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

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

To cite: Fujiwara T, Kanemitsu T, Tajima K, *et al*. 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. *BMJ Open* 2022;**12**:e055459. doi:10.1136/bmjopen-2021-055459

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-055459).

Received 15 July 2021
Accepted 31 May 2022

For numbered affiliations see end of article.

Correspondence to
Dr Hironobu Tokumasu; tokumasu@rwdata.co.jp

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 electronic 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% CIs were calculated.

Participants The validation cohort included patients with lung (n=2257), breast (n=1121), colorectal (n=1773), ovarian (n=216) and bladder (n=575) cancer who visited the hospital between January 2014 and December 2018, and those with prostate cancer (n=3491) 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 to 100.0) and 81.0% (74.9 to 86.2); breast, 100.0% (96.3 to 100.0) and 74.0% (67.3 to 79.9); colorectal, 100.0% (96.6 to 100.0) and 80.5% (74.3 to 85.8); ovarian, 89.8% (77.8 to 96.6) and 75.9% (62.8 to 86.1); bladder, 78.6% (63.2 to 89.7) and 67.3% (52.5 to 0.1); prostate, 100.0% (93.2 to 100.0) and 79.0% (69.7 to 86.5). Sensitivity and PPV for death were as follows: lung, 97.0% (84.2 to 99.9) and 100.0% (84.2 to 100.0); breast, 100.0% (1.3 to 100.0) and 100.0% (1.3 to 100.0); colorectal, 100.0% (28.4 to 100.0) and 100.0% (28.4 to 100.0); ovarian, 100.0% (35.9 to 100.0) and 100.0% (35.9 to 100.0); bladder, 100.0% (9.4–100.0) and 100.0% (9.4 to 100.0); prostate, 75.0% (19.4 to 99.4) and 100.0% (19.4 to 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.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ To our knowledge, this is the first study in oncology in Japan that validates disease and adverse event (AE) 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 generalisability 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 AE 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.

Trial registration number University Hospital Medical Information Network (UMIN) Clinical Trials Registry (UMIN000039345).

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 postmarketing surveillance to investigate the safety and efficacy of pharmaceutical products is being recognised 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 (online supplemental 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 anonymised and sent to Real World Data Co, the vendor for HCEI. The verified dataset was linked deterministically to claims data and EMRs originally derived from the hospital.

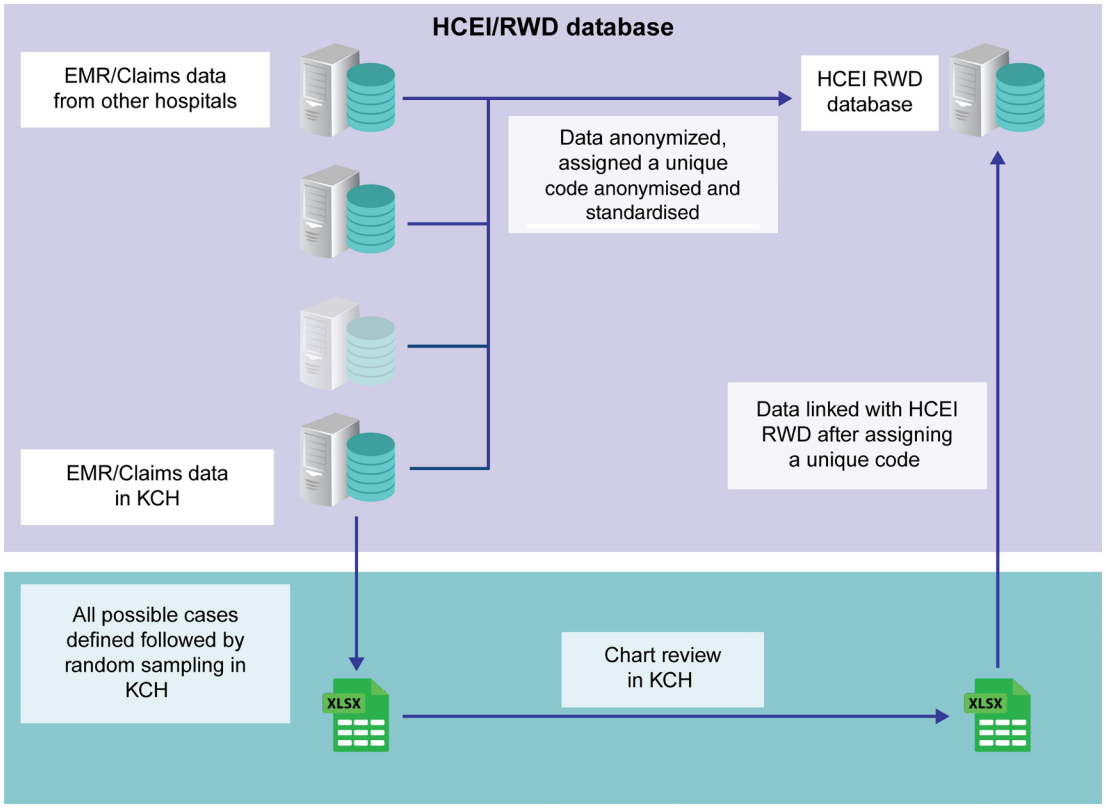

Figure 1 Health, Clinic, and Education Information Evaluation Institute/real-world database. EMR, electronic medical record; HCEI, Health, Clinic, and Education Information Evaluation Institute; KCH, Kurashiki Central Hospital; RWD, real-world database.

Fujiwara T, *et al*. *BMJ Open* 2022;**12**:e055459. doi:10.1136/bmjopen-2021-055459

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 (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, anonymised, linked to a unique code and standardised (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 36 37]

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 (online supplemental figures S1–S5), and those with prostate cancer (online supplemental 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 online supplemental 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 1 January 2014 to 31 January 2014 and from 1 November 2018 to 31 December 2018, and that for prostate cancer from 1 January 2009 to 31 January 2009 and from 1 November 2018 to 31 December 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 31 December 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 31 December 2013, and that for prostate cancer before 31 December 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 1 January 2014 and 31 December 2018; and those with prostate cancer between 1 January 2009 and 31 December 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/diarrhoea, 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 (online supplemental 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% 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 ICD-10, 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 (online supplemental table S2). Statistical analyses were conducted using R V.4.0.2 software.

RESULTS

Patient disposition

Of the 256 418 patients who received medical treatment from 2014 to 2018, 2257 with lung cancer (online supplemental figure S1), 1121 with breast cancer (online supplemental figure S2), 1773 with colorectal cancer (online supplemental figure S3), 216 with ovarian cancer (online supplemental figure S4) and 575 with bladder cancer (online supplemental figure S5) were included as all possible cases (table 1). From 2009 to 2018, 3491 patients with prostate cancer of 413 631 patients receiving medical treatment (online supplemental figure S6) were included as all possible cases (table 1).

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 (online supplemental table S2).

Lung cancer

The kappa value in chart reviews for diagnosis definitions was 0.982 (95% CI 0.947 to 1.017) for primary lung cancer, 0.979 (95% CI 0.950 to 1.008) for non-small cell lung cancer (NSCLC), 1.00 for small cell lung cancer (SCLC) and 0.982 (95% CI 0.947 to 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%; online supplemental table S4).

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	2257	162
Breast cancer	January 2014 to December 2018	256 418	253 358	1121	148
Colorectal cancer	January 2014 to December 2018	256 418	252 733	1773	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	3491	79

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

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

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

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

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

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

C. Primary breast cancer (kappa value [95% CI]: 1.000 [1.000 to 1.000])

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

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

		Reference standard: Positive (n)	Reference standard: Negative (n)	
Index test (β2)	Positive (n)	161	39	**PPV (%)** = 80.5 95% CI: 74.3 to 85.8
	Negative (n)	0	28 309	**NPV (%)** = 100.0 95% CI: 100.0 to 100.0
		Sensitivity (%) = 100.0 95% CI: 96.6 to 100.0	**Specificity (%)** = 99.9 95% CI: 98.8 to 99.9	

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

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

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

		Reference standard: Positive (n)	Reference standard: Negative (n)	
Index test (α1)	Positive (n)	33	16	**PPV (%)** = 67.3 95% CI: 52.5 to 80.1
	Negative (n)	9	44 206	**NPV (%)** = 100.0 95% CI: 100.0 to 100.0
		Sensitivity (%) = 78.6 95% CI: 63.2 to 89.7	**Specificity (%)** = 100.0 95% CI: 99.9 to 100.0	

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

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

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

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 to 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 (online supplemental 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 (online supplemental table S5).

Colorectal cancer

The kappa value in the chart review for both diagnosis definitions and death was 0.953 (95% CI 0.900 to 1.006). There were 39 false positives in β2 (figure 2); 15 were diagnosed with colorectal cancer before 2014, 2 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 to 0.997) and 0.940 (95%

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

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

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

B. Breast cancer (kappa value [95% CI]: 0.961 [0.917 to 0.005])

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

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

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

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

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

D. Ovarian cancer (kappa value [95% CI]: 0.940 [0.940 to 1.007])

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

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

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

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

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

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

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

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

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

CI 0.873 to 1.007) for death. PPV was higher with γ1 than with γ2 (75.9% vs 49.5%; online supplemental table S4). Sensitivity was higher with γ2 than with γ1 (100.0% vs 89.8%; online supplemental 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 to 0.985) and 0.878 (95% CI 0.784 to 0.973) for death. Sensitivity was 100% in ε2, but PPV was as low as 42.0% (online supplemental table S4). PPV was higher with ε1 than with ε2 (67.3% vs 42.0%; online supplemental 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 to 0.995) and 0.9045 (95% CI 0.798 to 1.011) for death. PPV was 100% in δ1 (online supplemental 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/diarrhoea 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 2 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 the 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).

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

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	2477	1728 (69.8)	75.0 (9.9)	72.8 (10.2)	801.4 (626.7)	1985024
All hospitals	19861	13136 (66.1)	74.8 (10.2)	73.5 (10.4)	523.9 (552.4)	10405993
Breast cancer						
Kurashiki Central Hospital	1166	10 (0.9)	67.0 (13.3)	64.1 (13.3)	1022.6 (650.8)	1192400
All hospitals	18289	131 (0.7)	64.7 (14.1)	62.6 (14.1)	780.5 (618.6)	14274791
Colorectal cancer						
Kurashiki Central Hospital	1684	989 (58.7)	73.6 (11.3)	71.1 (11.6)	930.5 (613.5)	1566924
All hospitals	23501	13836 (58.9)	74.1 (11.3)	72.1 (11.5)	770.6 (596.2)	18110552
Ovarian cancer						
Kurashiki Central Hospital	265	34 (12.8)	66.4 (15.4)	63.9 (15.5)	896.2 (653.5)	237497
All hospitals	2592	145 (5.6)	64.1 (14.9)	62.3 (15.1)	667.3 (581.1)	1729551
Bladder cancer						
Kurashiki Central Hospital	568	446 (78.5)	77.6 (10.0)	75.0 (10.5)	991.3 (611.8)	563042
All hospitals	7408	5810 (78.4)	76.9 (10.4)	74.9 (10.6)	799.9 (595.8)	5925496
Prostate cancer						
Kurashiki Central Hospital	3131	3057 (97.6)	76.5 (8.4)	71.9 (8.7)	1703.1 (1118.3)	5332446
All hospitals	32136	28690 (89.3)	77.7 (8.9)	74.2 (9.2)	1341.3 (1041.6)	43105126

ICD-10, International Classification of Diseases, 10th revision.

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 patients with NSCLC 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 organisation data in the USA 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 tumour 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 generalised. 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 minimise 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 randomised 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 generalisability 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.

Author affiliations
[1]Department of Management, Clinical Research Center, Kurashiki Central Hospital, Kurashiki, Japan
[2]Department of Otolaryngology/Head and Neck Surgery, Kurashiki Central Hospital, Kurashiki, 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

Twitter Takashi Fujiwara @TFujiwarbi

Acknowledgements The following persons from Kurashiki Central Hospital Clinical Research Centre (Department of Management, Clinical Research Centre, 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 provided support for statistical analysis and Yusuke Miyoshi of Chugai Pharmaceuticals Co 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.

Contributors TK, KT, AY and HT conceptualised the original idea. TF, TK, KT, AY and HT designed the study. TF, MI and HT collected the study data. TF and HT had access to all the study data. TF, KT and YO contributed to the analyses. TK drafted the initial manuscript. TF, KT, YA, MI, YO and HT provided critical interpretation and contributed to the revision of the manuscript. All authors provided final approval for the version to be published.

Funding This study was funded by Chugai Pharmaceutical Co.

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

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

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 Hospital (KCH3301), and conducted under the tenets of the Declaration of Helsinki, Act on the Protection of Personal Information, and Ethical Guidelines for Medical and Health Research Involving Human Subjects. It was conducted under a joint research agreement between Kurashiki Central Hospital, Chugai Pharmaceutical Co and HCEI. Target patients at Kurashiki Central Hospital could opt, on the hospital's website, to not disclose their information.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iD
Takashi Fujiwara http://orcid.org/0000-0002-6790-8713

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;103:202–5.
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;34:581–8.
4 Lin Y-S, Shen Y-C, Wu C-Y, *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 J-M, Lin C-C, Liu K-L, *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:e83–9.
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;39:1733–50.
12 Smith GL, Shih Y-CT, Giordano SH, *et al*. A method to predict breast cancer stage using Medicare claims. *Epidemiol Perspect Innov* 2010;7:1.
13 Yen TWF, 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 Japan Pharmaceutical Manufacturers Association. Chapter 4: post-marketing surveillance of drugs. Pharmaceutical regulations in Japan, 2020. Available: https://www.jpma.or.jp/english/about/parj/eki4g6000000784o-att/2020e_ch04.pdf [Accessed 21 Dec 2021].
19 Basic concept of validation of outcome definition used in post-marketing database survey: pharmaceuticals and medical devices agency, Japan, 2020. Available: https://www.pmda.go.jp/files/000235927.pdf [Accessed 22 Dec 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: http://www.jspe.jp/committee/020/0271_1/ [Accessed 13 Jan 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;91:1944–8.
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: http://www.jspe.jp/mt-static/FileUpload/files/JSPE_DB_TF_E.pdf [Accessed 26 Oct 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: https://www.ppc.go.jp/files/pdf/APPI_The_Every_Three_Year_Review_Outline_of_the_System_Reform.pdf [Accessed 22 Dec 2021].
37 Ministry of Health, Labour and Welfare, Japan. Ethical guidelines for medical and health research involving human subjects. Available: https://www.mhlw.go.jp/file/06-Seisakujouhou-10600000-Daijinkanboukouseikagakuka/0000080278.pdf [Accessed 22 Dec 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: http://www.zengankyo.ncc.go.jp/etc/index.html [Accessed 26 Oct 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: https://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryou/iryouhoken/dl/shidou_kansa_01.pdf [Accessed 22 Dec 2021].