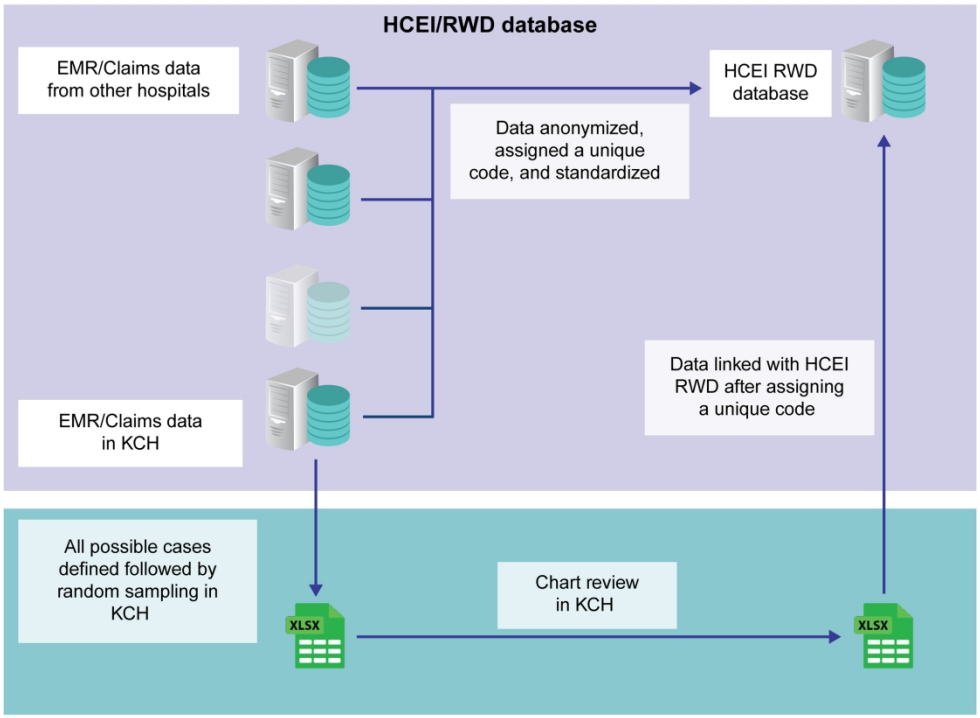
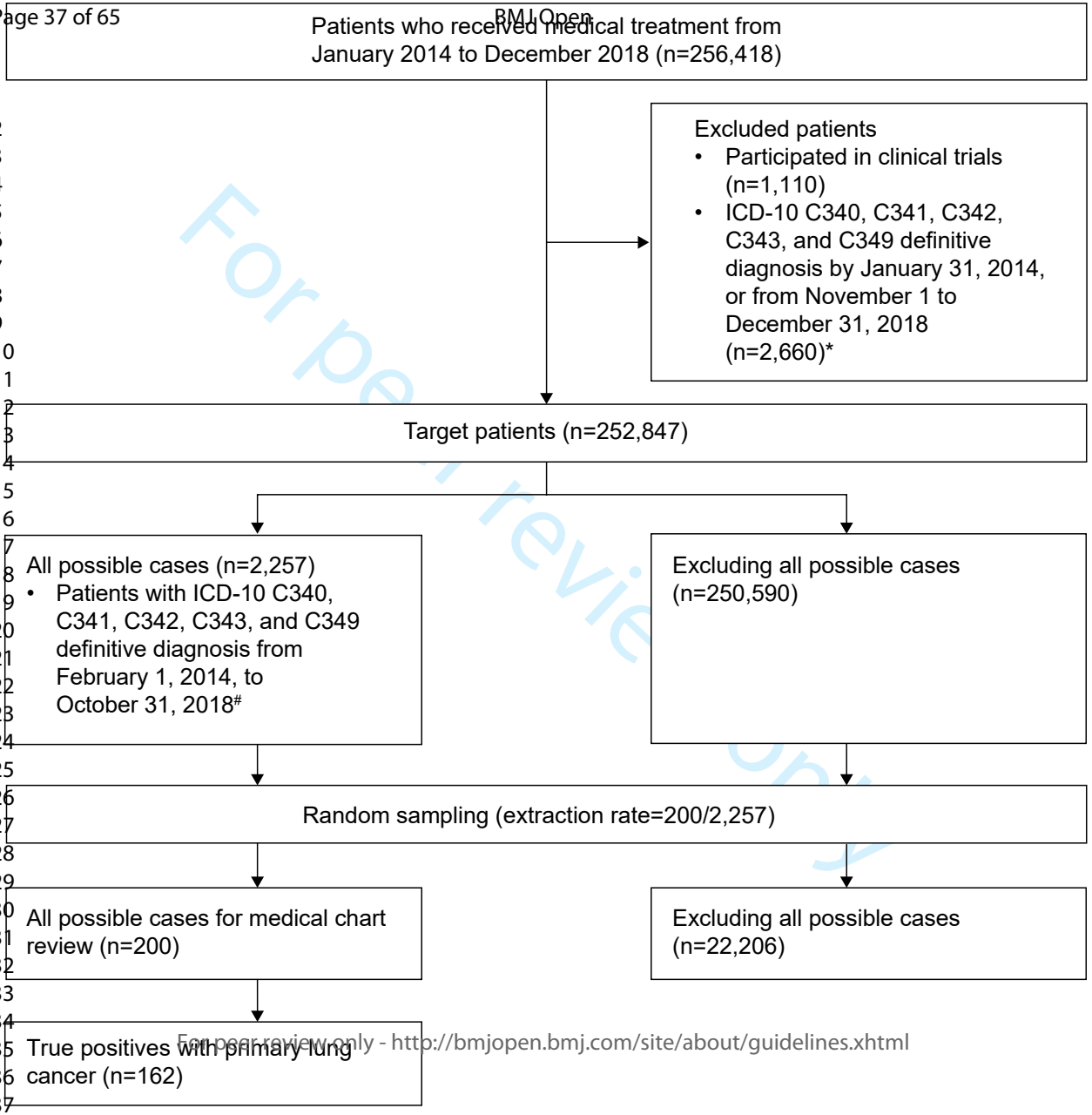
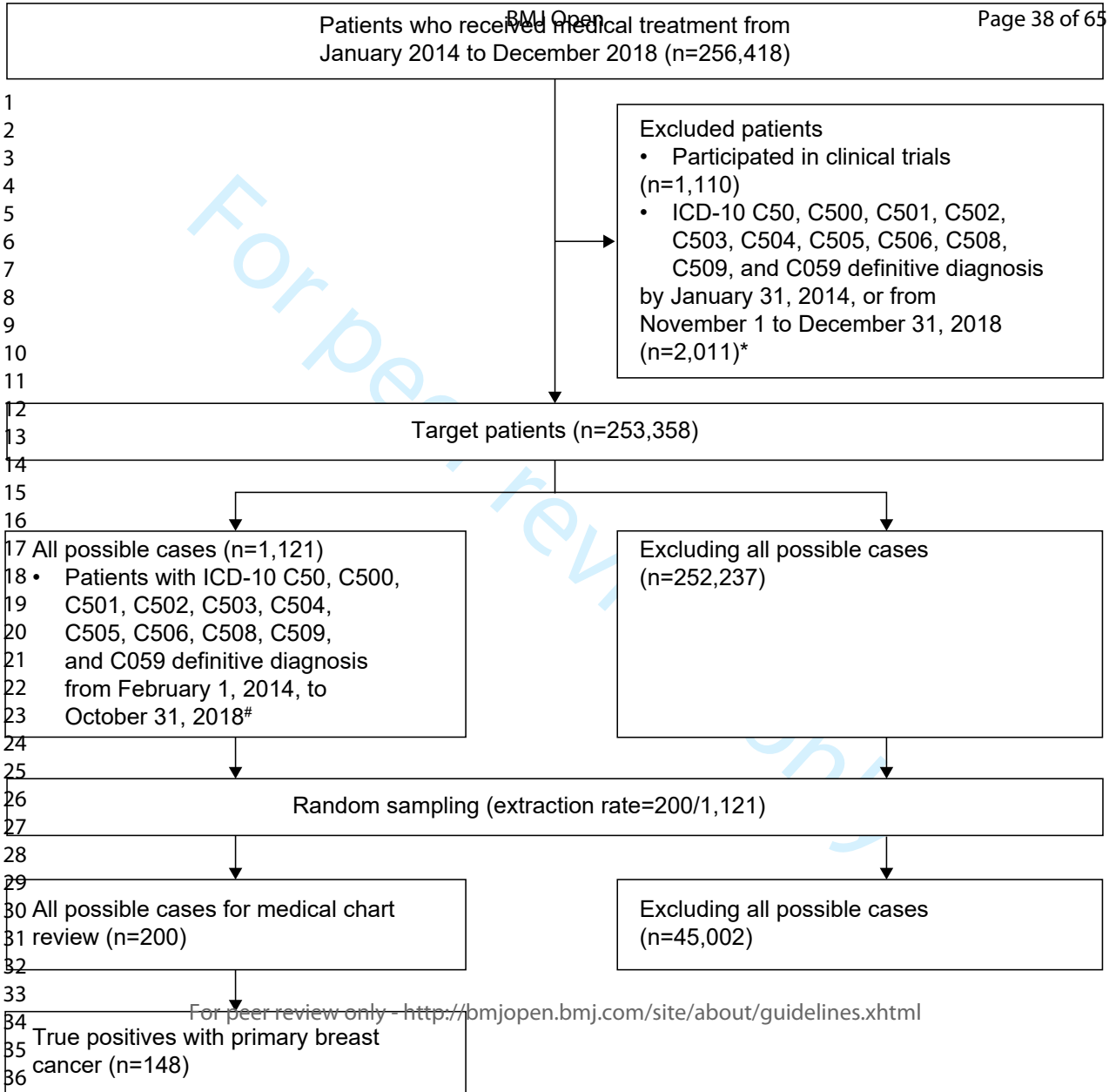
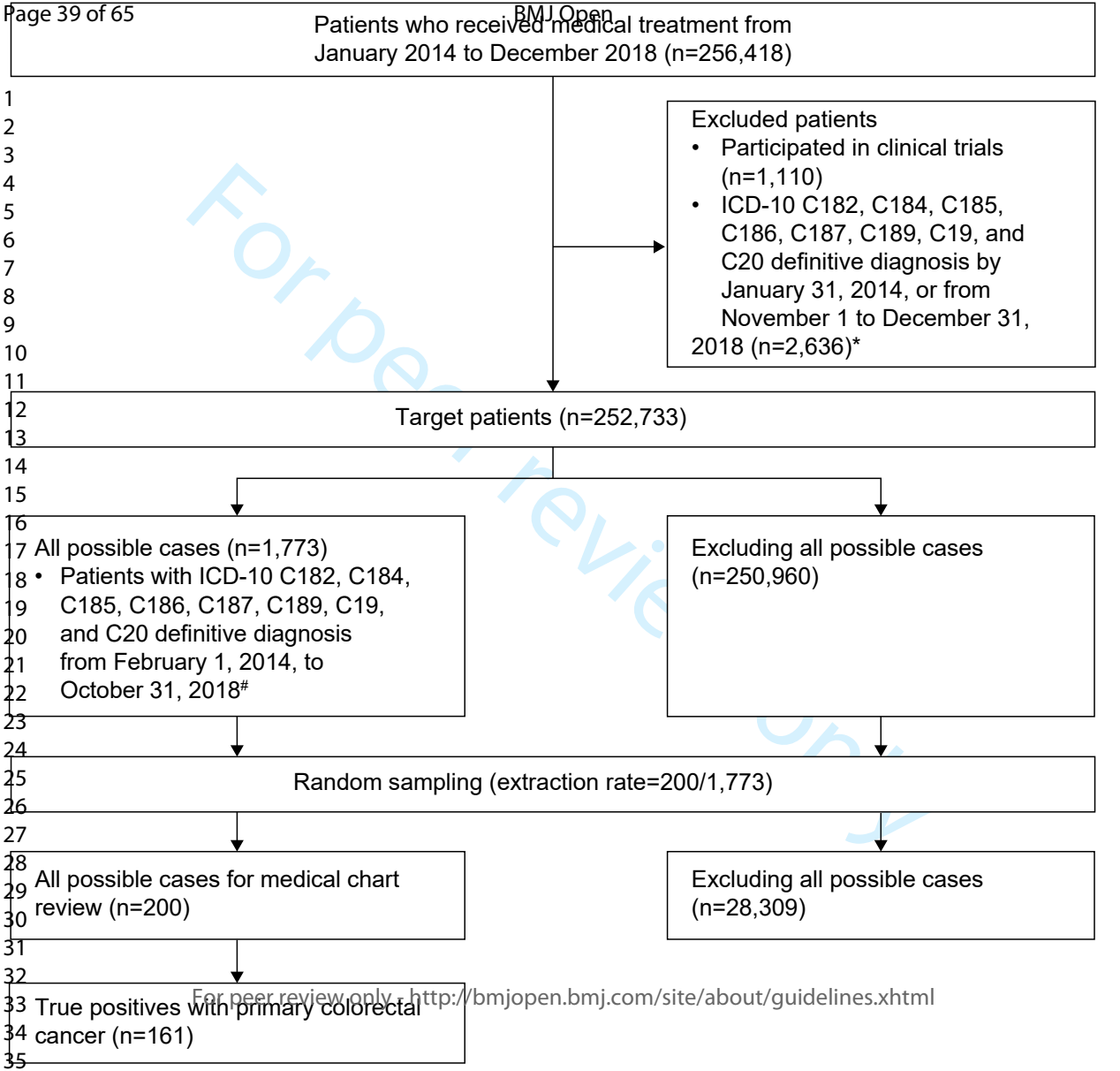


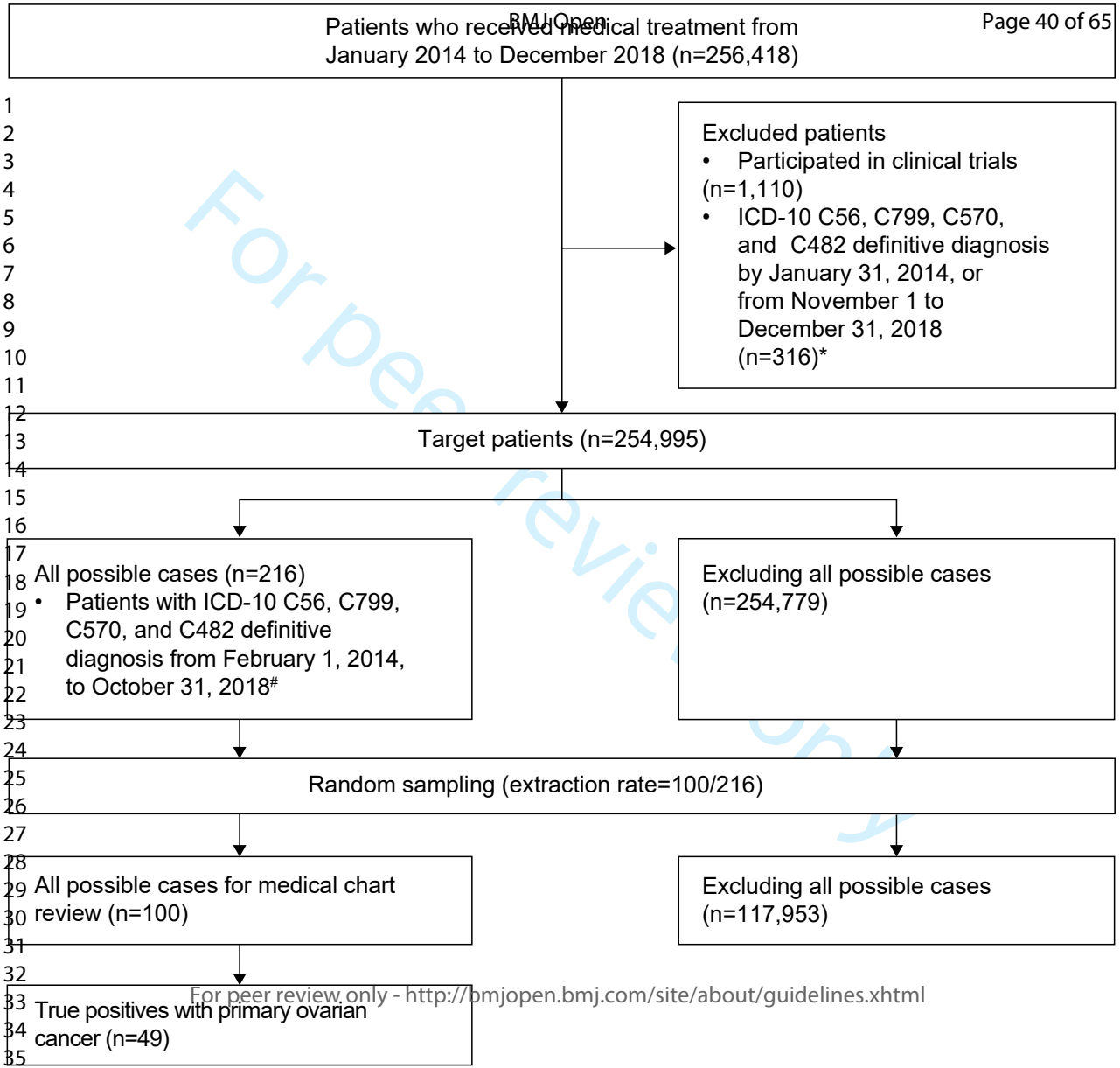
Figure 1. Health, Clinic, and Education Information Evaluation Institute/real-world database
EMR: Electronic medical records; HCEI: Health, Clinic, and Education Information Evaluation Institute; KCH: Kurashiki Central Hospital; RWD: real-world database

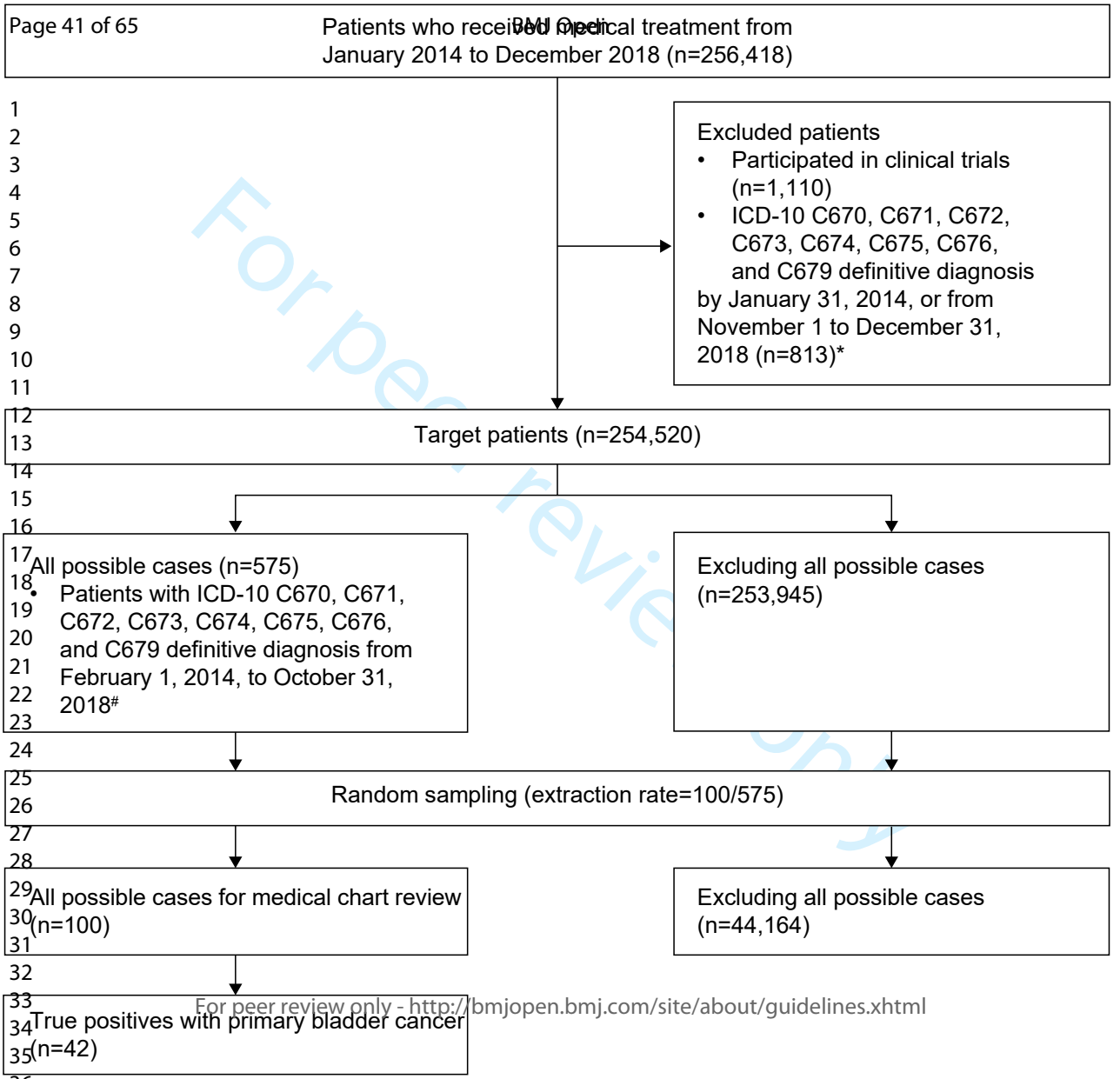
189x139mm (300 x 300 DPI)

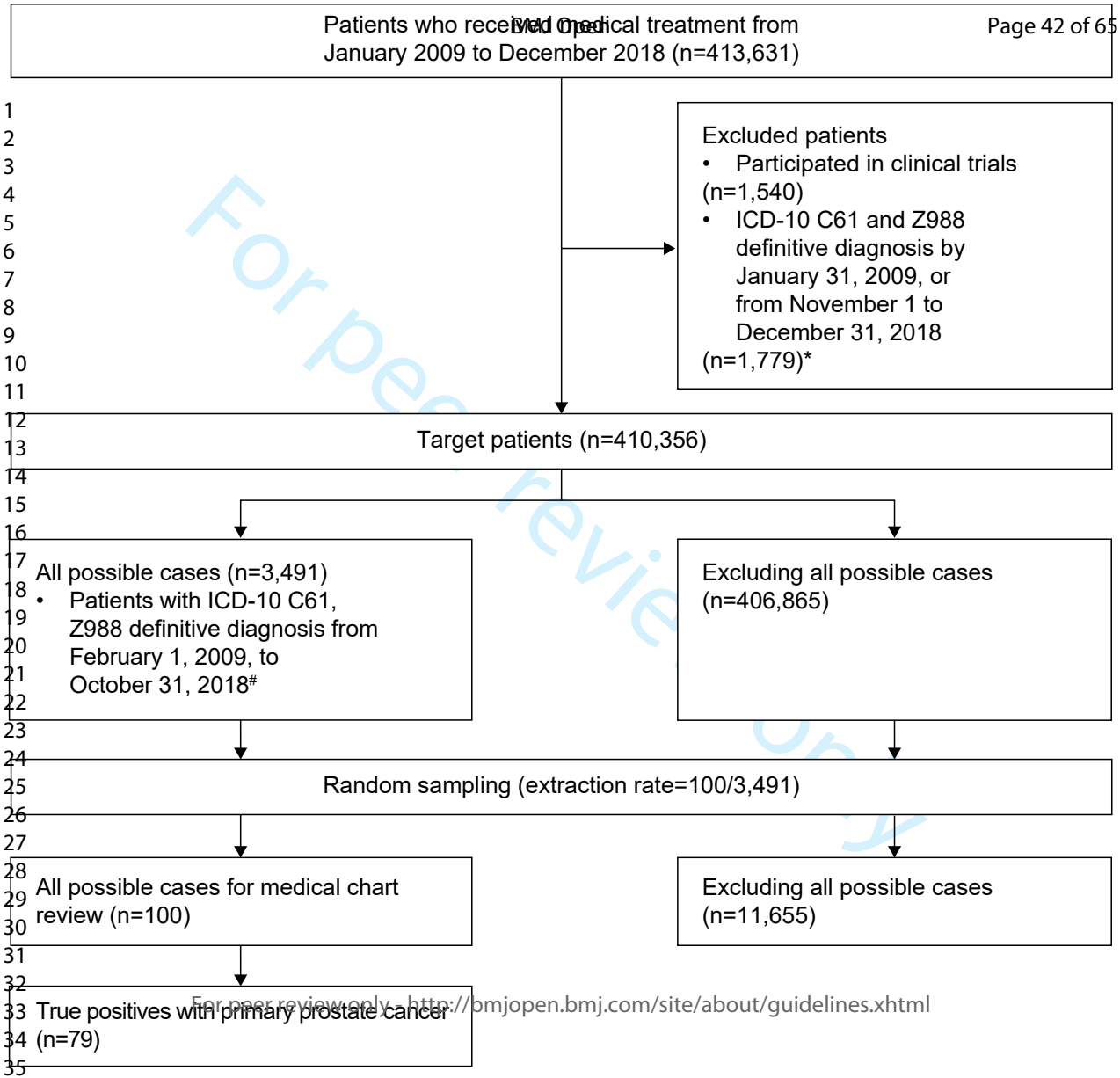












Supplemental Tables

Table S1. Outcome definitions

Outcome	Definition	
A. Primary lung cancer	A1	• Definitive diagnosis of lung cancer (ICD-10: C340, C341, C342, C343, or C349) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.
	A2	• Definitive diagnosis of lung cancer (ICD-10: C340, C341, C342, C343, or C349) recorded between 2014 and 2018 in EMR data.
	A3	• Diagnosis of lung cancer (Japanese original diagnostic code: 1629063) recorded between 2014 and 2018 in EMR data.
	A4	• Definitions written in A1 and specimen examination for laboratory diagnosis (Japanese original procedural code: 160060170, 160060270, 160171470, 160185110, 160214310, 160209750, 160214710, 160214810, 160190270, 160190370, 160190470, 160190570, 160214470, 160214970, or 160062310) recorded between 2014 and 2018 in claims data.
B. Non-small cell lung cancer	B1	• Diagnosis of non-small cell lung cancer (Japanese original diagnostic code: 8847272, 8847732, 8849238, 8847598, 8847637, 8847664, or 8842053) recorded between 2014 and 2018 in EMR data.
	B2	• Diagnosis of non-small cell lung cancer (Japanese original diagnostic code: 8842835, 8847676, 8847677,

Outcome	Definition	
		8847678, 8847679, 8835493, 8847634, 8847635, 8847636, 8847637, 8837666, 8847661, 8847662, 8847663, 8847664, 8831458, 8847595, 8847596, 8847597, 8847598, 8833932, 1629003, 1629006, 1629009, 8838805, 8838844, 8838852, 8838898, 8838901, 8842058, 8842831, 8842832, 8842833, 8842834, 8847272, 8847732, 8849238, 8849788, or 2312002) recorded between 2014 and 2018 in EMR data.
C. Small cell lung cancer	C1	• Diagnosis of small cell lung cancer (Japanese original diagnostic code: 8847594, 8842185, 8847633, 8847660, or 8847675) recorded between 2014 and 2018 in EMR data.
α. Primary breast cancer	α1	• Definitive diagnosis of breast cancer (ICD-10: C500, 501, 502, 503, 504, 505, 506, 508, 509, or D059) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.
	α2	• Definitive diagnosis of breast cancer (ICD-10: C500, 501, 502, 503, 504, 505, 506, 508, 509, or D059) recorded between 2014 and 2018 in EMR data.
	α3	• Diagnosis of breast cancer (Japanese original diagnostic code: 8849699) recorded between 2014 and 2018 in EMR data.
β. Primary colorectal cancer	β1	• Definitive diagnosis of colorectal cancer (ICD-10: C179, C182, C184, C186, C187, C189, C19, or C20) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or

Outcome	Definition	
		most resource-consuming diagnosis.
	β2	• Definitive diagnosis of colorectal cancer (ICD-10: C179, C182, C184, C186, C187, C189, C19, or C20) recorded between 2014 and 2018 in EMR data.
	β3	• Diagnosis of breast cancer (Japanese original diagnostic code: 8847915 or 8847916) recorded between 2014 and 2018 in EMR data.
γ. Primary ovarian cancer	γ1	• Definitive diagnosis of ovarian cancer (ICD-10: C56, C799, C570, or C482) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.
	γ2	• Definitive diagnosis of ovarian cancer (ICD-10: C56, C799, C570, or C482) recorded between 2014 and 2018 in EMR data.
ε. Primary bladder cancer	ε1	• Definitive diagnosis of bladder cancer (ICD-10: C670, C671, C672, C673, C674, C675, C676, or C679) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.
	ε2	• Definitive diagnosis of bladder cancer (ICD-10: C670, C671, C672, C673, C674, C675, C676, or C679) recorded between 2014 and 2018 in EMR data.
δ. Primary prostate cancer	δ1	• Definitive diagnosis of prostate cancer (ICD-10: C61 or Z988) recorded between 2009 and 2018 in DPC

Outcome	Definition	
		data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.
	δ2	• Definitive diagnosis of prostate cancer (ICD-10: C61 or Z988) recorded between 2009 and 2018 in EMR data.
D. Performance status 2 or higher at the start of chemotherapy	D1	Medical treatment of rehabilitation for cancer patients (Japanese original diagnostic code: 180033110) recorded between 2014 and 2018 in claims data, given in the same index month as the prescription month of the therapeutic drug described in Table S2.
	D2	• Medical treatment of rehabilitation for disuse syndrome (Japanese original diagnostic code: H001-02, 180044610, 180044710, 180044810, 180044910, 180045010, 180045110, 180045210, 180045310, 180045410, 180045530, 180045630, 180045730, 180051530, 180051630, 180051730, 180051830, 180051930, 180052030, 180052130, 180052230, 180052330, 180052430, 180052530, or 180052630) recorded between 2014 and 2018 in claims data, given in the same index month as the prescription month of the therapeutic drug described in Table S2.
E. Death	E1	• Date of death in EMR data.
	E2	• Date of death in DPC data.
	E3	• Medical treatment of death for patients (Japanese original diagnostic code: 114007270, 114018670, or 114019970) recorded between 2014 and 2018 in claims data.

Outcome	Definition	
	E4	• 30 days before and after definitions written in E1.
	E5	• 30 days before and after definitions written in E2.
	E6	• 30 days before and after definitions written in E3.
F. First recurrence/progression	F1	• Date of disease name with "recurrence" as a modifier in Japanese original diagnostic code.
	F2	• Second specimen examination for laboratory diagnosis (Japanese original procedural code: 160060170, 160060270, 160171470, 160185110, 160214310, 160209750, 160214710, 160214810, 160190270, 160190370, 160190470, 160190570, 160214470, 160214970, or 160062310) recorded between 2014 and 2018 in claims data.
	F3	• Definitions written in F2 and patients with no history of surgery for the purpose of excision (with or without surgery for the purpose of examination).
	F4	• Month of definitions written in F1.
	F5	• Month of definitions written in F2.
	F6	• Month of definitions written in F3.
G. Second recurrence/progression	G1	• Date of administration of the drug described in Appendix 2 after definitions written in F1.
	G2	• Third specimen examination for laboratory diagnosis (Japanese original procedural code: 160060170, 160060270, 160171470, 160185110, 160214310, 160209750, 160214710, 160214810, 160190270,

Outcome	Definition	
		160190370, 160190470, 160190570, 160214470, 160214970, or 160062310) recorded between 2014 and 2018 in claims data.
	G3	• Month of definitions written in G1.
	G4	• Month of definitions written in G2.
H. Third recurrence/progression	H1	• Date of administration of the drug described in Appendix 2 after G1.
	H2	• Forth specimen examination for laboratory diagnosis (Japanese original procedural code: 160060170, 160060270, 160171470, 160185110, 160214310, 160209750, 160214710, 160214810, 160190270, 160190370, 160190470, 160190570, 160214470, 160214970, or 160062310) recorded between 2014 and 2018 in claims data.
	H3	• Month of definitions written in H1.
	H4	• Month of definitions written in H2.
Adverse events		
I. Interstitial pneumonia	I1	• Definitive diagnosis of interstitial pneumonia (ICD-10: J702, J703, J704, J841 or J849) recorded in EMR data and Medical treatment (ATC code: H02AB04 or H02AB06 [excludes topical drugs]).
	I2	• Definitive diagnosis of interstitial pneumonia (ICD-10: J448, J700, J701, J702, J704, J82, J841, J849, or M0510) recorded in EMR data.

Outcome	Definition	
	I3	• Definitions written in I2 plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
J. Hepatic failure	J1	• Definitive diagnosis of hepatic failure (ICD-10: K720, K712, or K723) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	J2	• Laboratory data abnormality in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	J3	• Definitive diagnosis of hepatic failure (ICD:10: K710, K711, K712, K716, K717, K718, K719, K720, K729, K739, K740, K741, K743, K744, K745, K746, K750, K751, K752, K753, K754, K758, K759, K760, K761, K762, K763, K764, K765, K767, K768, K769, R18, R609, R945, or S361) recorded in EMR data.
	J4	• Definitions written in J3 plus prescription of medical treatment (ATC code: H02AB04, H02AB06, A05AA02, or A05BA08) recorded in claims data.
K. Colitis・diarrhea	K1	• Definitive diagnosis of colitis・diarrhea (ICD:10: A090 or A099) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of

Outcome	Definition	
		external medicine) recorded in claims data.
	K2	• Definitive diagnosis of colitis・diarrhea (ICD-10: A099, K501, K509, K510, K512, K513, K515, K518, K519, K521, K522, K528, K529, K550, K551, K552, K559, K566, K591, K628, K638, K921, K922, M321, or R101) recorded in EMR data.
	K3	• Definitions written in K2 plus prescription of medical treatment (ATC codes: H02AB04, H02AB06, A07A, A07F, A07E, A07D, or A07X) recorded in claims data.
L. Type 1 diabetes	L1	• Prescription of medical treatment (ATC code: A10AB, A10AC, A10AD, or A10AE)
	L2	• Definitive diagnosis of type 1 diabetes (ICD-10: E10, E100, E101, E102, E103, E104, E105, or E106) recorded in EMR data.
M. Encephalitis・meningitis	M1	• Definitive diagnosis of encephalitis・meningitis (ICD-10: G040, G048, G049, or G934) recorded in EMR data.
	M2	• Definitive diagnosis of encephalitis・meningitis (ICD-10: G040, G048, G049, or G934) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	M3	• Definitive diagnosis of encephalitis. • Meningitis (ICD-10: R291) recorded in EMR data.

Outcome	Definition	
	M4	• Definitions written in M3 plus prescription Meningitis (ICD-10: R291) recorded in EMR data of medical treatment (ATC code: J05AB, J01, or J02A) recorded in claims data.
N. Nerve disorder（excludes paresthesia）	N1	• Definitive diagnosis of nerve disorder (excludes paresthesia) (ICD-10: G500, G501, G508, G509, G511, G512, G513, G514, G518, G519, G520, G521, G522, G523, G527, G528, G529, G540, G541, G542, G543, G544, G545, G560, G561, G562, G563, G564, G568, G569, G570, G571, G572, G573, G574, G575, G576, G579, G580, G587, G588, G589, G603, G608, G609, G618, G620, G622, G629, G64, G723, G810, G811, G819, G820, G821, G822, G823, G824, G825, G830, G831, G832, G833, G839, G900, G902, G903, G904, G908, G909, H812, H919, H933, M7924, M7926, M7929, M8900, M998, R252, R253, or R258) recorded in EMR data.
	N2	• Definitions written in N1 and medical treatment (ATC code H02AB04 or H02AB06) recorded in claims data.
O. Myasthenia gravis	O1	• Definitive diagnosis of myasthenia gravis (ICD-10: G700) recorded in EMR data.
	O2	• Definitive diagnosis of myasthenia gravis (ICD-10: G700) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	O3	• Definitive diagnosis of myasthenia gravis (ICD-10: G700, G701, G709) recorded in EMR data.

Outcome	Definition	
	O4	• Definitions written in O3 and medical treatment (ATC code: H02AB04, H02AB06, or H07AA02) recorded in claims data.
P. Guillain-Barré syndrome	P1	• Definitive diagnosis of Guillain-Barré syndrome (ICD-10: G610) recorded in EMR data.
	P2	• Definitions written in P1 plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	P3	• Definitions written in P1 plus prescription of methylprednisolone (ATC code: H02AB04), prednisolone (ATC code: H02AB06 with exception of external medicine), or immunoglobulin recorded in claims data.
	P4	• Definitions written in P1 and medical treatment (ATC code: H02AB04, H02AB06, J06BA, J06BB, or J06BC) recorded in claims data.
Q. Skin disorders	Q1	• Definitive diagnosis of skin disorders (ICD-10: H605, H738, I831, L000, L010, L011, L020, L021, L022, L023, L024, L028, L029, L030, L031, L032, L033, L038, L039, L080, L081, L089, L100, L101, L102, L103, L104, L105, L108, L109, L110, L111, L119, L120, L121, L123, L129, L130, L131, L138, L139, L200, L208, L210, L219, L233, L238, L239, L26, L270, L271, L279, L280, L281, L282, L290, L291, L292, L298, L299, L300, L301, L302, L303, L304, L305, L309, L400, L401, L402, L403, L404, L408, L409, L410, L411, L413, L414, L415, L418, L419, L42, L430, L431, L433, L438, L439, L440, L441, L442, L443, L449, L500, L501, L502, L504, L508, L509, L510, L511, L512, L518, L519, L52, L530,

Outcome	Definition	
		L531, L532, L538, L539, L560, L561, L562, L563, L564, L568, L570, L571, L572, L574, L578, L580, L589, L590, L598, L700, L701, L702, L703, L708, L709, L710, L711, L718, L719, L730, L731, L738, L739, L80, L810, L811, L812, L813, L814, L816, L817, L818, L819, L82, L83, L850, L851, L852, L853, L858, L859, L870, L871, L872, L879, L88, L890, L891, L892, L893, L899, L900, L906, L908, L909, L919, L920, L921, L928, L929, L930, L931, L932, L940, L941, L942, L943, L944, L945, L946, L950, L951, L97, L980, L981, L982, L983, L984, L985, L986, L988, R02, R21, R238, or T783) recorded in EMR data.
	Q2	• Definitions written in Q1 and medical treatment (ATC codes: H02AB04, H02AB06, D04AA, or R01AC [excludes steroidal drugs]) recorded in claims data.
R. Rhabdomyolysis	R1	• "Drug-induced rhabdomyolysis" or "rhabdomyolysis" in definitive diagnosis of rhabdomyolysis (ICD-10: M6289) recorded in EMR data.
	R2	• Definitive diagnosis of rhabdomyolysis (ICD-10: D868, G718, G720, G722, G724, G729, M331, M332, M339, M353, M358, M6019, M6091, M6092, M6095, M6098, M6099, M6105, M6109, M6119, M6129, M6155, M6159, M6289, M7900, M7910, M7911, M7912, M7913, M7915, M7916, M7918, M7919, or M7979) recorded in EMR data.
	R3	• Definitions written in R2 plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone

Outcome	**Definition**	
		(ATC code: H02AB06 with exception of external medicine) recorded in claims data.
S. Myocarditis	S1	• Definitive diagnosis of myocarditis (ICD-10: I401, I408, I409, I514) recorded in EMR data.
	S2	• Definitive diagnosis of myocarditis (ICD-10: I401, I408, I409, I514) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	S3	• Definitive diagnosis of myocarditis (ICD-10: D868, E854, E888, E889, I010, I011, I012, I018, I019, I050, I051, I052, I058, I059, I060, I061, I062, I069, I070, I071, I078, I079, I080, I081, I082, I083, I088, I089, I090, I091, I092, I099, I200, I201, I208, I209, I210, I212, I213, I214, I219, I220, I221, I228, I229, I230, I231, I232, I233, I234, I235, I236, I238, I240, I248, I249, I251, I252, I253, I254, I255, I256, I258, I259, I300, I308, I309, I319, I339, I340, I341, I348, I350, I351, I352, I358, I359, I360, I361, I362, I369, I370, I371, I372, I379, I38, I401, I408, I409, I420, I421, I422, I423, I424, I425, I426, I427, I428, I429, I440, I441, I442, I443, I444, I445, I446, I447, I451, I452, I453, I454, I455, I456, I458, I459, I460, I461, I469, I470, I471, I472, I479, I480, I481, I482, I489, I490, I491, I492, I493, I494, I495, I498, I499, I500, I501, I509, I513, I514, I515, I518, I519, R000, R001, R008, R570, R571, R579, or R943) recorded in EMR data.
	S4	• Definitions written in S3 plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone

Outcome	Definition	
		(ATC code: H02AB06 with exception of external medicine) recorded in claims data.
T. Gastrointestinal perforation	T1	• Definitive diagnosis of gastrointestinal perforation (ICD-10: K255, K265, K631, K65S, or K639) recorded in EMR data.
U. Adrenal insufficiency	U1	• Definitive diagnosis of adrenal insufficiency in Japanese original diagnostic code including the words "autoimmune adrenitis" recorded in claims data and "hypoadrenocorticism" plus medical treatment (ATC: code H02AB09) recorded in claims data.
	U2	• Definitive diagnosis of adrenal insufficiency (ICD-10: E271, E272, E273, E274, E275 or E278) recorded in EMR data.
	U3	• Definitions written in U2 plus medical treatment (ATC code H02AB09) recorded in claims data.
X. Febrile neutropenia	X1	• Definitive diagnosis of febrile neutropenia (ICD-10: D70) recorded in EMR data and medical treatment (Table S2) recorded in claims data.

ATC, Anatomical Therapeutic Chemical; DPC, Diagnosis Procedure Combination; EMR, electronic medical record; ICD-10, ICD-10, International Classification of Diseases, 10th revision

Table S2. Drug codes

ATC code	Common name
L01XC32	Atezolizumab
L01XC17	Nivolumab
L01XC18	Pembrolizumab
L01XC31	Avelumab
L01XC28	Durvalumab
L01XC06	Cetuximab
L01XC08	Panitumumab
L01XE02	Gefitinib
L01XE35	Osimertinib
L01XE47	Dacomitinib
L01XE13	Afatinib
L01XE03	Erlotinib
L01XE36	Alectinib
L01XE44	Lorlatinib
L01XE28	Ceritinib
L01XE16	Crizotinib
L01XC07	Bevacizumab (includes related biosimilars)
L01XC13	Pertuzumab
L01XC14	Trastuzumab emtansine
L01XE07	Lapatinib
L01XE33	Palbociclib
L01XE50	Abemaciclib
L01XE10, L04AA18	Everolimus
L01XX46	Olaparib
L01XC08	Panitumumab
L01XE21	Regorafenib

ATC code	Common name
L01	Anti-malignant tumor drugs excluding talaporfin sodium (620001918), porfimer sodium (620007468), anagrelide hydrochloride hydrate (622379001), and sterile talc (622293901)
L02	Hormone therapy
L04	Immunosuppressive drug
J01CR05	Tazobactam and piperacillin
J01DD02	Ceftazidime hydrate
J01DE03	Cefozopran hydrochloride
J01DE01	Cefepime dihydrochloride hydrate
J01DE02	Cefpirome sulfate
J01DH05	Biapenem
J01DH02	Meropenem hydrate
J01DH51	Imipenem hydrate, cilastatin sodium
J01DH04	Doripenem hydrate
J01DH55	Panipenem and betamipron

Table S3. Accuracy of diagnosis definitions

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Lung cancer								
Primary lung cancer								
A1	132	7	22,237	30	81.5 (74.6–87.1)	100.0 (99.9–100.0)	95.0 (89.9–98.0)	99.9 (99.8–99.9)
A2	162	38	22,206	0	100.0 (96.6–100.0)	99.8 (99.8–99.9)	81.0 (74.9–86.2)	100.0 (100.0–100.0)
A3	19	1	22,243	143	11.7 (7.2–17.7)	100.0 (100.0–100.0)	95.0 (75.1–99.9)	99.4 (99.2–99.5)
A4	128	7	22, 237	34	79.0 (71.8–85.0)	100.0 (99.9–100)	94.8 (89.6–97.9)	99.8 (99.8–99.9)
Non-small cell lung cancer								
B1	46	6	22,280	74	38.3 (29.6–47.6)	100.0 (99.9–100.0)	88.5 (76.6–95.6)	99.7 (99.6–99.7)
B2	46	6	22,280	74	38.3 (29.6–47.6)	100.0 (99.9–100.0)	88.5 (76.6–95.6)	99.7 (99.6–99.7)
Small cell lung cancer								
C1	10	0	22,395	1	90.9 (58.7–99.8)	100.0 (100.0–100.0)	100.0 (58.7–100.0)	100.0 (100.0–100.0)
Breast cancer								
Primary breast cancer								
α1	93	18	45,036	55	62.8 (54.5–70.6)	100.0 (99.9–100.0)	83.8 (75.6–90.1)	99.9 (99.8–99.9)
α2	148	52	45,002	0	100.0 (96.3–100.0)	99.9 (99.8–99.9)	74.0 (67.3–79.9)	100.0 (100.0–100.0)
α3	0	0	45,054	148	0.0 (0.0–3.7)	100.0 (100.0–100.0)	NA	99.7 (99.6–99.7)

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Colorectal cancer								
Primary colorectal cancer								
β1	108	8	28,340	53	67.1 (59.2–74.3)	100.0 (99.9–100.0)	93.1 (86.9–97.0)	99.8 (99.8–99.9)
β2	161	39	28,309	0	100.0 (96.6–100.0)	99.9 (99.8–99.9)	80.5 (74.3–85.8)	100.0 (100.0–100.0)
β3	0	0	28,348	161	0.0 (0.0–3.4)	100.0 (100.0–100.0)	NA	99.4 (99.3–99.5)
Ovarian cancer								
Primary ovarian cancer								
γ1	44	14	11,692	5	89.8 (77.8–96.6)	99.9 (99.8–99.9)	75.9 (62.8–86.1)	100.0 (99.7–100.0)
γ2	49	50	11,656	0	100.0 (89.4–100.0)	99.6 (99.4–99.7)	49.5 (39.3–59.7)	100.0 (100.0–100.0)
Bladder cancer								
Primary bladder cancer								
ε1	33	16	44,206	9	78.6 (63.2–89.7)	100.0 (99.9–100.0)	67.3 (52.5–80.1)	100.0 (100.0–100.0)
ε2	42	58	44,164	0	100.0 (87.7–100.0)	99.9 (99.8–99.9)	42.0 (32.2–52.3)	99.9 (99.8–99.9)
Prostate cancer								
Primary prostate cancer								
δ1	17	0	11,676	62	21.5 (12.1–32.2)	100.0 (100.0–100.0)	100.0 (72.7–100.0)	99.5 (99.3–99.6)
δ2	79	21	11,655	0	100.0 (93.2–100.0)	99.8 (99.7–99.9)	79.0 (69.7–86.5)	100.0 (100.0–100.0)

CI, confidence interval; NA, not available; NPV, negative predictive value; PPV, positive predictive value

Table S4. Accuracy of death definitions

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Lung cancer								
E1	32	0	40	1	97.0 (84.2–99.9)	100.0 (87.1–100.0)	100.0 (84.2–100.0)	97.6 (87.1–99.9)
E2	9	0	40	24	27.3 (13.3–45.5)	100.0 (87.1–100.0)	100.0 (55.5–100.0)	62.5 (49.5–74.3)
E3	0	0	40	33	0.0 (0.0–15.3)	100.0 (87.1–100.0)	NA	54.8 (4.7–66.5)
E4	32	0	40	1	97.0 (84.2–99.9)	100.0 (87.1–100.0)	100.0 (84.2–100.0)	97.6 (87.1–99.9)
E5	9	0	40	24	27.3 (13.3–45.5)	100.0 (87.1–100.0)	100.0 (55.5–100.0)	62.5 (49.5–74.3)
E6	0	0	40	33	0.0 (0.0–15.3)	100.0 (87.1–100.0)	NA	54.8 (4.7–66.5)
Breast cancer								
E1	1	0	104	0	100.0 (1.3–100.0)	100.0 (94.8–100.0)	100.0 (1.3–100.0)	100.0 (94.8–100.0)
E2	0	0	104	1	0.0 (0.0–98.7)	100.0 (94.8–100.0)	NA	99.0 (94.8–100.0)
E3	0	0	104	1	0.0 (0.0–98.7)	100.0 (94.8–100.0)	NA	99.0 (94.8–100.0)
E4	1	0	104	0	100.0 (1.3–100.0)	100.0 (94.8–100.0)	100.0% (1.3–100.0)	100.0 (94.8–100.0)
E5	0	0	104	1	0.0 (0.0–98.7)	100.0 (94.8–100.0)	NA	99.0 (94.8–100.0)
E6	0	0	104	1	0.0 (0.0–98.7)	100.0 (94.8–100.0)	NA	99.0 (94.8–100.0)
Colorectal cancer								
E1	4	0	53	0	100.0 (28.4–100.0)	100.0 (90.1–100.0)	100.0 (28.4–100.0)	100.0 (90.1–100.0)

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
E2	2	0	53	2	50.0 (6.8–93.2)	100.0 (90.1–100.0)	100.0 (9.4–100.0)	96.4 (87.5–99.6)
E3	0	0	53	4	0.0 (0.0–71.6)	100.0 (90.1–100.0)	NA	93.0 (83.0–98.1)
E4	4	0	53	0	100.0 (28.4–100.0)	100.0 (90.1–100.0)	100.0 (28.4–100.0)	100.0 (90.1–100.0)
E5	2	0	53	2	50.0 (6.8–93.2)	100.0 (90.1–100.0)	100.0 (9.4–100.0)	96.4 (87.5–99.6)
E6	0	0	53	4	0.0 (0.0–71.6)	100.0 (90.1–100.0)	NA	93.0 (83.0–98.1)
Ovarian cancer								
E1	5	0	16	0	100.0 (35.9–100.0)	100.0 (71.3–100.0)	100.0 (35.9-100.0)	100.0 (71.3–100.0)
E2	2	0	16	3	40.0 (5.3–85.3)	100.0 (71.3–100.0)	100.0 (9.4–100.0)	84.2 (60.4–96.6)
E3	0	0	16	5	0.0 (0.0–64.1)	100.0 (71.3–100.0)	NA	76.2 (52.8–91.8)
E4	5	0	16	0	100.0 (35.9–100.0)	100.0 (71.3–100.0)	100.0 (35.9-100.0)	100.0 (71.3–100.0)
E5	2	0	16	3	40.0 (5.3–85.3)	100.0 (71.3–100.0)	100.0 (9.4–100.0)	84.2 (60.4–96.6)
E6	0	0	16	5	0.0 (0.0–64.1)	100.0 (71.3–100.0)	NA	76.2 (52.8–91.8)
Bladder cancer								
E1	2	0	8	0	100.0 (9.4–100.0)	100.0 (51.8–100.0)	100.0 (9.4–100.0)	100.0 (51.8–100.0)
E2	1	0	8	1	50.0 (1.3–98.7)	100.0 (51.8–100.0)	100.0 (51.8–100.0)	100.0 (1.3–100.0)
E3	0	0	8	2	0.0 (0.0–90.6)	100.0 (51.8–100.0)	NA	80.0 (44.4–97.5)
E4	2	0	8	0	100.0 (9.4–100.0)	100.0 (51.8–100.0)	100.0 (9.4–100.0)	100.0 (51.8–100.0)

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
E5	1	0	8	1	50.0 (1.3–98.7)	100.0 (51.8–100.0)	100.0 (51.8–100.0)	100.0 (1.3–100.0)
E6	0	0	8	2	0.0% (0.0–90.6)	100.0 (51.8–100.0)	NA	80.0 (44.4–97.5)
Prostate cancer								
E1	3	0	32	1	75.0 (19.4–99.4)	100.0 (94.2–100.0)	100.0 (19.4–100.0)	97.0 (84.2–99.9)
E2	0	0	32	4	0.0 (0.0–71.6)	100.0 (84.2–100.0)	NA	88.9 (73.9–96.9)
E3	0	0	32	4	0.0 (0.0–71.6)	100.0 (84.2–100.0)	NA	88.9 (73.9–96.9)
E4	3	0	32	1	75.0 (19.4–99.4)	100.0 (94.2–100.0)	100.0 (19.4–100.0)	97.0 (84.2–99.9)
E5	0	0	32	4	0.0 (0.0–71.6)	100.0 (84.2–100.0)	NA	88.9 (73.9–96.9)
E6	0	0	32	4	0.0 (0.0–71.6)	100.0 (84.2–100.0)	NA	88.9 (73.9–96.9)

CI, confidence interval; NA, not available; NPV, negative predictive value; PPV, positive predictive value

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	3
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	3
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	5
	4	Study objectives and hypotheses	5 and 6
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	6
Participants	**6**	Eligibility criteria	8
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	8
	8	Where and when potentially eligible participants were identified (setting, location and dates)	8
	9	Whether participants formed a consecutive, random or convenience series	
Test methods	**10a**	Index test, in sufficient detail to allow replication	9
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	5
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	10
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	10
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	8
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	8
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	10-11
	15	How indeterminate index test or reference standard results were handled	10-11
	16	How missing data on the index test and reference standard were handled	10-11
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	Not applicable
	18	Intended sample size and how it was determined	Page 9
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Supplementary figures 1- 6
	20	Baseline demographic and clinical characteristics of participants	Table 4
	21a	Distribution of severity of disease in those with the target condition	Not applicable
	21b	Distribution of alternative diagnoses in those without the target condition	Not applicable
	22	Time interval and any clinical interventions between index test and reference standard	
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Table 2, Table 3, Table S3, Table S4
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	Along with each result in corresponding tables
	25	Any adverse events from performing the index test or the reference standard	Not applicable
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	Page 25
	27	Implications for practice, including the intended use and clinical role of the index test	Page 26
OTHER INFORMATION			

	28	Registration number and name of registry	Page 8
	29	Where the full study protocol can be accessed	No
	30	Sources of funding and other support; role of funders	Page 26

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

BMJ Open

Accuracy of Algorithms to Identify Patients With a Diagnosis of Major Cancers and Cancer Related Adverse Events in an administrative database: A validation study in an acute care hospital in Japan

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-055459.R1
Article Type:	Original research
Date Submitted by the Author:	22-Jan-2022
Complete List of Authors:	Fujiwara, Takashi; Kurashiki Central Hospital, Department of Management, Clinical Research Center; Kurashiki Central Hospital, Department of Otolaryngology/Head and Neck Surgery Kanemitsu, Takashi; Chugai Pharmaceutical Co Ltd, Medical Affairs Division Tajima, Kosei ; Chugai Pharmaceutical Co Ltd, Clinical Development Division Yuri, Akinori; Chugai Pharmaceutical Co Ltd, Drug Safety Division Iwasaku, Masahiro; Kurashiki Central Hospital, Department of Management, Clinical Research Center Okumura, Yasuyuki; Real world Data Co., Ltd. Tokumasu, Hironobu; Kurashiki Central Hospital, Department of Management, Clinical Research Center; Real World Data Co., Ltd.
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Oncology
Keywords:	Adult oncology < ONCOLOGY, Breast tumours < ONCOLOGY, Gynaecological oncology < GYNAECOLOGY, Respiratory tract tumours < ONCOLOGY, Urological tumours < ONCOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title page

Title

Accuracy of Algorithms to Identify Patients With a Diagnosis of Major Cancers and Cancer-Related Adverse Events in an administrative database: A validation study in an acute care hospital in Japan

Authors

Takashi Fujiwara[1,2], Takashi Kanemitsu[3], Kosei Tajima[4], Akinori Yuri[5], Masahiro Iwasaku[6], Yasuyuki Okumura[7], Hironobu Tokumasu[6,7]

Affiliations

1. Department of Public Health Research, Kurashiki Clinical Research Institute, Okayama, Japan

2. Department of Otolaryngology/Head and Neck Surgery, Kurashiki Central Hospital, Okayama, Japan

3. Medical Affairs Division, Chugai Pharmaceutical Co., Ltd., Tokyo, Japan

4. Clinical Development Division, Chugai Pharmaceutical Co., Ltd., Tokyo, Japan

5. Drug Safety Division, Chugai Pharmaceutical Co., Ltd., Tokyo, Japan

6. Department of Management, Clinical Research Center, Kurashiki Central Hospital, Okayama, Japan

7. Real World Data Co., Ltd., Kyoto, Japan

Corresponding author

Name: Hironobu Tokumasu

Address: Department of Management, Clinical Research Center, Kurashiki Central Hospital, 1-1-1 Miwa, Kurashiki, Okayama, 710-8602 Japan

Phone number: +81-86-422-0210

Fax number: +81- 86-421-3424

Email: tokumasu@rwdata.co.jp

Word count: 4013

Abstract

Objectives: Validation studies in oncology are limited in Japan. This study was conducted to evaluate the accuracy of diagnosis and adverse event (AE) definitions for specific cancers in a Japanese health administrative real-world database (RWD).

Design and setting: Retrospective observational validation study to assess the diagnostic accuracy of electrical medical records (EMRs) and claim coding regarding oncology diagnosis and AEs based on medical record review in the RWD. The sensitivity and positive predictive value (PPV) with 95% confidence intervals (CIs) were calculated.

Participants: The validation cohort included patients with lung (n=2,257), breast (n=1,121), colorectal (n=1,773), ovarian (n=216), and bladder (n=575) cancer who visited the hospital between January 2014 and December 2018, and those with prostate cancer (n=3,491) visiting between January 2009 and December 2018, who were identified using EMRs.

Outcomes: Key outcomes included primary diagnosis, deaths, and AEs.

Results: For primary diagnosis, sensitivity and PPV for the respective cancers were as follows: lung, 100.0% (96.6–100.0) and 81.0% (74.9–86.2); breast, 100.0% (96.3–100.0) and 74.0% (67.3–79.9); colorectal, 100.0% (96.6–100.0) and 80.5% (74.3–85.8); ovarian, 89.8% (77.8–96.6) and 75.9 (62.8–86.1); bladder, 78.6% (63.2–89.7) and 67.3% (52.5–80.1); prostate, 100.0% (93.2–100.0) and 79.0% (69.7–86.5). Sensitivity and PPV for death were as follows: lung, 97.0% (84.2–99.9) and 100.0% (84.2–100.0); breast, 100.0% (1.3–100.0) and 100.0% (1.3–100.0); colorectal, 100.0% (28.4–100.0) and 100.0% (28.4–100.0); ovarian, 100.0% (35.9–100.0) and 100.0% (35.9–100.0); bladder, 100.0% (9.4–100.0) and 100.0% (9.4–100.0); prostate, 75.0% (19.4–99.4) and 100.0% (19.4–100.0). Overall, PPV tended to be low, with the definition based on International Classification of Diseases, 10th revision alone for AEs.

Conclusion: Diagnostic accuracy was not so high, and therefore needs to be further

investigated.

Trial registration: University hospital Medical Information Network (UMIN) Clinical Trials Registry; UMIN000039345.

Strengths and limitations of this study

- To our knowledge, this is the first study in oncology in Japan that validates disease and adverse event definitions in a health administrative real-world database (RWD) using chart review based on electronic medical records data from a hospital as the reference standard.
- Validation was performed at a single facility, which may limit generalizability and transportability of the results.
- Study results are limited by the inherent issues related to the use of an RWD, which primarily stores medical information for the purpose of insurance claims.
- The diagnosis and adverse event definitions used in this study may not be the most suitable; thus, there is an opportunity to further deepen these definitions.
- Study methods for the consolidation of true positives for events with low incidence need to be further investigated as it was challenging to investigate outcomes with extremely low incidence.

Keywords

database, electronic medical record, health administrative, real-world database, validation study

INTRODUCTION

In recent years, evidence from routine clinical practice using data from real-world databases (RWDs) has increasingly gained importance in decision-making in healthcare, research, and drug development.[1] In addition, RWD studies can help generate evidence for advancement

in precision medicine and facilitation of targeted and efficient patient care.[2] In line with this trend, evidence related to several aspects, such as health technology, expenditure forecasting, survival outcomes, time to therapy, and treatment efficacy, is increasingly being collected from RWD studies in oncology.[3-6]

However, it is important to validate case-identification algorithms to evaluate the accuracy of information sourced from RWDs, which is usually collected for purposes other than research.[7] To this end, several studies have been conducted outside of Japan to evaluate the accuracy of algorithms based on health administrative data in identifying cancer diagnoses or other outcomes using databases, such as registries, population-based cohorts, chart reviews, and electronic medical records (EMRs) as reference standards.[8-17]

The implementation of the revised ordinance of Good Postmarketing Study Practice by the Pharmaceuticals and Medical Devices Agency (PMDA) of Japan in 2018 suggests that the importance of using RWDs in post-marketing surveillance to investigate the safety and efficacy of pharmaceutical products is being recognized in Japan as well.[18] To encourage validation studies, the PMDA of Japan and Japan Society for Pharmacoepidemiology established a basic concept for conducting validation studies to verify diagnosis codes and other outcome definitions in Japanese RWDs.[19, 20] However, to our knowledge, only a few claims-based validation studies [21-32] have reported on outcomes in cancer [32, 33] to date. Thus, this necessitates validation studies on a wider range of cancer types in Japan using a reliable database as a reference standard. This study was conducted for validation of diagnosis and adverse event (AE) definitions for specific cancers in a Japanese RWD using a chart review by EMR.

PATIENTS AND METHODS

Study design

This was a validation study of diagnosis and AE definitions in the health administrative RWD of the Health, Clinic, and Education Information Evaluation Institute (HCEI) conducted by chart review of EMRs from Kurashiki Central Hospital, Japan, as the reference standard.

Data collection

Data were collected retrospectively from EMRs at the Kurashiki Central Hospital, Japan (**Figure 1**), which were the primary data source. All possible cases that met the diagnosis and AE definitions and cases other than all possible cases were identified using International Classification of Diseases, 10th revision (ICD-10) codes (**Figures S1–S6**) from the EMRs. Further, these cohorts were randomly sampled to verify the diagnoses and related events. EMRs were manually reviewed to verify the diagnosis of all possible cases. This verified dataset was anonymized and sent to Real World Data Co. Ltd., the vendor for HCEI. The verified dataset was linked deterministically to claims data and EMRs originally derived from the hospital.

Chart review based on EMR

A chart review for all possible cases was conducted by medical professionals, including medical doctors involved in the management of cancer patients and four clinical research coordinators (CRCs) at the Kurashiki Central Hospital, Japan. The diagnosis of cancer was made primarily by histopathological tests, followed by radiological diagnosis and findings based on the physician's clinical examination. At least two CRCs conducted chart reviews independently. Any disagreements were resolved by the two CRCs and by a medical doctor, if still unresolved.

HCEI database

HCEI is an integrated RWD initiated in Japan and supported by Real World Data Co., Ltd. (Kyoto).[34] As of August 2020, HCEI was collecting information from approximately

20 million patients from 190 medical institutions in Japan, including Kurashiki Central Hospital. The HCEI database covers 1.2% of the overall Japanese population and includes data from 1.3 million outpatients and 0.21 million inpatients in 2019.[34] Medical information is extracted from EMRs, claims, and Diagnosis Procedure Combination (DPC) in the HCEI database. Patient-level data from DPC, EMRs, and claims are integrated in advance at the hospital, anonymized, linked to a unique code, and standardized (**Figure 1**). The linked data are then provided to HCEI for storage on their server. Information on procedures (such as surgery) is obtained from claims, while information on laboratory tests and treatments is obtained from EMRs. Diagnosis data are obtained from both claims and EMRs. Per HCEI's security policy, personal identifiable information (such as date of birth) is not collected during data extraction. Master lists are constructed based on the national standards of the Ministry of Health, Labour and Welfare (MHLW) of Japan.[35]

Study ethics

This study was approved by the Research Institute of Healthcare Data Science (https://rihds.org/ethic/) (RI2019010) and the institutional ethics committee of Kurashiki Central Hospital (KCH3301), and conducted under the tenets of the Declaration of Helsinki, Act on the Protection of Personal Information,[36] and Ethical Guidelines for Medical and Health Research Involving Human Subjects.[37] It was conducted under a joint research agreement between Kurashiki Central Hospital, Chugai Pharmaceutical Co., Ltd., and HCEI, and is registered at the UMIN Clinical Trials Registry (UMIN000039345). Target patients at Kurashiki Central Hospital could opt, on the hospital's website, to not disclose their information.

Patient and public involvement in research

Patients or the public were not involved in the design or conduct, reporting or dissemination

plans of our research.

Patient selection

Patients with lung, breast, colorectal, ovarian, and bladder cancer who visited Kurashiki Central Hospital between January 2014 and December 2018 (**Figures S1–S5**), and those with prostate cancer (**Figure S6**) who visited the hospital between January 2009 and December 2018 were eligible for the study. Further information on inclusion criteria is provided in **Table S1**. Patients participating in clinical trials during the data extraction periods and those who were assigned the respective ICD-10 code for lung, colorectal, breast, ovarian, and bladder cancer from January 1, 2014, to January 31, 2014, and from November 1, 2018, to December 31, 2018, and that for prostate cancer from January 1, 2009, to January 31, 2009, and from November 1, 2018, to December 31, 2018, were excluded from the study. Patients diagnosed during these periods were excluded to avoid bias due to the time lag between suspected diagnosis by medical examination and confirmation of diagnosis by biopsy, when the outcome definition was potentially met.

The cohort entry date was the date when the respective cancer was diagnosed—January 2014 for lung, breast, colorectal, ovarian, and bladder cancer and January 2009 for prostate cancer—and the end date was December 31, 2018. To avoid selection of cases diagnosed before the cohort entry date, patients who were assigned the respective ICD-10 code for lung, colorectal, breast, ovarian, and bladder cancer before December 31, 2013, and that for prostate cancer before December 31, 2008, were excluded.

Eligible patients were stratified by random sampling as all possible and not possible cases. All possible cases included patients who met the ICD-10 code for the respective support during the specified data extraction period. Patients who were never assigned an ICD-10 code for the respective cancer; those with lung, colorectal, breast, ovarian, and bladder cancer who visited the hospital between January 1, 2014, and December 31, 2018;

and those with prostate cancer between January 1, 2009, and December 31, 2018, were stratified as not possible cases. Overall, 200 cases each with lung, breast, or colorectal cancer and 100 cases each with ovarian, bladder, or prostate cancer were targeted and randomly selected from all possible cases for the EMR review, and not possible cases were also randomly selected using the same proportions.

Outcomes and assessment of accuracy

Outcomes for validation included primary diagnosis, performance status (PS) ≥2,[38] first/second/third recurrence or exacerbation, death, and AEs, particularly immune-related AEs (irAEs), associated with new diagnoses for patients with lung, breast, colorectal, ovarian, bladder, and prostate cancer. AEs included interstitial pneumonia, liver dysfunction, colitis/diarrhea, type 1 diabetes mellitus (T1DM), encephalitis/meningitis, nerve disorders (excluding paresthesia), myasthenia gravis, Guillain-Barré syndrome, skin disorder, rhabdomyolysis, myocarditis, perforation of digestive tract/fistula, hypoadrenocorticism, and febrile neutropenia.

Outcomes were defined by separate algorithms (**Tables S2 and S3**) for each cancer type using one variable or a combination of ≥2 variables, such as diagnoses, treatments, procedures, and laboratory test results. Lung cancer was further classified as primary, non-small cell, and small cell.

Statistical analysis

The target sample size for random sampling was determined based on the feasibility of chart review. If ≥100 patients each meet the definition of primary diagnosis and true positives, the 95% confidence intervals (CIs) for positive predictive value (PPV) and sensitivity can be estimated with a precision of up to ±10% for lung, breast, and colorectal cancer.[39] The sample size for ovarian, bladder, and prostate cancer was half that for lung, breast, and

colorectal cancer.

In the dataset submitted by HCEI, accuracy for each cancer type was evaluated using sensitivity, specificity, PPV, and negative predictive value (NPV) for primary diagnosis, first recurrence/exacerbation, and death. Other outcomes were evaluated using only PPV to determine if the cases were true for those meeting the outcome definition. AEs were validated in patients with true primary cancer who had received chemotherapy. PPV was calculated only after confirming whether the outcome occurred within (before or after) 30 days of the patient meeting the outcome definition.

All possible cases refer to the population that is assumed to include all true patients,[19, 40-42] and included patients who met the ICD-10 code for the respective cancer in EMRs during the specified data extraction period. True positives were defined as patients in whom the outcomes occurred based on HCEI information and EMR review. In addition, patients were randomly selected from cases other than all possible cases at the same extraction rate as that for "all possible cases" to calculate the specificity and NPV for primary diagnosis, first recurrence/exacerbation, and death. The data extraction period for different cancer types was estimated based on the national survival rate survey of 2019 conducted by the National Cancer Center Council,[43] in which the survival period was 10 years for prostate cancer and 5 years for other cancer types. Likewise, a longer data extraction period was considered for prostate cancer to allow for the collection of true positives.

The frequency and 95% CIs were calculated for sensitivity, specificity, PPV, and NPV. 95% CIs were calculated by the symmetric CI method. The degree of agreement between two chart reviewers was evaluated using the kappa coefficient. Extrapolability of the Kurashiki Central Hospital database to that of other hospitals in HCEI database was assessed by comparing the distribution of patient characteristics (age at data extraction, sex, age at time of granting ICD10, observation periods). Outcome definitions used for identification of patients

were as follows: A1 for lung cancer, α1 for breast cancer, β1 for colorectal cancer, γ1 for ovarian cancer, ε1 for bladder cancer, and δ1 for prostate cancer (**Table S2**). Statistical analyses were conducted using R-4.0.2 software.

RESULTS

Patient disposition

Of the 256,418 patients who received medical treatment from 2014 to 2018, 2,257 with lung cancer (**Figure S1**), 1,121 with breast cancer (**Figure S2**), 1,773 with colorectal cancer (**Figure S3**), 216 with ovarian cancer (**Figure S4**), and 575 with bladder cancer (**Figure S5**) were included as all possible cases (**Table 1**). From 2009 to 2018, 3,491 patients with prostate cancer of 413,631 patients receiving medical treatment (**Figure S6**) were included as all possible cases (**Table 1**).

Table 1. Study cohort

Cancer type	Study period for patient selection and chart review	Patients who underwent medical treatment during the study periods, n	Target patients, n	All possible cases, n	True cases, n
Lung cancer	January 2014 to December 2018	256,418	252,847	2,257	162
Breast cancer	January 2014 to December 2018	256,418	253,358	1,121	148
Colorectal cancer	January 2014 to December 2018	256,418	252,733	1,773	161
Ovarian cancer	January 2014 to December 2018	256,418	254,995	216	49
Bladder cancer	January 2014 to December 2018	256,418	254,520	575	42
Prostate cancer	January 2009 to December 2018	413,631	410,356	3,491	79

For identifying patients with each cancer type, the following outcome definitions were used: A1 for lung cancer, α1 for breast cancer, β1 for colorectal cancer, γ1 for ovarian cancer, ε1 for bladder cancer, and δ1 for prostate cancer (Table S2).

Lung cancer

The kappa value in chart reviews for diagnosis definitions was 0.982 (95% CI: 0.947–1.017) for primary lung cancer, 0.979 (95% CI: 0.950–1.008) for non-small cell lung cancer (NSCLC), 1.00 for small cell lung cancer (SCLC), and 0.982 (95% CI: 0.947–1.017) for death. There were 30 false negatives and 132 true positives for A1 using DPC diagnosis (**Figure 2**). Sensitivity was 100% with A2 using related definitive diagnosis (**Figure 2**). Although specificity, PPV, and NPV for NSCLC were high for B1 and B2 using cancer-related diagnosis codes, sensitivity was low (38.3%; **Table S4**). Accuracy was high for all statistical parameters for SCLC (**Figure 2**). Data on death could be extracted with high accuracy using EMR definitions (E1; **Figure 3**).

Breast cancer

The kappa value in the chart review for diagnosis definitions was 1.000 and 0.961 (95% CI: 0.917–1.005) for death. The sensitivity was 100% for α2 using EMR diagnosis (**Figure 2**) Sensitivity was as low as 62.8% and there were 55 false negatives in α1 using DPC diagnosis (**Table S4**). The accuracy of death definitions for breast cancer was challenging to calculate because outcome events were very few owing to good disease prognosis (**Table S5**).

Colorectal cancer

The kappa value in the chart review for both diagnosis definitions and death was 0.953 (95% CI: 0.900–1.006). There were 39 false positives in β2 (**Figure 2**); 15 were diagnosed with colorectal cancer before 2014, two had malignancies that were excluded, and the remaining patients were diagnosed with another cancer on subsequent EMR examination. Death occurred in 4/57 target patients, and sensitivity and specificity of E1 were 100% each (**Figure 3**).

Ovarian cancer

The kappa value in the chart review for diagnosis definitions was 0.920 (95% CI: 0.843–0.997) and 0.940 (95% CI: 0.873–1.007) for death. PPV was higher with γ1 than with γ2 (75.9% vs 49.5%; **Table S4**). Sensitivity was higher with γ2 than with γ1 (100.0% vs 89.8%; **Table S4**). Death occurred in 5/21 target patients, and the sensitivity and specificity of E1 were 100% each (**Figure 3**).

Bladder cancer

The kappa value in the chart review for diagnosis definitions was 0.898 (95% CI: 0.812–0.985) and 0.878 (95% CI: 0.784–0.973) for death. Sensitivity was 100% in ε2, but PPV was as low as 42.0% (**Table S4**). PPV was higher with ε1 than with ε2 (67.3% vs 42.0%; **Table S4**). Death occurred in 2/10 target patients, and the sensitivity and specificity of E1 were 100% each (**Figure 3**).

Prostate cancer

The kappa value in the chart review for diagnosis definitions was 0.875 (95% CI: 0.755–0.995) and 0.9045 (95% CI: 0.798–1.011) for death. PPV was 100% in δ1 (**Table S4**), and sensitivity was 100% in δ2 (**Figure 2**). Death occurred in 4/36 target patients, and the sensitivity and specificity of E1 were 75% and 100%, respectively (**Figure 3**).

Adverse events

The overall PPV for all cancer types was <50%: 47.1% for interstitial pneumonia, 34.6% for liver disorders, 25.5% for colitis/diarrhea, and 13.3% for nerve disorders (excluding paresthesia) by related ICD-10 definitive diagnosis. Although PPV was 100% for encephalitis/meningitis and gastrointestinal perforation by related ICD-10 definitive diagnosis, only one case each was identified as these are rare AEs. For skin disorders, PPV was 76.4% by related ICD-10 definitive diagnosis and 70.4% when treatments were combined in the definition. A combination of related ICD-10 definitive diagnosis and treatments resulted in a PPV of 87.5% for liver disorders. By ICD-10-related definitive diagnosis and intravenous antibiotics use, PPV was 76.9%–100% for febrile neutropenia. PPV was 0% for T1DM.

No events of myasthenia gravis, Guillain-Barré syndrome, rhabdomyolysis, adrenal hypofunction, and myocarditis were identified in this analysis.

Other outcomes

Only one true positive case was extracted for PS ≥2 for lung cancer using the definition of rehabilitation status. Of 51 patients who had received chemotherapy, the PS was 0–1 for 33 patients, 2–4 for 16 patients, and unclear for two patients. Thus, only 1 (6.3%) true positive case with PS ≥2 was extracted using the definition of chemotherapy. Therefore, despite a PPV of 100.0%, it could be challenging to use the current definition of PS ≥2 in an administrative database study. Similarly, the accuracy of the definition of first recurrence/exacerbation was extremely low for all cancer types owing to very few true positives. Since the accuracy of the second and third recurrence/exacerbation was calculated based on the number of true positives during first recurrence/exacerbation, it could not be evaluated.

Extrapolability of EMR data

Sex and age of all possible cases at the Kurashiki Central Hospital and all hospitals were similar (**Table 2**).

Table 2. Demographic and observation period of study population

	All possible cases, n	Male, n (%)	Age (years) at data extraction, mean (SD)	Age (years) at the time of granting ICD-10, mean (SD)	Observation period (days) mean (SD)	Observation period (days) person-years
Lung cancer						
Kurashiki Central Hospital	2,477	1,728 (69.8)	75.0 (9.9)	72.8 (10.2)	801.4 (626.7)	1,985,024
All hospitals	19,861	13,136 (66.1)	74.8 (10.2)	73.5 (10.4)	523.9 (552.4)	10,405,993
Breast cancer						
Kurashiki Central Hospital	1,166	10 (0.9)	67.0 (13.3)	64.1 (13.3)	1,023.6 (650.8)	1,192,400
All hospitals	18,289	131 (0.7)	64.7 (14.1)	62.6 (14.1)	780.5 (618.6)	14,274,791
Colorectal cancer						
Kurashiki Central Hospital	1,684	989 (58.7)	73.6 (11.3)	71.1 (11.6)	930.5 (613.5)	1,566,924
All hospitals	23,501	13,836 (58.9)	74.1 (11.3)	72.1 (11.5)	770.6 (596.2)	18,110,552
Ovarian cancer						
Kurashiki Central Hospital	265	34 (12.8)	66.4 (15.4)	63.9 (15.5)	896.2 (653.5)	237,497

	All possible cases, n	Male, n (%)	Age (years) at data extraction, mean (SD)	Age (years) at the time of granting ICD-10, mean (SD)	Observation period (days) mean (SD)	Observation period (days) person-years
All hospitals	2,592	145 (5.6)	64.1 (14.9)	62.3 (15.1)	6672 (581.1)	1,729,551
Bladder cancer						
Kurashiki Central Hospital	568	446 (78.5)	77.6 (10.0)	75.0 (10.5)	9910 (611.8)	563,042
All hospitals	7,408	5,810 (78.4)	76.9 (10.4)	74.9 (10.6)	7999 (595.8)	5,925,496
Prostate cancer						
Kurashiki Central Hospital	3,131	3,057 (97.6)	76.5 (8.4)	71.9 (8.7)	1,7031 (1,118.3)	5,332,446
All hospitals	32,136	28,690 (89.3)	77.7 (8.9)	74.2 (9.2)	1,3413 (1,041.6)	43,105,126

ICD-10, International Classification of Diseases, 10th revision; SD, standard deviation

DISCUSSION

To our knowledge, this is the first study in oncology in Japan that validates disease names and AE definitions in an RWD by using chart review based on EMR as the gold standard. The diagnostic accuracy of primary diagnosis definitions by ICD-10 code in EMRs and DPC was evaluated. The PPV of diagnosis definition by DPC was relatively high, but sensitivity tended to be low. Although the diagnosis definition using DPC showed false negatives, it can be used for identifying patients with the respective disease. In the definitions using a definitive diagnosis from claims, PPV tended to decrease, but sensitivity tended to increase, thereby suggesting the importance of selecting outcome definition according to the purpose of the study.

The diagnostic accuracy of lung cancer by histological classification varied, with a sensitivity of 90.9% and PPV of 100.0% for SCLC and a sensitivity of 38.3% and PPV of 88.5% for NSCLC. Since the database is used primarily for insurance purposes, precise histological classification of lung cancer in EMR was likely not considered an important documentation item by physicians; therefore, only 38.3% of NSCLC patients received ICD-10 code of NSCLC. In SCLC, further studies to investigate improved methods of extracting false negatives are warranted.

The sensitivity for the EMR definition of breast cancer was 100% and DPC definition was as low as 62.8%. However, specificity was high with both EMR and DPC, and PPV ranged between 74.0% and 83.8%. In a previous study,[33] high sensitivity, specificity, and PPV were observed using definitions obtained by combining diagnostic and procedure codes in a Japanese claims database, suggesting that a combination of codes may result in higher accuracy.

The accuracy of the evaluation for death was high (97.0% sensitivity and 100.0% PPV) using the EMR definition for lung cancer. Although the sensitivity was high using the

EMR definition for other cancers as well, further studies with a larger sample size are needed for confirmation. In cancer types other than lung cancer, which generally have a short survival according to the national cancer survival rate survey,[43] high sensitivity and PPV were observed with some definitions. The number of true negatives was high due to a longer survival at Kurashiki Central Hospital than expected, resulting in fewer deaths, which made the evaluation challenging. Thus, further investigation is necessary. In Japan, a death notification is submitted to the city office in case of death, but it is not linked to the hospital information system and EMRs. Therefore, there is a high likelihood of death data getting missed. However, Kurashiki Central Hospital follows up patients to check their health status, including death, and the likelihood of missing death data was therefore minimal.

Identification of cases with "recurrence/exacerbation" was extremely difficult in all cancer types by definition using items such as diagnoses with "recurrent" as a modifier, pathology-related medical practice code, or relevant surgical history. A previous validation study in breast cancer conducted using cancer registry and health maintenance organization data in the United States suggested that the quality of recurrence data may improve by using multiple recurrence algorithms, and a second cancer record in a cancer registry may potentially improve the diagnostic accuracy of recurrence [17] In another validation study conducted in Canada, Xu et al assessed the recurrence of breast cancer using data extracted from discharge abstracts, physician billing claims, and the National Ambulatory Care Reporting System.[15] They achieved a sensitivity of 94.2% and a PPV of 79.2% using definitions based on second round of chemotherapy, diagnostic procedures, treatment, visit to oncologists, patient age, and tumor stage.[15] True positives may be identified if specific therapies are used for the first recurrence/exacerbation, but further investigation is required. Similarly, PS ≥2, an important variable for cancer, needs further investigation as it was extremely difficult to identify in this study.

For AEs, PPV tended to be low overall with a definition based on ICD-10 alone, suggesting that a combination of definitions based on specific treatment modalities for AEs could be more appropriate. The definitions of febrile neutropenia and skin disorders had high PPVs and, therefore, can be generalized. The validation of T1DM as an AE was challenging as it was difficult to differentiate whether it was an existing comorbidity or developed newly. Moreover, T1DM as a primary diagnosis is rarely found, as the treatment usually targets complications of T1DM. For a few AEs, no true positives were identified, possibly because the outcome definition was developed for irAEs. However, owing to the absence of any reference standard for irAEs in clinical practice, chart review was instead conducted for AEs in general. For AEs with a low incidence, further large studies with a more appropriate validation method are required.

Since RWDs contain a large volume of information, it is not realistic to perform validation of multiple outcomes using all cases; instead, representative samples should be used as much as possible. However, such investigations are possible only in a small number of medical facilities. An efficient and precise validation dataset that comprehensively represents the database of a medical facility is required to minimize bias. Furthermore, definition of the disease and outcomes with low incidence should allow for the collection of as many true positives as possible.

In our study, all possible cases were extracted using the related ICD-10 code from medical information available in the study institution. The Health Insurance Bureau of the MHLW requires that a suspected diagnosis is changed to a definitive diagnosis as soon as a diagnosis is confirmed.[44] Since the RWD used in this study is a health insurance database, patients with a definitive diagnosis identified by ICD-10 code were deemed as all possible cases. To confirm the robustness of this hypothesis, 100 cases for each cancer type were randomly sampled from cases other than all possible cases to ensure that no patients with a

primary diagnosis were included. A more efficient method is warranted for validation before a pharmacoepidemiology study using information from an RWD. In randomized controlled trials (RCTs), the efficacy and safety of treatments are assessed objectively; therefore, assessments are preset. However, in daily clinical practice, treatment decisions are subjective and based on the availability and type of medical resources, capabilities, treatment cost, and patient needs. Therefore, diagnosis and outcome definitions based on efficacy and safety assessments used in RCTs may not be suitable in RWD studies and should be carefully evaluated for use in daily clinical practice.

In this study, validation was performed at a single facility, potentially limiting generalizability and transportability of the results. Further, the results are limited by the inherent issues related to use of an RWD, which primarily stores medical information for the purpose of insurance claims. Moreover, ICD-10 codes for patients diagnosed or treated in other hospitals could be missing from EMRs at Kurashiki Central Hospital. Furthermore, chart review of all patients was not conducted in this study. Therefore, patients with a primary diagnosis among other than all possible cases could have been misclassified as true negatives, potentially underestimating the number of false negatives. Moreover, the diagnosis and AE definitions used in this study may not be the most suitable, and there is an opportunity to further deepen the definitions. For instance, the definition of AE in this study was developed based on treatment-associated irAEs and information on therapeutic agents such as steroids and treatments for allergy; however, definitions based on therapies used for general AE treatment could have been more appropriate. Furthermore, it was challenging to investigate outcomes with an extremely low incidence, for example, certain AEs. Therefore, study methods for consolidation of true positives for events with low incidence need to be investigated.

CONCLUSIONS

The results from our study suggest that diagnostic accuracy was not so high. DPC data could identify only a limited proportion of patients with cancer, while claims or DPC data could identify only a limited proportion of deceased patients. Since the number of cases was limited in this study, further investigation is required to validate the definitions using DPC and claims data. In view of the current claims process in Japan, EMR data are deemed appropriate to comprehensively identify patients with cancer or deceased patients for postmarketing surveillance using RWD. Although a high PPV was observed for a few AEs, precision could have been low owing to the low incidence of AEs, and therefore, validation of AEs warrants further investigation.

Acknowledgments

The following persons from Kurashiki Central Hospital Clinical Research Center (Department of Management, Clinical Research Center, Kurashiki Central Hospital, Okayama, Japan) provided additional support: Maki Satomi coordinated at the study site for implementation of protocol procedures and Ryo Ishida, Emi Sato, Mami Yamaguchi, and Yuri Komatsubara contributed to the chart review. Takeshi Kimura of Real World Data Co., Ltd. provided support for statistical analysis and Yusuke Miyoshi of Chugai Pharmaceuticals Co., Ltd. provided administrative support. Akihiro Seki of Chugai Pharmaceuticals supported in developing the outcome definitions.

Editorial support in the form of medical writing, assembling tables, and creating high-resolution images based on the authors' detailed directions, collating author comments, copyediting, fact checking, and referencing was provided by Dr. Deepali Garg, MBBS, PGDHA, of Cactus Life Sciences (part of Cactus Communications) and funded by Chugai Pharmaceutical Co., Ltd.

Funding

This study was funded by Chugai Pharmaceutical Co., Ltd.

Competing interests

TK, KT, and AY are employees of Chugai Pharmaceutical Co., Ltd. TF reports personal fee for statistical analysis from Real World Data Co., Ltd. during the conduct of the study; personal fee for collaborative research from Chugai Pharmaceutical Co., Ltd.; and personal fee for statistical analysis from Real World Data Co., Ltd. outside the submitted work. MI has nothing to disclose. YO is an employee of Real World Data Co., Ltd. and reports personal fees from MSD K.K., Otsuka Pharmaceutical, and Kurashiki Central Hospital, outside the submitted work. HT reports personal fees for lecture from AYUMI Pharmaceutical Corporation and Chugai Pharmaceutical Co., Ltd., outside the submitted work and is an employee of Kurashiki Central Hospital and the Director of Real World Data, Co., Ltd.

Author contributions

TF contributed to the study concept and design, and collection, analysis, and interpretation of data. TK, KT, YA and HT contributed to study concept and design, and data interpretation. MI contributed to collection and interpretation of data. YO and YA contributed to analysis and interpretation of data. All authors provided final approval for the version to be published.

Data sharing statement

Data are available upon reasonable request.

Figure legends

Figure 1. Health, Clinic, and Education Information Evaluation Institute/real-world database

EMR, electronic medical record; HCEI, Health, Clinic, and Education Information Evaluation Institute; KCH, Kurashiki Central Hospital; RWD, real-world database

Figure 2. Diagnosis definitions with high* accuracy

CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value

*All accuracy values included for a definition are approximately 70% or more.

Figure 3. Death definitions with high* accuracy

CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value

*All accuracy values included for a definition are >70%.

Figure S1. Patient disposition: Lung cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 199 duplicates; #Study observation periods lasted from January 1, 2014, to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods;

100 patients were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer.

Random sampling was performed based on the extraction percentage.

Figure S2. Patient disposition: Breast cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 61 duplicates; #Study observation periods lasted from January 1, 2014 to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer.

Random sampling was performed based on the extraction percentage.

Figure S3. Patient disposition: Colorectal cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 61 duplicates; #Study observation periods lasted from January 1 to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer.

Random sampling was performed based on the extraction percentage.

Figure S4. Patient disposition: Ovarian cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including three duplicates; #Study observation periods lasted from January 1 to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer.

Random sampling was performed based on the extraction percentage.

Figure S5. Patient disposition: Bladder cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 25 duplicates; #Study observation periods lasted from January 1, 2014, to January 31, 2014, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients were randomly sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer.

Random sampling was performed based on the extraction percentage.

Figure S6. Patient disposition: Prostate cancer

ICD-10, International Classification of Diseases, 10th Revision

*Including 44 duplicates; #Study observation periods lasted from January 1, 2009, to January 31, 2009, and from November 1, 2018, to December 31, 2018, but were excluded as study washout periods; 100 patients were sampled from patients other than all possible cases (patients given a suspected diagnosis of related ICD-10) to confirm non-diagnosis of primary cancer.

Random sampling was performed based on the extraction percentage.

References

1 Miksad RA, Abernethy AP. Harnessing the power of real-world evidence (RWE): A checklist to ensure regulatory-grade data quality. *Clin Pharmacol Ther* 2018;103:202–05.

2 Tsai CJ, Riaz N, Gomez SL. Big data in cancer research: Real-world resources for precision oncology to improve cancer care delivery. *Semin Radiat Oncol* 2019;29:306–10.

3 Hess LM, Cui ZL, Mytelka DS, et al. Treatment patterns and survival outcomes for patients receiving second-line treatment for metastatic colorectal cancer in the USA. *Int J*

Colorectal Dis 2019;34:581–88.

4 Lin YS, Shen YC, Wu CY, et al. Danshen improves survival of patients with breast cancer and dihydroisotanshinone I induces ferroptosis and apoptosis of breast cancer cells. *Front Pharmacol* 2019;10:1226.

5 Liu JM, Lin CC, Liu KL, et al. Second-line hormonal therapy for the management of metastatic castration-resistant prostate cancer: A real-world data study using a claims database. *Sci Rep* 2020;10:4240.

6 Piccinni C, Dondi L, Ronconi G, et al. HR+/HER2- metastatic breast cancer: Epidemiology, prescription patterns, healthcare resource utilisation and costs from a large Italian real-world database. *Clin Drug Investig* 2019;39:945–51.

7 Mahajan R. Real world data: Additional source for making clinical decisions. *Int J Appl Basic Med Res* 2015;5:82.

8 Bronson MR, Kapadia NS, Austin AM, et al. Leveraging linkage of cohort studies with administrative claims data to identify individuals with cancer. *Med Care* 2018;56:e83–e89.

9 Fenton JJ, Onega T, Zhu W, et al. Validation of a medicare claims-based algorithm for identifying breast cancers detected at screening mammography. *Med Care* 2016;54:e15–22.

10 Gold HT, Do HT. Evaluation of three algorithms to identify incident breast cancer in Medicare claims data. *Health Serv Res* 2007;42:2056–69.

11 Nattinger AB, Laud PW, Bajorunaite R, et al. An algorithm for the use of Medicare claims data to identify women with incident breast cancer. *Health Serv Res* 2004;39:1733–49.

12 Smith GL, Shih YC, Giordano SH, et al. A method to predict breast cancer stage using Medicare claims. *Epidemiol Perspect Innov* 2010;7:1.

13 Yen TW, Laud PW, Sparapani RA, et al. An algorithm to identify the development of

lymphedema after breast cancer treatment. *J Cancer Surviv* 2015;9:161–71.

14 Nordstrom BL, Whyte JL, Stolar M, et al. Identification of metastatic cancer in claims data. *Pharmacoepidemiol Drug Saf* 2012;21(Suppl 2):21–8.

15 Xu Y, Kong S, Cheung WY, et al. Development and validation of case-finding algorithms for recurrence of breast cancer using routinely collected administrative data. *BMC Cancer* 2019;19:210.

16 Du XL, Key CR, Dickie L, et al. External validation of Medicare claims for breast cancer chemotherapy compared with medical chart reviews. *Med Care* 2006;44:124–31.

17 Kroenke CH, Chubak J, Johnson L, et al. Enhancing breast cancer recurrence algorithms through selective use of medical record data. *J Natl Cancer Inst* 2016;108:djv336.

18 Chapter 4: Post-marketing surveillance of drugs. Pharmaceutical regulations in Japan: Japan Pharmaceutical Manufacturers Association; 2020. Available at: https://www.jpma.or.jp/english/about/parj/eki4g6000000784o-att/2020e_ch04.pdf Accessed December 21, 2021.

19 Basic concept of validation of outcome definition used in post-marketing database survey: Pharmaceuticals and Medical Devices Agency, Japan; 2020 Available at: https://www.pmda.go.jp/files/000235927.pdf. Accessed December 22, 2021.

20 Task force on validation of indicators obtained from claims centered on injury and illness names in Japan: Japan Society for Pharmacoepidemiology; 2018 Available at: http://www.jspe.jp/committee/020/0271_1/. Accessed January 13, 2022.

21 Ando T, Ooba N, Mochizuki M, et al. Positive predictive value of ICD-10 codes for acute myocardial infarction in Japan: A validation study at a single center. *BMC Health Serv Res* 2018;18:895.

22 Imai S, Yamana H, Inoue N, et al. Validity of administrative database detection of previously resolved hepatitis B virus in Japan. *J Med Virol* 2019;91:1944–48.

23 Iwamoto M, Higashi T, Miura H, 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;45:1036–41.

24 Lee J, Imanaka Y, Sekimoto M, et al. Validation of a novel method to identify healthcare-associated infections. *J Hosp Infect* 2011;77:316–20.

25 Ooba N, Setoguchi S, Ando T, et al. Claims-based definition of death in Japanese claims database: Validity and implications. *PLoS One* 2013;8:e66116.

26 Takeda T, Mihara N, Murata T, et al. Estimating the ratio of patients with a certain disease between hospitals for the allocation of patients to clinical trials using health insurance claims data in Japan. *Stud Health Technol Inform* 2016;228:537–41.

27 Tanaka S, Hagino H, Ishizuka A, et al. Validation study of claims-based definitions of suspected atypical femoral fractures using clinical information. *Jpn J Pharmacoepidemiol* 2016;21:13–19.

28 Yamana H, Moriwaki M, Horiguchi H, et al. Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. *J Epidemiol* 2017;27:476–82.

29 Koretsune Y, Yamashita T, Yasaka M, et al. Usefulness of a healthcare database for epidemiological research in atrial fibrillation. *J Cardiol* 2017;70:169–79.

30 Sakai M, Ohtera S, Iwao T, et al. Validation of claims data to identify death among aged persons utilizing enrollment data from health insurance unions. *Environ Health Prev Med* 2019;24:63.

31 Ono Y, Taneda Y, Takeshima T, et al. Validity of Claims Diagnosis Codes for Cardiovascular Diseases in Diabetes Patients in Japanese Administrative Database. *Clin Epidemiol* 2020;12:367–75.

32 Shigemi D, Morishima T, Yamana H, et al. Validity of initial cancer diagnoses in the Diagnosis Procedure Combination data in Japan. *Cancer Epidemiol* 2021;74:102016.

33 Sato I, Yagata H, Ohashi Y. The accuracy of Japanese claims data in identifying breast cancer cases. *Biol Pharm Bull* 2015;38:53–7.

34 Databases available for pharmacoepidemiology researches in Japan (information obtained from survey answers as of August 2020) Japanese Society for Pharmacoepidemiology; 2020 Available at: http://www.jspe.jp/mt-static/FileUpload/files/JSPE_DB_TF_E.pdf. Accessed October 26, 2020.

35 Kimura E, Ueno S. Trends in health information and communication standards in Japan. *J Natl Inst Public Health* 2020;69 52–62.

36 Act on the Protection of Personal Information "The Every-Three-Year Review" Outline of the System Reform 2019 Available at: https://www.ppc.go.jp/files/pdf/APPI_The_Every_Three_Year_Review_Outline_of_the_System_Reform.pdf. Accessed December 22, 2021.

37 Ethical Guidelines for Medical and Health Research Involving Human Subjects: Ministry of Health, Labour and Welfare, Japan; Available at: https://www.mhlw.go.jp/file/06-Seisakujouhou-10600000-Daijinkanboukouseikagakuka/0000080278.pdf. Accessed December 22, 2021.

38 Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982;5:649–55.

39 Cutrona SL, Toh S, Iyer A, et al. Design for validation of acute myocardial infarction cases in Mini-Sentinel. *Pharmacoepidemiol Drug Saf* 2012;21(Suppl 1):274–81.

40 Krysko KM, Ivers NM, Young J, et al. Identifying individuals with multiple sclerosis in an electronic medical record. *Mult Scler* 2015;21:217–24.

41 Widdifield J, Ivers NM, Young J, et al. Development and validation of an administrative data algorithm to estimate the disease burden and epidemiology of multiple sclerosis in Ontario, Canada. *Mult Scler* 2015;21:1045–54.

42 Iwagami M, Aoki S, Akazawa M, et al. Task force related to validation of indicators obtained from receipt information focusing on disease names in Japan. *Pharmacoepidemiology* 2018;23:95–123.

43 National Cancer Center Council. Survival rate survey Japanese Association of Clinical Cancer Centers; 2019 Available at: http://www.zengankyo.ncc.go.jp/etc/index.html. Accessed October 26, 2020.

44 For the understanding of health insurance treatment [medical department] Guidance and Audit Office, Medical Economics Division, Health Insurance Bureau of the MHLW; 2018 Available at: https://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryou/iryouhoken/dl/shidou_kansa_01.pdf. Accessed December 22, 2021.

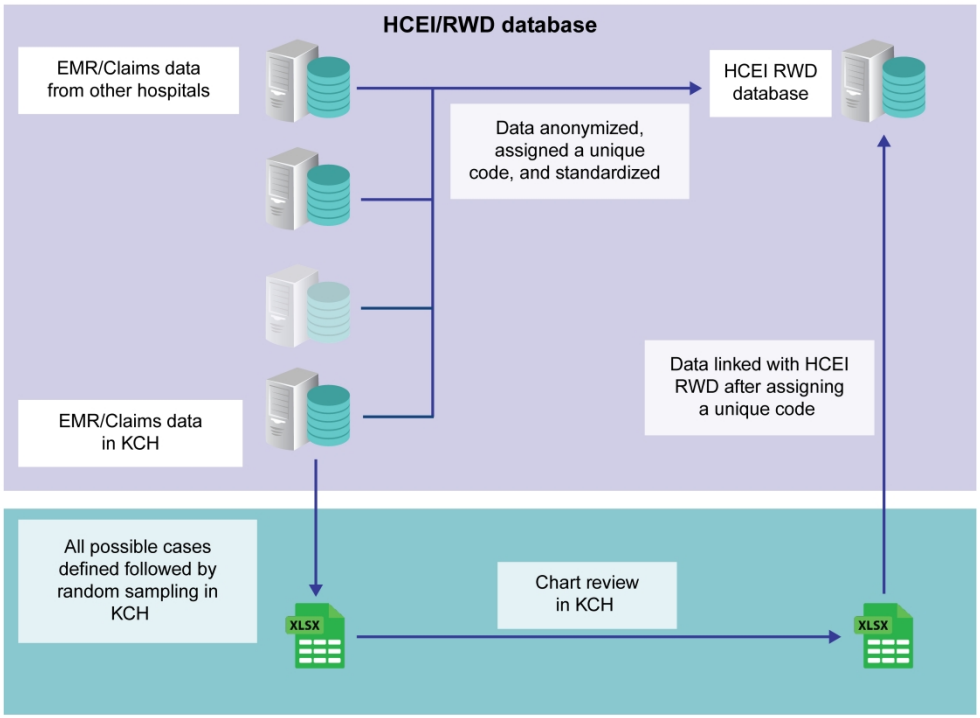


Figure 1. Health, Clinic, and Education Information Evaluation Institute/real-world database
EMR: Electronic medical records; HCEI: Health, Clinic, and Education Information Evaluation Institute; KCH: Kurashiki Central Hospital; RWD: real-world database

A. Primary lung cancer (Kappa value [95% CI]: 0.982 [0.947–1.017])

		Index test (A1)		
		Positive (n)	**Negative (n)**	
Reference standard	Positive (n)	132	7	**PPV (%)** = 95.0 95% CI: 89.9–98.0
	Negative (n)	30	22,237	**NPV (%)** = 99.9 95% CI: 99.8–99.9
		Sensitivity (%) = 81.5 95% CI: 74.6–87.1	**Specificity (%)** = 100.0 95% CI: 99.9–100.0	

		Index test (A2)		
		Positive (n)	**Negative (n)**	
Reference standard	Positive (n)	162	38	**PPV (%)** = 81.0 95% CI: 74.9–86.2
	Negative (n)	0	22,206	**NPV (%)** = 100.0 95% CI: 100.0–100.0
		Sensitivity (%) = 100.0 95% CI: 96.6–100.0	**Specificity (%)** = 99.8 95% CI: 99.8–99.9	

		Index test (A4)		
		Positive (n)	**Negative (n)**	
Reference standard	Positive (n)	128	7	**PPV (%)** = 94.8 95% CI: 89.6–97.9
	Negative (n)	34	22,237	**NPV (%)** = 99.8 95% CI: 99.8–99.9
		Sensitivity (%) = 79. 95% CI: 71.8–85.0	**Specificity (%)** = 100.0 95% CI: 99.9–100	

B. Small cell lung cancer (Kappa value [95% CI]: 1.000 [1.000–1.000])

		Index test (C1)		
		Positive (n)	**Negative (n)**	
Reference standard	Positive (n)	10	0	**PPV (%)** = 100.0 95% CI: 58.7–100.0
	Negative (n)	1	22,395	**NPV (%)** = 100.0 95% CI: 100.0–100.0
		Sensitivity (%) = 90.9 95% CI: 58.7–99.8	**Specificity (%)** = 100.0 95% CI: 100.0–100.0	

C. Primary breast cancer (Kappa value [95% CI]: 1.000 [1.000–1.000])

		Index test (α2)		
		Positive (n)	**Negative (n)**	
Reference standard	Positive (n)	148	52	**PPV (%)** = 74.0 95% CI: 67.3–79.9
	Negative (n)	0	45,002	**NPV (%)** = 100.0 95% CI: 100.0–100.0
		Sensitivity (%) = 100.0 95% CI: 96.3–100.0	**Specificity (%)** = 99.9 95% CI: 99.8–99.9	

D. Primary colorectal cancer (Kappa value [95% CI]: 0.953 [0.900–1.006])

		Index test (β2)		
		Positive (n)	**Negative (n)**	
Reference standard	Positive (n)	161	39	**PPV (%)** = 80.5 95% CI: 74.3–85.8
	Negative (n)	0	28,309	**NPV (%)** = 100.0 95% CI: 100.0–100.0
		Sensitivity (%) = 100.0 95% CI: 96.6–100.0	**Specificity (%)** = 99.9 95% CI: 99.8–99.9	

E. Primary ovarian cancer (Kappa value [95% CI]: 0.920 [0.843–0.997])

		Index test (γ1)		
		Positive (n)	**Negative (n)**	
Reference standard	Positive (n)	44	14	**PPV (%)** = 75.9 95% CI: 62.8–86.1
	Negative (n)	5	11,692	**NPV (%)** = 100.0 95% CI: 99.7–100.0
		Sensitivity (%) = 89.8 95% CI: 77.8–96.6	**Specificity (%)** = 99.9 95% CI: 99.8–99.9	

F. Primary bladder cancer (Kappa value [95% CI]: 0.898 [0.812–0.985])

		Index test (ε1)		
		Positive (n)	**Negative (n)**	
Reference standard	Positive (n)	33	16	**PPV (%)** = 67.3 95% CI: 52.5–80.1
	Negative (n)	9	44,206	**NPV (%)** = 100.0 95% CI: 100.0–100.0
		Sensitivity (%) = 78.6 95% CI: 63.2–89.7	**Specificity (%)** = 100.0 95% CI: 99.9–100.0	

G. Primary prostate cancer (Kappa value [95% CI]: 0.875 [0.755–0.995])

		Index test (δ2)		
		Positive (n)	**Negative (n)**	
Reference standard	Positive (n)	79	21	**PPV (%)** = 79.0 95% CI: 69.7–86.5
	Negative (n)	0	11,655	**NPV (%)** = 100.0 95% CI: 100.0–100.0
		Sensitivity (%) = 100.0 95% CI: 93.2–100.0	**Specificity (%)** = 99.8 95% CI: 99.7–99.9	

Figure 2. Diagnosis definitions with high* accuracy
CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value
*All accuracy values included for a definition are approximately 70% or more.

189x198mm (300 x 300 DPI)

A. Lung cancer (Kappa value [95% CI]: 0.982 [0.947–1.017])

		Index test (E1)			Index test (E4)		
		Positive (n)	Negative (n)		Positive (n)	Negative (n)	
Reference standard	Positive (n)	32	0	**PPV (%)** = 100.0 95% CI: 84.2–100.0	32	0	**PPV (%)** = 100.0 95% CI: 84.2–100.0
	Negative (n)	1	40	**NPV (%)** = 97.6 95% CI: 87.1–99.9	1	40	**NPV (%)** = 97.6 95% CI: 87.1–99.9
		Sensitivity (%) = 97.0 95% CI: 84.2–99.9	**Specificity (%)** = 100.0 95% CI: 87.1–100.0		**Sensitivity (%)** = 97.0 95% CI: 84.2–99.9	**Specificity (%)** = 100.0 95% CI: 87.1–100.0	

B. Breast cancer (Kappa value [95% CI]: 0.961 [0.917–1.005])

		Index test (E1)			Index test (E4)		
		Positive (n)	Negative (n)		Positive (n)	Negative (n)	
Reference standard	Positive (n)	1	0	**PPV (%)** = 100.0 95% CI: 1.3–100.0	1	0	**PPV (%)** = 100.0 95% CI: 1.3–100.0
	Negative (n)	0	104	**NPV (%)** = 100.0 95% CI: 94.8–100.0	0	104	**NPV (%)** = 100.0 95% CI: 94.8–100.0
		Sensitivity (%) = 100.0 95% CI: 1.3–100.0	**Specificity (%)** = 100.0 95% CI: 94.8–100.0		**Sensitivity (%)** = 100.0 95% CI: 1.3–100.0	**Specificity (%)** = 100.0 95% CI: 94.8–100.0	

C. Colorectal cancer (Kappa value [95% CI]: 0.953 [0.900-1.000])

		Index test (E1)			Index test (E4)		
		Positive (n)	Negative (n)		Positive (n)	Negative (n)	
Reference standard	Positive (n)	4	0	**PPV (%)** = 100.0 95% CI: 28.4–100.0	4	0	**PPV (%)** = 100.0 95% CI: 28.4–100.0
	Negative (n)	0	53	**NPV (%)** =100.0 95% CI: 90.1–100.0	0	53	**NPV (%)** = 100.0 95% CI: 90.1–100.0
		Sensitivity (%) = 100.0 95% CI: 28.4–100.0	**Specificity (%)** = 100.0 95% CI: 90.1–100.0		**Sensitivity (%)** = 100.0 95% CI: 28.4–100.0	**Specificity (%)** = 100.0 95% CI: 90.1–100.0	

D. Ovarian cancer (Kappa value [95% CI]: 0.940 [0.873–1.007])

		Index test (E1)			Index test (E4)		
		Positive (n)	Negative (n)		Positive (n)	Negative (n)	
Reference standard	Positive (n)	5	0	**PPV (%)** = 100.0 95% CI: 35.9–100.0	5	0	**PPV (%)** = 100.0 95% CI: 35.9–100.0
	Negative (n)	0	16	**NPV (%)** = 100.0 95% CI: 71.3–100.0	0	16	**NPV (%)** = 100.0 95% CI: 71.3–100.0
		Sensitivity (%) = 100.0 95% CI: 35.9–100.0	**Specificity (%)** = 100.0 95% CI: 71.3–100.0		**Sensitivity (%)** = 100.0 95% CI: 35.9–100.0	**Specificity (%)** = 100.0 95% CI: 71.3–100.0	

E. Bladder cancer (Kappa value [95% CI]: 0.878 [0.784-0.973])

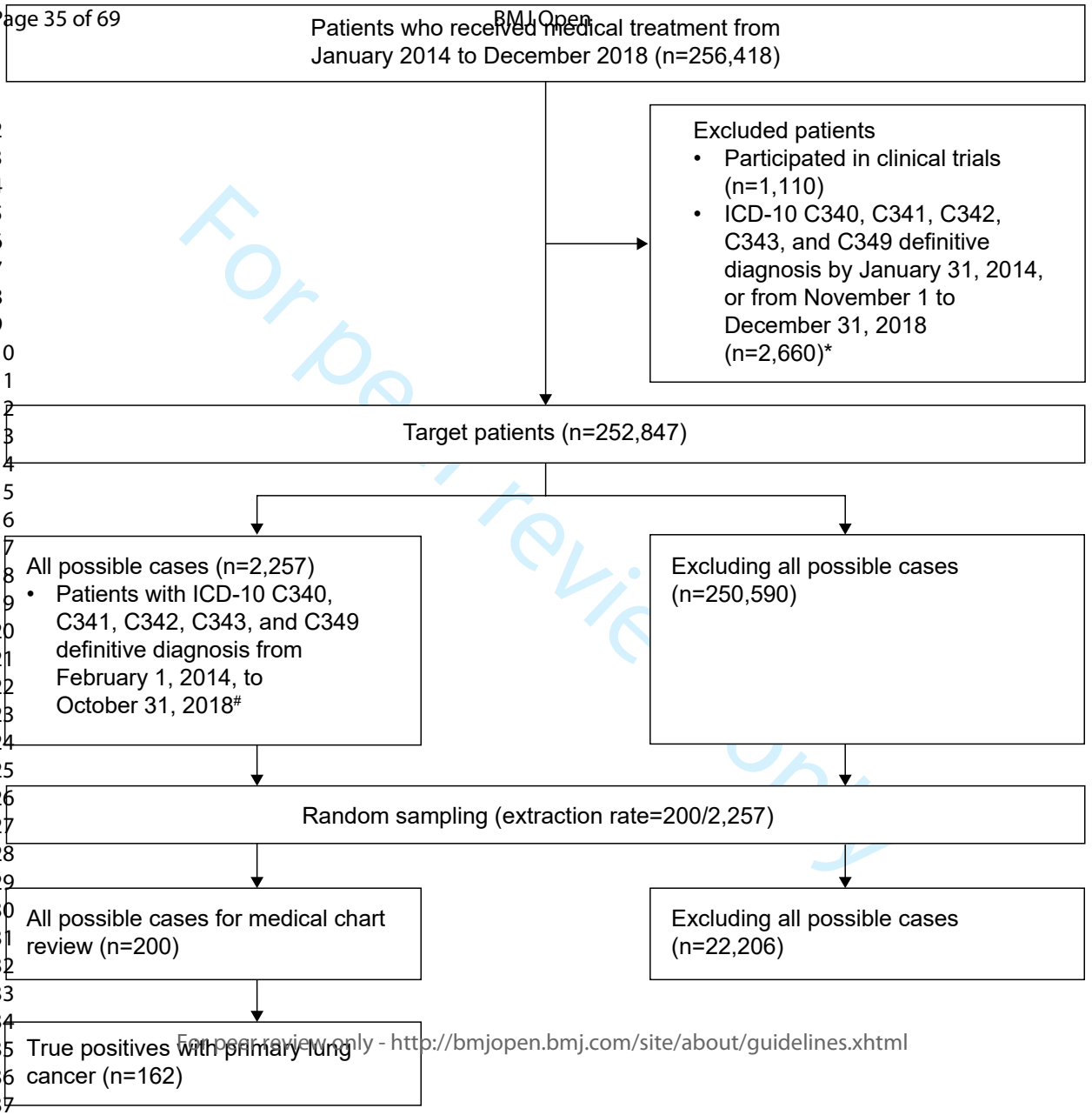
		Index test (E1)			Index test (E4)		
		Positive (n)	Negative (n)		Positive (n)	Negative (n)	
Reference standard	Positive (n)	2	0	**PPV (%)** = 100.0 95% CI: 9.4–100.0	2	0	**PPV (%)** = 100.0 95% CI: 9.4–100.0
	Negative (n)	0	8	**NPV (%)** = 100.0 95% CI: 51.8–100.0	0	8	**NPV (%)** = 100.0 95% CI: 51.8–100.0
		Sensitivity (%) = 100.0 95% CI: 9.4–100.0	**Specificity (%)** = 100.0 95% CI: 51.8–100.0		**Sensitivity (%)** = 100.0 95% CI: 9.4–100.0	**Specificity (%)** = 100.0 95% CI: 51.8–100.0	

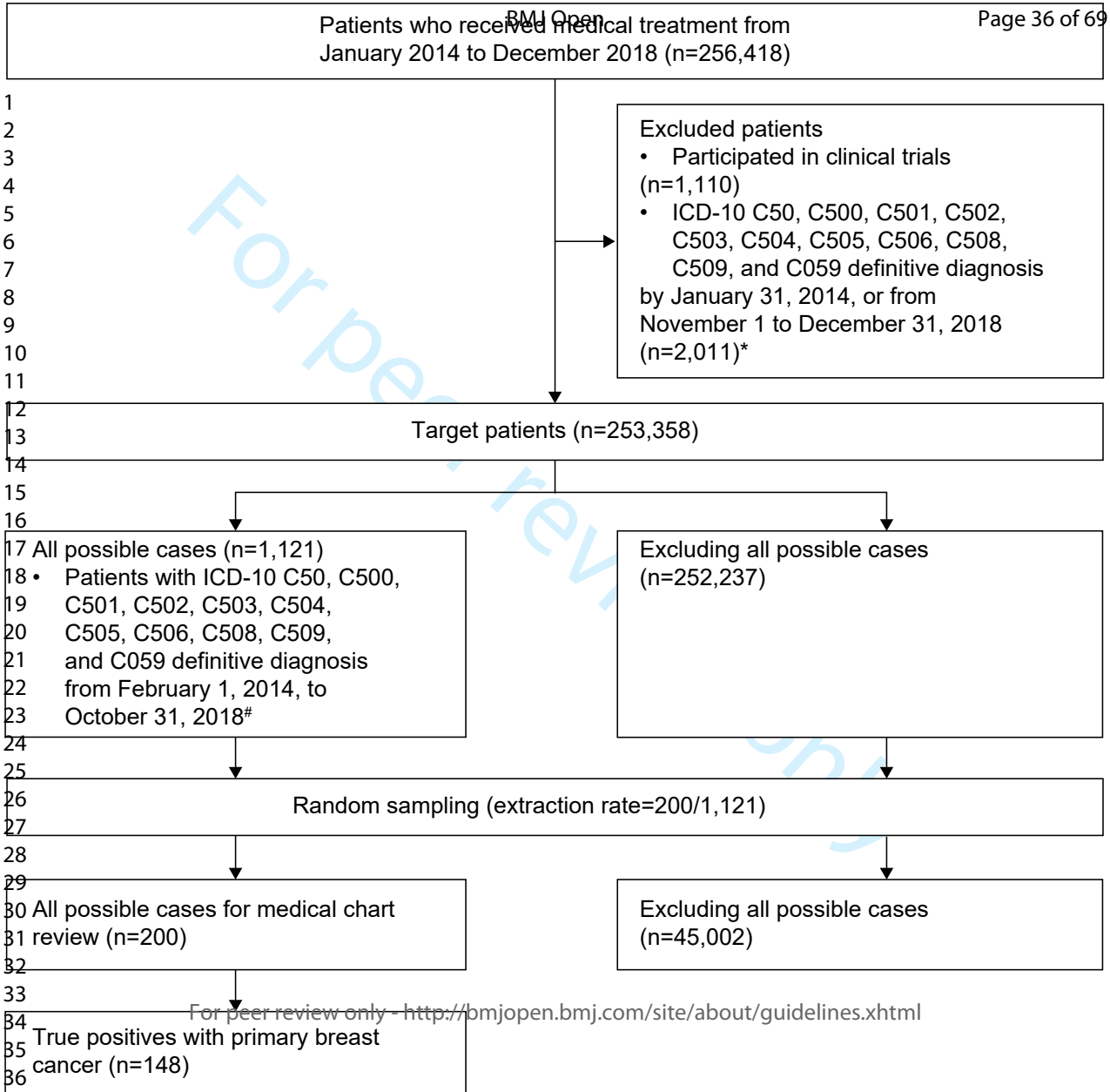
F. Prostate cancer (Kappa value [95% CI]: 0.905 [0.798–1.011])

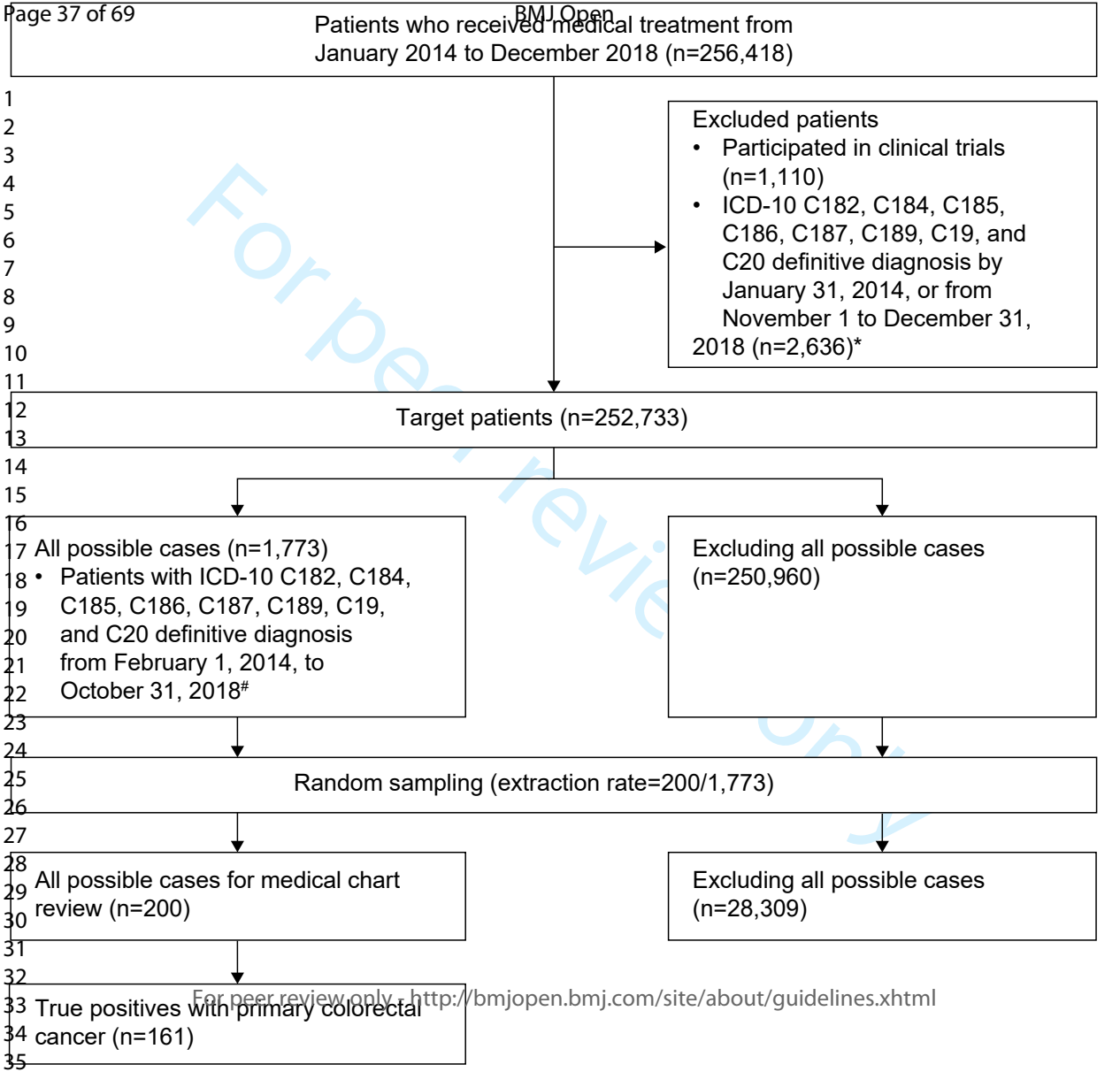
		Index test (E1)			Index test (E4)		
		Positive (n)	Negative (n)		Positive (n)	Negative (n)	
Reference standard	Positive (n)	3	0	**PPV (%)** = 100.0 95% CI: 19.4–100.0	3	0	**PPV (%)** = 100.0 95% CI: 19.4–100.0
	Negative (n)	1	32	**NPV (%) = 97.0** 95% CI: 84.2–99.9	1	32	**NPV (%) = 97.0** 95% CI: 84.2–99.9
		Sensitivity (%) = 75 95% CI: 19.4–99.4	**Specificity (%)** = 100.0 95% CI: 94.2–100.0		**Sensitivity (%)** = 75 95% CI: 19.4–99.4	**Specificity (%)** = 100.0 95% CI: 94.2–100.0	

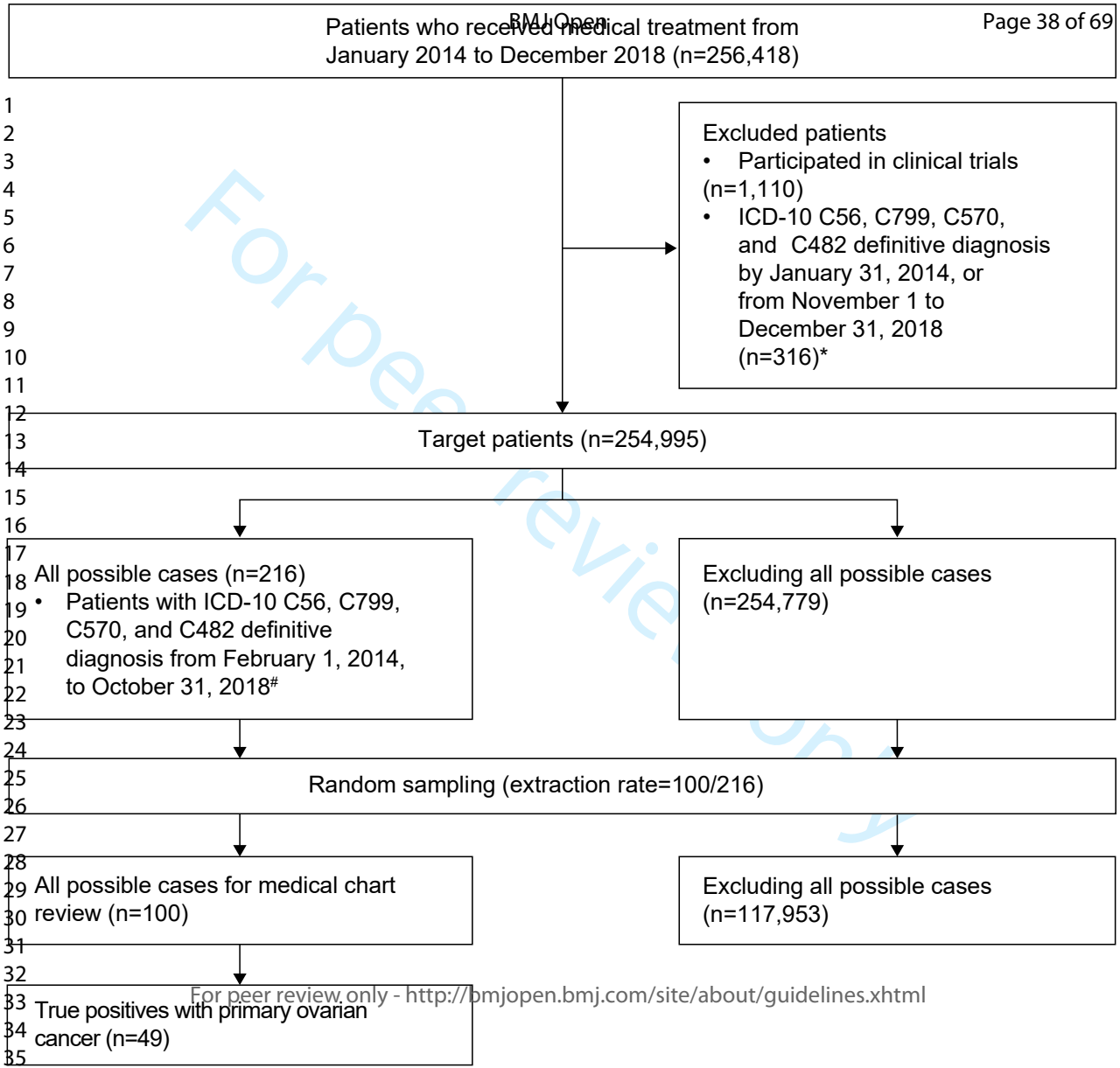
Figure 3. Death definitions with high* accuracy
CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value
*All accuracy values included for a definition are >70%.

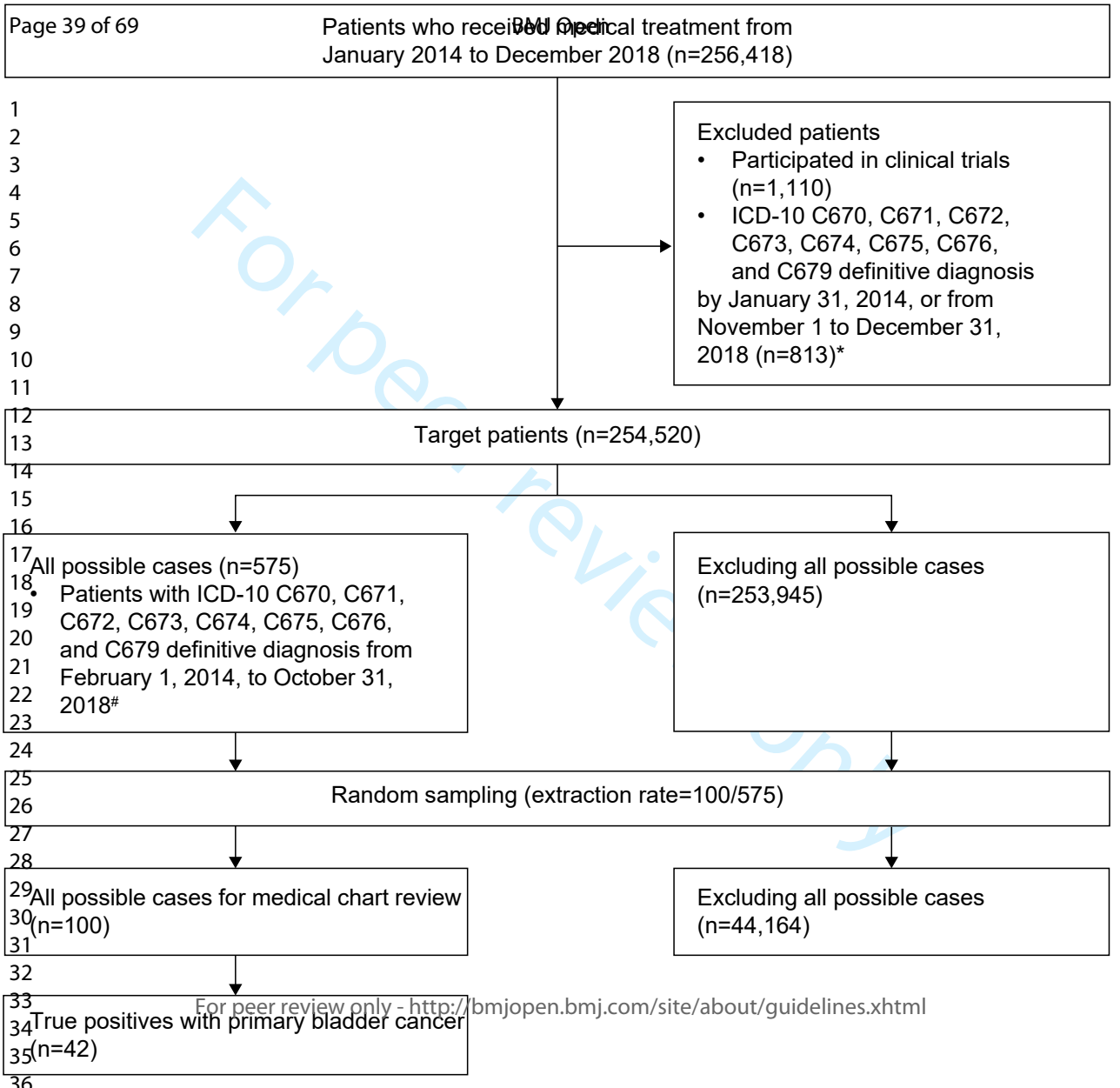
189x206mm (300 x 300 DPI)

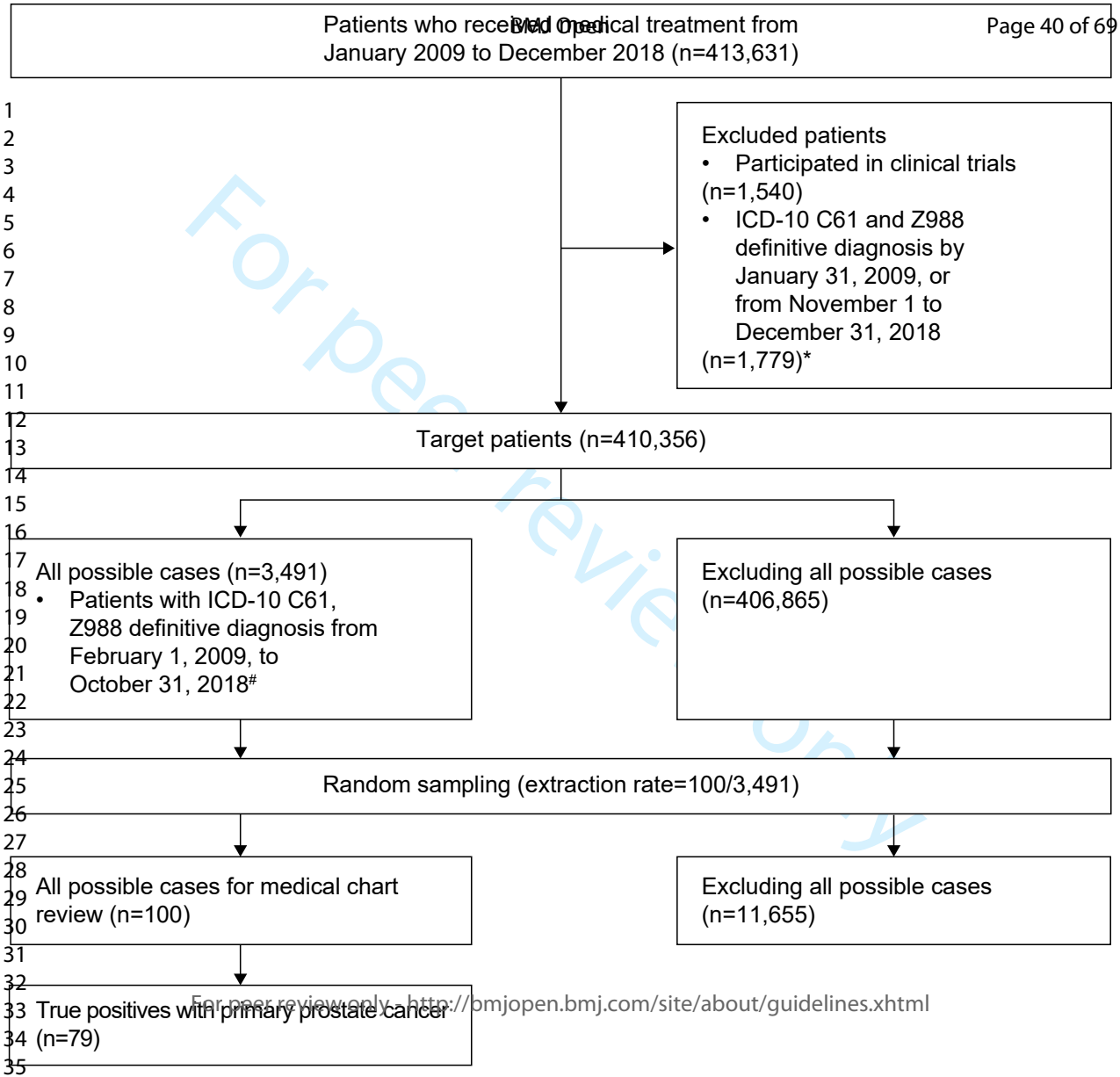












Supplemental Tables

Table S1. Inclusion criteria for lung, breast, colorectal, ovarian, bladder, and prostate cancer

Conventional classification	WHO classification	Patient criteria
True primary lung cancer in this study*		
Lung tumor	Tumors of the lung	
Epithelial tumor	Epithelial tumors	
Adenocarcinoma	Adenocarcinoma	Inclusion as non-small cell carcinoma (excluding atypical adenomatoid familial of pre-invasive lesions)
Squamous cell carcinoma	Squamous cell carcinoma	Inclusion as non-small cell carcinoma (excluding atypia of pre-invasive lesions)
Neuroendocrine tumors	Neuroendocrine tumors	
Small cell carcinoma	Small cell carcinoma	Inclusion as small cell cancer
Large cell neuroendocarcinoma	Large cell neuroendocrine carcinoma	Exclusion
Carcinoid tumor	Carcinoid tumors	Exclusion
Pre-invasive lesion	Preinvasive lesion	Exclusion
Large cell carcinoma	Large cell carcinoma	Inclusion as non-small cell carcinoma
Adenosquamous carcinoma	Adenosquamous carcinoma	Inclusion as non-small cell carcinoma
Sarcomatoid carcinoma	Sarcomatoid carcinoma	Inclusion as non-small cell carcinoma
Unclassified carcinoma	Other and unclassified carcinoma	Exclusion
Salivary gland type tumor	Salivary gland-type tumors	Exclusion
Papilloma	Papillomas	Exclusion
Adenoma	Adenomas	Exclusion
Mesenchymal tumor	Mesenchymal tumors	Exclusion
Lymphohistiocytic tumor	Lymphohistiocytic tumors	Exclusion
Tumors of ectopic origin	Tumors of ectopic origin	Exclusion

Lung metastases	Metastases to the lung	Exclusion
Pleural tumor	Tumor of the pleura	
Mesothelial tumor	Mesothelial tumors	Exclusion
Lymphoproliferative disorders	Lymphoproliferative disorders	Exclusion
Mesenchymal tumor	Mesenchymal tumors	Exclusion
True primary breast cancer in this study[#]		
Mammary gland tumor		
Epithelial tumor	Epithelial tumors	
Benign tumor	Benign tumors	Exclusion
Malignant tumor	Malignant tumors (carcinomas)	
Noninfiltrating carcinoma	Noninvasive carcinoma	Exclusion
Microinvasive carcinoma	Microinvasive carcinoma	Inclusion
Invasive carcinoma	Invasive breast carcinoma	Inclusion
Paget's disease	Paget's disease of the nipple	Exclusion
Mixed connective and epithelial tumors	Mixed connective tissue and epithelial tumors	Exclusion
Nonepithelial tumor	Nonepithelial tumors	Exclusion
Other	Others	Exclusion
So-called mammary gland disease	So-called mastopathy	Exclusion
Hamartoma	Hamartoma	Exclusion
Inflammatory lesions	Inflammatory lesion	Exclusion
Mammary fibrosis	Fibrous disease	Exclusion
Gynecomastia	Gynecomastia	Exclusion
Accessory milk	Accessory mammary gland	Exclusion
Metastatic tumors	Metastatic tumor	Exclusion
Other	Others	Exclusion
True primary colorectal cancer in this study[†]		
Benign epithelial tumor		Exclusion
Malignant epithelial tumor		
Adenocarcinoma (adenocarcinoma)		Inclusion

Adenosquamous carcinoma (adenosquamous carcinoma)		Inclusion
Squamous cell carcinoma (squamous carcinoma)		Inclusion
Carcinoid tumour (carcinoid tumor)		Exclusion
Endocrine carcinoma (endocrine cell carcinoma)		Exclusion
Miscellaneous (miscellaneous histological types of malignant epithelial tumors)		Exclusion
Nonepithelial tumor		Exclusion
Lymphoma (lymphoma)		Exclusion
Unclassifiable tumor		Exclusion
Metastatic tumors		Exclusion
Tumor-like lesions		Exclusion
Hereditary neoplasms and gastrointestinal polyposis		Exclusion
Appendix		Exclusion
Anal canal (including perianal skin)		Exclusion
True primary ovarian cancer in this study‡		
Ovarian tumor	Ovarian tumors	
Epithelial tumor	Epithelial tumors	
Serous tumor	Serous tumors	
Benign	Benign	Exclusion
Borderline malignancy	Borderline	Exclusion
Malignant	Malignant	Inclusion
Mucinous neoplasms	Mucinous tumors	
Benign	Benign	Exclusion
Borderline malignancy	Borderline	Exclusion
Malignant	Malignant	Inclusion

Endometrioid tumor	Endometrioid tumors	
Benign	Benign	Exclusion
Borderline malignancy	Borderline	Exclusion
Malignant	Malignant	Inclusion
Clear cell tumors	Clear cell tumors	
Benign	Benign	Exclusion
Borderline malignancy	Borderline	Exclusion
Malignant	Malignant	Inclusion
Brenner's tumor	Brenner tumors	
Benign	Benign	Exclusion
Borderline malignancy	Borderline	Exclusion
Malignant	Malignant	Inclusion
Seromucosal tumor	Seromucinous tumors	
Benign	Benign	Exclusion
Borderline malignancy	Borderline	Exclusion
Malignant	Malignant	Inclusion
Anaplastic Carcinoma	Undifferentiated carcinoma	Inclusion
Mesenchymal tumor	Mesenchymal tumors	Exclusion
Mixed epithelial mesenchymal tumor	Mixed epithelial and Mesenchymal tumors	Exclusion
Sex cord–stromal tumor	Sex cord–stromal tumors	Exclusion
Mixed sex cord–stromal tumor	Mixed sex cord–stromal tumors	Exclusion
Germ cell tumor	Germ cell tumors	Exclusion
Somatic tumors associated with monodermal teratomas and dermoid cysts	Monodermal teratoma and somatic-type tumors arising from dermoid cyst	Exclusion
Germ cell and policy stromal tumors	Germ cell-sex cord-stromal tumors	Exclusion
Other tumors	Miscellaneous tumors	Exclusion
Mesothelial tumor	Mesothelial tumors	Exclusion

Soft tissue	Soft tissue tumors	Exclusion
Neoplastic lesions	Tumor-like lesions	Exclusion
Lymphoid and myeloid neoplasms	Lymphoid and myeloid tumors	Exclusion
Secondary tumors	Secondary tumors	Exclusion
Tubal tumor	Tubal tumors	Inclusion
Peritoneal tumor	Peritoneal tumors	Inclusion
Epithelial tumor	Epithelial tumors	Inclusion*
Mesothelial tumor	Mesothelial tumors	Exclusion
Smooth muscle tumors	Smooth muscle tumors	Exclusion
Tumors of unknown origin	Tumors of uncertain origin	Exclusion
Other primary tumors	Miscellaneous primary tumors	Exclusion
Secondary tumors	Secondary tumors	Exclusion
True primary prostate cancer in this study[£]		
Malignant tumor		
Adenocarcinoma	Adenocarcinoma	Inclusion
Rare adenocarcinoma	Adenocarcinoma rare type	Inclusion
Urothelial carcinoma	Urothelial carcinoma	Inclusion
Squamous cell carcinoma	Squamous carcinoma	Inclusion
Adenosquamous carcinoma	Adenosquamous carcinoma	Inclusion
Basal cell carcinoma	Basal cell carcinoma	Inclusion
Small cell carcinoma	Small cell carcinoma	Inclusion
Anaplastic carcinoma	Undifferentiated carcinoma	Inclusion
Other malignant tumors	Other malignant tumors	
Sarcoma	Sarcoma	Exclusion
Metastatic tumors	Metastatic tumor	Exclusion
Tumor unclassifiable	Unclassified tumor	Exclusion
Borderline and associated lesions		Exclusion

True primary prostate cancer in this study		
Malignant tumor		
Adenocarcinoma	Adenocarcinoma	Inclusion
Rare adenocarcinoma	Adenocarcinoma rare type	Inclusion
Urothelial carcinoma	Urothelial carcinoma	Inclusion
Squamous cell carcinoma	Squamous carcinoma	Inclusion
Adenosquamous carcinoma	Adenosquamous carcinoma	Inclusion
Basal cell carcinoma	Basal cell carcinoma	Inclusion
Small cell carcinoma	Small cell carcinoma	Inclusion
Anaplastic carcinoma	Undifferentiated carcinoma	Inclusion
Other malignant tumors	Other malignant tumors	
Sarcoma	Sarcoma	Exclusion
Metastatic tumors	Metastatic tumor	Exclusion
Tumor unclassifiable	Unclassified tumor	Exclusion
Borderline and associated lesions		Exclusion
True primary bladder cancer in this study		
Bladder cancer		
Urothelial tumors		
Noninvasive flat urothelial carcinoma in situ (urothelial carcinoma in situ)		Exclusion
Papillary urothelial carcinoma in situ (noninvasive papillary urothelial carcinoma)		Exclusion
Invasive urothelial carcinoma (invasive urothelial carcinoma)		Inclusion
Squamous cell neoplasia		Inclusion
Glandular tumors		Inclusion
Tumors related to the ureteral membrane		Inclusion
Neuroendocrine tumors		Exclusion

Anaplastic carcinoma		Exclusion
Pigmented tumor		Exclusion
Mesenchymal tumor		Exclusion
Lymphohematopoietic tumors		Exclusion

*For true primary lung cancer, based on the classification tables (p70-73) of the 8th edition of the Clinical/Pathological Handling Code of the Japanese Lung Cancer Society (original publication 2016).

#For true primary breast cancer, based on the histological classification table (p24-25) of the 18th Edition of the Clinical and Pathological Handling Code of the Japanese Breast Cancer Society " (Gold Original Publication 2018) and the comparison table (P65-67) between the WHO classification and the handling conventional classification of the year of publication.

†For true primary colorectal cancers, based on the classification tables (p30-31) of the 9th edition of the Clinical/Pathological Handling Code (original publication 2018) of the Colon Cancer Study Group

‡For true primary ovarian cancers, based on the classification tables (p22-27) of the first edition of the Clinical and Pathological Handling Code (original publication 2016) of the Japanese Society of Obstetrics and Gynecology/Japanese Society of Pathology

£For true primary prostate cancer, based on the classification table (p.61) of the Japanese Society of Urological Sciences/Japan Society of Pathology/Japan Society of Medical Radiology, 4th edition of the Covenant on Clinical and Pathological Handling (Kanehara Publishing, 2010).

Table S2. Outcome definitions

Outcome	Definition	
A. Primary lung cancer	A1	• Definitive diagnosis of lung cancer (ICD-10: C340, C341, C342, C343, or C349) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.
	A2	• Definitive diagnosis of lung cancer (ICD-10: C340, C341, C342, C343, or C349) recorded between 2014 and 2018 in EMR data.
	A3	• Diagnosis of lung cancer (Japanese original diagnostic code: 1629003) recorded between 2014 and 2018 in EMR data.
	A4	• Definitions written in A1 and specimen examination for laboratory diagnosis (Japanese original procedural code: 160060170, 160060270, 160171470, 160185110, 160214310, 160209750, 160214710, 160214810, 160190270, 160190370, 160190470, 160190570, 160214470, 160214970, or 160062310) recorded between 2014 and 2018 in claims data.
B. Non-small cell lung cancer	B1	• Diagnosis of non-small cell lung cancer (Japanese original diagnostic code: 8847272, 8847732, 8849238, 8847598, 8847637, 8847664, or 8842053) recorded between 2014 and 2018 in EMR data.
	B2	• Diagnosis of non-small cell lung cancer (Japanese original diagnostic code: 8842835, 8847676, 8847677, 8847678, 8847679, 8835493, 8847634, 8847635, 8847636, 8847637, 8837666, 8847661, 8847662,

Outcome	Definition	
		8847663, 8847664, 8831458, 8847595, 8847596, 8847597, 8847598, 8833932, 1629003, 1629006, 1629009, 8838805, 8838844, 8838852, 8838898, 8838901, 8842053, 8842831, 8842832, 8842833, 8842834, 8847272, 8847732, 8849238, 8849788, or 2312002) recorded between 2014 and 2018 in EMR data.
C. Small cell lung cancer	C1	• Diagnosis of small cell lung cancer (Japanese original diagnostic code: 8847594, 8842185, 8847633, 8847660, or 8847675) recorded between 2014 and 2018 in EMR data.
α. Primary breast cancer	α1	• Definitive diagnosis of breast cancer (ICD-10: C500, 501, 502, 503, 504, 505, 506, 508, 509, or D059) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.
	α2	• Definitive diagnosis of breast cancer (ICD-10: C500, 501, 502, 503, 504, 505, 506, 508, 509, or D059) recorded between 2014 and 2018 in EMR data.
	α3	• Diagnosis of breast cancer (Japanese original diagnostic code: 8849699) recorded between 2014 and 2018 in EMR data.
β. Primary colorectal cancer	β1	• Definitive diagnosis of colorectal cancer (ICD-10: C179, C182, C184, C186, C187, C189, C19, or C20) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.

Outcome	Definition	
	β2	• Definitive diagnosis of colorectal cancer (ICD-10: C179, C182, C184, C186, C187, C189, C19, or C20) recorded between 2014 and 2018 in EMR data.
	β3	• Diagnosis of breast cancer (Japanese original diagnostic code: 8847015 or 8847916) recorded between 2014 and 2018 in EMR data.
γ. Primary ovarian cancer	γ1	• Definitive diagnosis of ovarian cancer (ICD-10: C56, C799, C570, or C482) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.
	γ2	• Definitive diagnosis of ovarian cancer (ICD-10: C56, C799, C570, or C482) recorded between 2014 and 2018 in EMR data.
ε. Primary bladder cancer	ε1	• Definitive diagnosis of bladder cancer (ICD-10: C670, C671, C672, C673, C674, C675, C676, or C679) recorded between 2014 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.
	ε2	• Definitive diagnosis of bladder cancer (ICD-10: C670, C671, C672, C673, C674, C675, C676, or C679) recorded between 2014 and 2018 in EMR data.
δ. Primary prostate cancer	δ1	• Definitive diagnosis of prostate cancer (ICD-10: C61 or Z988) recorded between 2009 and 2018 in DPC data. Primary diagnosis, admission-precipitating diagnosis, or most resource-consuming diagnosis.

Outcome	Definition	
	δ2	• Definitive diagnosis of prostate cancer (ICD-10: C61 or Z988) recorded between 2009 and 2018 in EMR data.
D. Performance status 2 or higher at the start of chemotherapy	D1	Medical treatment of rehabilitation for cancer patients (Japanese original diagnostic code: 180033110) recorded between 2014 and 2018 in claims data, given in the same index month as the prescription month of the therapeutic drug described in Table S3.
	D2	• Medical treatment of rehabilitation for disuse syndrome (Japanese original diagnostic code: H001-02, 180044610, 180044710, 180044810, 180044910, 180045010, 180045110, 180045210, 180045310, 180045410, 180045530, 180045630, 180045730, 180051530, 180051630, 180051730, 180051830, 180051930, 180052030, 180052130, 180052230, 180052330, 180052430, 180052530, or 180052630) recorded between 2014 and 2018 in claims data, given in the same index month as the prescription month of the therapeutic drug described in Table S3.
E. Death	E1	• Date of death in EMR data.
	E2	• Date of death in DPC data.
	E3	• Medical treatment of death for patients (Japanese original diagnostic code: 114007270, 114018670, or 114019970) recorded between 2014 and 2018 in claims data.
	E4	• 30 days before and after definitions written in E1.

Outcome	Definition	
	E5	• 30 days before and after definitions written in E2.
	E6	• 30 days before and after definitions written in E3.
F. First recurrence/progression	F1	• Date of disease name with "recurrence" as a modifier in Japanese original diagnostic code.
	F2	• Second specimen examination for laboratory diagnosis (Japanese original procedural code: 160060170, 160060270, 160171470, 160185110, 160214310, 160209750, 160214710, 160214810, 160190270, 160190370, 160190470, 160190570, 160214470, 160214970, or 160062310) recorded between 2014 and 2018 in claims data.
	F3	• Definitions written in F2 and patients with no history of surgery for the purpose of excision (with or without surgery for the purpose of examination).
	F4	• Month of definitions written in F1.
	F5	• Month of definitions written in F2.
	F6	• Month of definitions written in F3.
G. Second recurrence/progression	G1	• Date of administration of the drug described in Appendix 2 after definitions written in F1.
	G2	• Third specimen examination for laboratory diagnosis (Japanese original procedural code: 160060170, 160060270, 160171470, 160185110, 160214310, 160209750, 160214710, 160214810, 160190270, 160190370, 160190470, 160190570, 160214470, 160214970, or 160062310) recorded between 2014 and

Outcome	Definition	
		2018 in claims data.
	G3	• Month of definitions written in G1.
	G4	• Month of definitions written in G2.
H. Third recurrence/progression	H1	• Date of administration of the drug described in Appendix 2 after G1
	H2	• Forth specimen examination for laboratory diagnosis (Japanese original procedural code: 160060170, 160060270, 160171470, 160185110, 160214310, 160209750, 160214710, 160214810, 160190270, 160190370, 160190470, 160190570, 160214470, 160214970, or 160062310) recorded between 2014 and 2018 in claims data.
	H3	• Month of definitions written in H1.
	H4	• Month of definitions written in H2.
Adverse events		
I. Interstitial pneumonia	I1	• Definitive diagnosis of interstitial pneumonia (ICD-10: J702, J703, J704, J841 or J849) recorded in EMR data and Medical treatment (ATC code: H02AB04 or H02AB06 [excludes topical drugs]).
	I2	• Definitive diagnosis of interstitial pneumonia (ICD-10: J448, J700, J701, J702, J704, J82, J841, J849, or M0510) recorded in EMR data.
	I3	• Definitions written in I2 plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone

Outcome	Definition	
		(ATC code: H02AB06 with exception of external medicine) recorded in claims data.
J. Hepatic failure	J1	• Definitive diagnosis of hepatic failure (ICD-10: K720, K712, or K743) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	J2	• Laboratory data abnormality in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	J3	• Definitive diagnosis of hepatic failure (ICD-10: K710, K711, K712, K716, K717, K718, K719, K720, K729, K739, K740, K741, K743, K744, K745, K746, K750, K751, K752, K753, K754, K758, K759, K760, K761, K762, K763, K764, K765, K767, K768, K769, R18, R609, R945, or S361) recorded in EMR data.
	J4	• Definitions written in J3 plus prescription of medical treatment (ATC code: H02AB04, H02AB06, A05AA02, or A05BA08) recorded in claims data.
K. Colitis・diarrhea	K1	• Definitive diagnosis of colitis・diarrhea (ICD-10: A090 or A099) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.

Outcome	Definition	
	K2	• Definitive diagnosis of colitis・diarrhea (ICD-10: A099, K501, K509, K510, K512, K513, K515, K518, K519, K521, K522, K528, K529, K550, K551, K552, K559, K566, K591, K628, K638, K921, K922, M321, or R101) recorded in EMR data.
	K3	• Definitions written in K2 plus prescription of medical treatment (ATC codes: H02AB04, H02AB06, A07A, A07F, A07E, A07D, or A07X) recorded in claims data.
L. Type 1 diabetes	L1	• Prescription of medical treatment (ATC code: A10AB, A10AC, A10AD, or A10AE)
	L2	• Definitive diagnosis of type 1 diabetes (ICD-10: E10, E100, E101, E102, E103, E104, E105, or E106) recorded in EMR data.
M. Encephalitis・meningitis	M1	• Definitive diagnosis of encephalitis・meningitis (ICD-10: G040, G048, G049, or G934) recorded in EMR data.
	M2	• Definitive diagnosis of encephalitis・meningitis (ICD-10: G040, G048, G049, or G934) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	M3	• Definitive diagnosis of encephalitis. • Meningitis (ICD-10: R291) recorded in EMR data.
	M4	• Definitions written in M3 plus prescription Meningitis (ICD-10: R291) recorded in EMR data of medical

Outcome	Definition	
		treatment (ATC code: J05AB, J01, or J02A) recorded in claims data.
N. Nerve disorder (excludes paresthesia)	N1	• Definitive diagnosis of nerve disorder (excludes paresthesia) (ICD-10: G500, G501, G508, G509, G511, G512, G513, G514, G518, G519, G520, G521, G522, G523, G527, G528, G529, G540, G541, G542, G543, G544, G545, G560, G561, G562, G563, G564, G568, G569, G570, G571, G572, G573, G574, G575, G576, G579, G580, G587, G588, G589, G603, G608, G609, G618, G620, G622, G629, G64, G723, G810, G811, G819, G820, G821, G822, G823, G824, G825, G830, G831, G832, G833, G839, G900, G902, G903, G904, G908, G909, H812, H919, H933, M7924, M7926, M7929, M8900, M998, R252, R253, or R258) recorded in EMR data.
	N2	• Definitions written in N1 and medical treatment (ATC code H02AB04 or H02AB06) recorded in claims data.
O. Myasthenia gravis	O1	• Definitive diagnosis of myasthenia gravis (ICD-10: G700) recorded in EMR data.
	O2	• Definitive diagnosis of myasthenia gravis (ICD-10: G700) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	O3	• Definitive diagnosis of myasthenia gravis (ICD-10: G700, G701, G709) recorded in EMR data.
	O4	• Definitions written in O3 and medical treatment (ATC code: H02AB04, H02AB06, or H07AA02)

Outcome	Definition	
		recorded in claims data.
P. Guillain-Barré syndrome	P1	• Definitive diagnosis of Guillain-Barré syndrome (ICD-10: G610) recorded in EMR data.
	P2	• Definitions written in P1 plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	P3	• Definitions written in P1 plus prescription of methylprednisolone (ATC code: H02AB04), prednisolone (ATC code: H02AB06 with exception of external medicine), or immunoglobulin recorded in claims data.
	P4	• Definitions written in P1 and medical treatment (ATC code: H02AB04, H02AB06, J06BA, J06BB, or J06BC) recorded in claims data.
Q. Skin disorders	Q1	• Definitive diagnosis of skin disorders (ICD-10: H605, H738, I831, L00, L010, L011, L020, L021, L022, L023, L024, L028, L029, L030, L031, L032, L033, L038, L039, L080, L081, L089, L100, L101, L102, L103, L104, L105, L108, L109, L110, L111, L119, L120, L121, L123, L129, L130, L131, L138, L139, L200, L208, L210, L219, L233, L238, L239, L26, L270, L271, L279, L280, L281, L282, L290, L291, L292, L298, L299, L300, L301, L302, L303, L304, L305, L309, L400, L401, L402, L403, L404, L408, L409, L410, L411, L413, L414, L415, L418, L419, L42, L430, L431, L433, L438, L439, L440, L441, L442, L443, L449, L500, L501, L502, L504, L508, L509, L510, L511, L512, L518, L519, L52, L530, L531, L532, L538, L539, L560, L561, L562, L563, L564, L568, L570, L571, L572, L574, L578, L580,

Outcome	Definition	
		L589, L590, L598, L700, L701, L702, L703, L708, L709, L710, L711, L718, L719, L730, L731, L738, L739, L80, L810, L811, L812, L813, L814, L816, L817, L818, L819, L82, L83, L850, L851, L852, L853, L858, L859, L870, L871, L872, L879, L88, L890, L891, L892, L893, L899, L900, L906, L908, L909, L919, L920, L921, L928, L929, L930, L931, L932, L940, L941, L942, L943, L944, L945, L946, L950, L951, L97, L980, L981, L982, L983, L984, L985, L986, L988, R02, R21, R238, or T783) recorded in EMR data.
	Q2	• Definitions written in Q1 and medical treatment (ATC codes: H02AB04, H02AB06, D04AA, or R01AC [excludes steroidal drugs]) recorded in claims data.
R. Rhabdomyolysis	R1	• "Drug-induced rhabdomyolysis" or "rhabdomyolysis" in definitive diagnosis of rhabdomyolysis (ICD-10: M6289) recorded in EMR data.
	R2	• Definitive diagnosis of rhabdomyolysis (ICD-10: D868, G718, G720, G722, G724, G729, M331, M332, M339, M353, M358, M6019, M6091, M6092, M6095, M6098, M6099, M6105, M6109, M6119, M6129, M6155, M6159, M6289, M7900, M7910, M7911, M7912, M7913, M7915, M7916, M7918, M7919, or M7979) recorded in EMR data.
	R3	• Definitions written in R2 plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.

Outcome	Definition	
S. Myocarditis	S1	• Definitive diagnosis of myocarditis (ICD-10: I401, I408, I409, I514) recorded in EMR data.
	S2	• Definitive diagnosis of myocarditis (ICD-10: I401, I408, I409, I514) recorded in EMR data plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.
	S3	• Definitive diagnosis of myocarditis (ICD-10: D868, E854, E888, E889, I010, I011, I012, I018, I019, I050, I051, I052, I058, I059, I060, I061, I062, I069, I070, I071, I078, I079, I080, I081, I082, I083, I088, I089, I090, I091, I092, I099, I200, I201, I208, I209, I210, I212, I213, I214, I219, I220, I221, I228, I229, I230, I231, I232, I233, I234, I235, I236, I238, I240, I248, I249, I251, I252, I253, I254, I255, I256, I258, I259, I300, I308, I309, I319, I339, I340, I341, I348, I350, I351, I352, I358, I359, I360, I361, I362, I369, I370, I371, I372, I379, I38, I401, I408, I409, I420, I421, I422, I423, I424, I425, I426, I427, I428, I429, I440, I441, I442, I443, I444, I445, I446, I447, I451, I452, I453, I454, I455, I456, I458, I459, I460, I461, I469, I470, I471, I472, I479, I480, I481, I482, I489, I490, I491, I492, I493, I494, I495, I498, I499, I500, I501, I509, I513, I514, I515, I518, I519, R000, R001, R008, R570, R571, R579, or R943) recorded in EMR data.
	S4	• Definitions written in S3 plus prescription of methylprednisolone (ATC code: H02AB04) or prednisolone (ATC code: H02AB06 with exception of external medicine) recorded in claims data.

Outcome	Definition	
T. Gastrointestinal perforation	T1	• Definitive diagnosis of gastrointestinal perforation (ICD-10: K255, K265, K631, K65S, or K639) recorded in EMR data.
U. Adrenal insufficiency	U1	• Definitive diagnosis of adrenal insufficiency in Japanese original diagnostic code including the words "autoimmune adrenitis" recorded in claims data and "hypoadrenocorticism" plus medical treatment (ATC: code H02AB09) recorded in claims data.
	U2	• Definitive diagnosis of adrenal insufficiency (ICD-10: E271, E272, E273, E274, E275 or E278) recorded in EMR data.
	U3	• Definitions written in U2 plus medical treatment (ATC code H02AB09) recorded in claims data.
X. Febrile neutropenia	X1	• Definitive diagnosis of febrile neutropenia (ICD-10: D70) recorded in EMR data and medical treatment (Table S3) recorded in claims data.

ATC, Anatomical Therapeutic Chemical; DPC, Diagnosis Procedure Combination; EMR, electronic medical record; ICD-10, ICD-10, International Classification of Diseases, 10th revision

Table S3. Drug codes

ATC code	Common name
L01XC32	Atezolizumab
L01XC17	Nivolumab
L01XC18	Pembrolizumab
L01XC31	Avelumab
L01XC28	Durvalumab
L01XC06	Cetuximab
L01XC08	Panitumumab
L01XE02	Gefitinib
L01XE35	Osimertinib
L01XE47	Dacomitinib
L01XE13	Afatinib
L01XE03	Erlotinib
L01XE36	Alectinib
L01XE44	Lorlatinib
L01XE28	Ceritinib
L01XE16	Crizotinib
L01XC07	Bevacizumab (includes related biosimilars)
L01XC13	Pertuzumab
L01XC14	Trastuzumab emtansine
L01XE07	Lapatinib
L01XE33	Palbociclib
L01XE50	Abemaciclib
L01XE10, L04AA18	Everolimus
L01XX46	Olaparib
L01XC08	Panitumumab
L01XE21	Regorafenib
L01	Anti-malignant tumor drugs excluding talaporfin sodium (620001918), porfimer sodium (620007468), anagrelide hydrochloride hydrate (622379001), and sterile talc (622293901)
L02	Hormone therapy

ATC code	**Common name**
L04	Immunosuppressive drug
J01CR05	Tazobactam and piperacillin
J01DD02	Ceftazidime hydrate
J01DE03	Cefozopran hydrochloride
J01DE01	Cefepime dihydrochloride hydrate
J01DE02	Cefpirome sulfate
J01DH05	Biapenem
J01DH02	Meropenem hydrate
J01DH51	Imipenem hydrate, cilastatin sodium
J01DH04	Doripenem hydrate
J01DH55	Panipenem and betamipron

Table S4. Accuracy of diagnosis definitions

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Lung cancer								
Primary lung cancer								
A1	132	7	22,237	30	81.5 (74.6–87.1)	100.0 (99.9–100.0)	95.0 (89.9–98.0)	99.9 (99.8–99.9)
A2	162	38	22,206	0	100.0 (96.6–100.0)	99.8 (99.8–99.9)	81.0 (74.9–86.2)	100.0 (100.0–100.0)
A3	19	1	22,243	143	11.7 (7.2–17.7)	100.0 (100.0–100.0)	95.0 (75.1–99.9)	99.4 (99.2–99.5)
A4	128	7	22, 237	34	79.0 (71.8–85.0)	100.0 (99.9–100)	94.8 (89.6–97.9)	99.8 (99.8–99.9)
Non-small cell lung cancer								
B1	46	6	22,280	74	38.3 (29.6–47.6)	100.0 (99.9–100.0)	88.5 (76.6–95.6)	99.7 (99.6–99.7)
B2	46	6	22,280	74	38.3 (29.6–47.6)	100.0 (99.9–100.0)	88.5 (76.6–95.6)	99.7 (99.6–99.7)
Small cell lung cancer								
C1	10	0	22,395	1	90.9 (58.7–99.8)	100.0 (100.0–100.0)	100.0 (58.7–100.0)	100.0 (100.0–100.0)
Breast cancer								
Primary breast cancer								
α1	93	18	45,036	55	62.8 (54.5–70.6)	100.0 (99.9–100.0)	83.8 (75.6–90.1)	99.9 (99.8–99.9)
α2	148	52	45,002	0	100.0 (96.3–100.0)	99.9 (99.8–99.9)	74.0 (67.3–79.9)	100.0 (100.0–100.0)
α3	0	0	45,054	148	0.0 (0.0–3.7)	100.0 (100.0–100.0)	NA	99.7 (99.6–99.7)
Colorectal cancer								
Primary colorectal cancer								

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
β1	108	8	28,340	53	67.1 (59.2–74.3)	100.0 (99.9–100.0)	93.1 (86.9–97.0)	99.8 (99.8–99.9)
β2	161	39	28,309	0	100.0 (96.6–100.0)	99.9 (99.8–99.9)	80.5 (74.3–85.8)	100.0 (100.0–100.0)
β3	0	0	28,348	161	0.0 (0.0–3.4)	100.0 (100.0–100.0)	NA	99.4 (99.3–99.5)
Ovarian cancer								
Primary ovarian cancer								
γ1	44	14	11,692	5	89.8 (77.8–96.6)	99.9 (99.8–99.9)	75.9 (62.8–86.1)	100.0 (99.7–100.0)
γ2	49	50	11,656	0	100.0 (89.4–100.0)	99.6 (99.4–99.7)	49.5 (39.3–59.7)	100.0 (100.0–100.0)
Bladder cancer								
Primary bladder cancer								
ε1	33	16	44,206	9	78.6 (63.2–89.7)	100.0 (99.9–100.0)	67.3 (52.5–80.1)	100.0 (100.0–100.0)
ε2	42	58	44,164	0	100.0 (87.7–100.0)	99.9 (99.8–99.9)	42.0 (32.2–52.3)	99.9 (99.8–99.9)
Prostate cancer								
Primary prostate cancer								
δ1	17	0	11,676	62	21.5 (12.1–32.2)	100.0 (100.0–100.0)	100.0 (72.7–100.0)	99.5 (99.3–99.6)
δ2	79	21	11,655	0	100.0 (93.2–100.0)	99.8 (99.7–99.9)	79.0 (69.7–86.5)	100.0 (100.0–100.0)

CI, confidence interval; NA, not available; NPV, negative predictive value; PPV, positive predictive value

Table S5. Accuracy of death definitions

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Lung cancer								
E1	32	0	40	1	97.0 (84.2–99.9)	100.0 (87.1–100.0)	100.0 (84.2–100.0)	97.6 (87.1–99.9)
E2	9	0	40	24	27.3 (13.3–45.5)	100.0 (87.1–100.0)	100.0 (55.5–100.0)	62.5 (49.5–74.3)
E3	0	0	40	33	0.0 (0.0–15.3)	100.0 (87.1–100.0)	NA	54.8 (4.7–66.5)
E4	32	0	40	1	97.0 (84.2–99.9)	100.0 (87.1–100.0)	100.0 (84.2–100.0)	97.6 (87.1–99.9)
E5	9	0	40	24	27.3 (13.3–45.5)	100.0 (87.1–100.0)	100.0 (55.5–100.0)	62.5 (49.5–74.3)
E6	0	0	40	33	0.0 (0.0–15.3)	100.0 (87.1–100.0)	NA	54.8 (4.7–66.5)
Breast cancer								
E1	1	0	104	0	100.0 (1.3–100.0)	100.0 (94.8–100.0)	100.0 (1.3–100.0)	100.0 (94.8–100.0)
E2	0	0	104	1	0.0 (0.0–98.7)	100.0 (94.8–100.0)	NA	99.0 (94.8–100.0)
E3	0	0	104	1	0.0 (0.0–98.7)	100.0 (94.8–100.0)	NA	99.0 (94.8–100.0)
E4	1	0	104	0	100.0 (1.3–100.0)	100.0 (94.8–100.0)	100.0% (1.3–100.0)	100.0 (94.8–100.0)
E5	0	0	104	1	0.0 (0.0–98.7)	100.0 (94.8–100.0)	NA	99.0 (94.8–100.0)
E6	0	0	104	1	0.0 (0.0–98.7)	100.0 (94.8–100.0)	NA	99.0 (94.8–100.0)
Colorectal cancer								
E1	4	0	53	0	100.0 (28.4–100.0)	100.0 (90.1–100.0)	100.0 (28.4–100.0)	100.0 (90.1–100.0)
E2	2	0	53	2	50.0 (6.8–93.2)	100.0 (90.1–100.0)	100.0 (9.4–100.0)	96.4 (87.5–99.6)

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
E3	0	0	53	4	0.0 (0.0–71.6)	100.0 (90.1–100.0)	NA	93.0 (83.0–98.1)
E4	4	0	53	0	100.0 (28.4–100.0)	100.0 (90.1–100.0)	100.0 (28.4–100.0)	100.0 (90.1–100.0)
E5	2	0	53	2	50.0 (6.8–93.2)	100.0 (90.1–100.0)	100.0 (9.4–100.0)	96.4 (87.5–99.6)
E6	0	0	53	4	0.0 (0.0–71.6)	100.0 (90.1–100.0)	NA	93.0 (83.0–98.1)
Ovarian cancer								
E1	5	0	16	0	100.0 (35.9–100.0)	100.0 (71.3–100.0)	100.0 (35.9-100.0)	100.0 (71.3–100.0)
E2	2	0	16	3	40.0 (5.3–85.3)	100.0 (71.3–100.0)	100.0 (9.4–100.0)	84.2 (60.4–96.6)
E3	0	0	16	5	0.0 (0.0–64.1)	100.0 (71.3–100.0)	NA	76.2 (52.8–91.8)
E4	5	0	16	0	100.0 (35.9–100.0)	100.0 (71.3–100.0)	100.0 (35.9-100.0)	100.0 (71.3–100.0)
E5	2	0	16	3	40.0 (5.3–85.3)	100.0 (71.3–100.0)	100.0 (9.4–100.0)	84.2 (60.4–96.6)
E6	0	0	16	5	0.0 (0.0–64.1)	100.0 (71.3–100.0)	NA	76.2 (52.8–91.8)
Bladder cancer								
E1	2	0	8	0	100.0 (9.4–100.0)	100.0 (51.8–100.0)	100.0 (9.4–100.0)	100.0 (51.8–100.0)
E2	1	0	8	1	50.0 (1.3–98.7)	100.0 (51.8–100.0)	100.0 (51.8–100.0)	100.0 (1.3–100.0)
E3	0	0	8	2	0.0 (0.0–90.6)	100.0 (51.8–100.0)	NA	80.0 (44.4–97.5)
E4	2	0	8	0	100.0 (9.4–100.0)	100.0 (51.8–100.0)	100.0 (9.4–100.0)	100.0 (51.8–100.0)
E5	1	0	8	1	50.0 (1.3–98.7)	100.0 (51.8–100.0)	100.0 (51.8–100.0)	100.0 (1.3–100.0)

Outcome definition	True positives, n	False positives, n	True negatives, n	False negatives, n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
E6	0	0	8	2	0.0% (0.0–90.6)	100.0 (51.8–100.0)	NA	80.0 (44.4–97.5)
Prostate cancer								
E1	3	0	32	1	75.0 (19.4–99.4)	100.0 (94.2–100.0)	100.0 (19.4–100.0)	97.0 (84.2–99.9)
E2	0	0	32	4	0.0 (0.0–71.6)	100.0 (84.2–100.0)	NA	88.9 (73.9–96.9)
E3	0	0	32	4	0.0 (0.0–71.6)	100.0 (84.2–100.0)	NA	88.9 (73.9–96.9)
E4	3	0	32	1	75.0 (19.4–99.4)	100.0 (94.2–100.0)	100.0 (19.4–100.0)	97.0 (84.2–99.9)
E5	0	0	32	4	0.0 (0.0–71.6)	100.0 (84.2–100.0)	NA	88.9 (73.9–96.9)
E6	0	0	32	4	0.0 (0.0–71.6)	100.0 (84.2–100.0)	NA	88.9 (73.9–96.9)

CI, confidence interval; NA, not available; NPV, negative predictive value; PPV, positive predictive value

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	3
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	3
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	5
	4	Study objectives and hypotheses	5 and 6
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	6
Participants	**6**	Eligibility criteria	8
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	8
	8	Where and when potentially eligible participants were identified (setting, location and dates)	8
	9	Whether participants formed a consecutive, random or convenience series	
Test methods	**10a**	Index test, in sufficient detail to allow replication	9
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	5
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	10
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	10
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	8
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	8
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	10-11
	15	How indeterminate index test or reference standard results were handled	10-11
	16	How missing data on the index test and reference standard were handled	10-11
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	Not applicable
	18	Intended sample size and how it was determined	Page 9
RESULTS			
Participants	**19**	Flow of participants, using a diagram	Supplementary figures 1- 6
	20	Baseline demographic and clinical characteristics of participants	Table 4
	21a	Distribution of severity of disease in those with the target condition	Not applicable
	21b	Distribution of alternative diagnoses in those without the target condition	Not applicable
	22	Time interval and any clinical interventions between index test and reference standard	
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Table 2, Table 3, Table S3, Table S4
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	Along with each result in corresponding tables
	25	Any adverse events from performing the index test or the reference standard	Not applicable
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	Page 25
	27	Implications for practice, including the intended use and clinical role of the index test	Page 26
OTHER INFORMATION			

	28	Registration number and name of registry	Page 8
	29	Where the full study protocol can be accessed	No
	30	Sources of funding and other support; role of funders	Page 26

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.