PRACTICAL CONSIDERATIONS OF JAPANESE HEALTH INSURANCE CLAIM DATABASES FOR REAL-WORLD EVIDENCE IN ONCOLOGY: INSIGHTS FROM A SYSTEMATIC LITERATURE REVIEW

RWD3

Wendy D'Souza[a], Sonali Dalwadi[a], Vidula Bhole[a], Haruyoshi Ogata[a], **Ivan D'Souza**[a]

[a]MedPro Clinical Research, 2-2-15 Minami Aoyama, Minato-ku, Minato City, Tokyo 107-0062, Japan

CONCLUSION

- Japanese healthcare databases are increasingly important sources of oncology RWE
- Database selection should be driven by research objectives, considering database-specific strengths and limitations
- Methodological fit is critical, particularly for machine learning and survival analyses
- This review provides evidence-based guidance for selecting the most appropriate Japanese claims database for oncology research
- Better-informed database selection can strengthen RWE study planning and reduce gaps or limitations in study design

INTRODUCTION

- **Growing real-world data (RWD) infrastructure** in Japan is accelerating real-world evidence (RWE) generation[1]
- **Major Japanese databases:** Diagnosis Procedure Combination (DPC) database, JMDC, Medical Data Vision (MDV), and Japanese National Database of Health Insurance Claims and Specific Health Checkups (NDB)[2,3]
- **Cancer burden in Japan:** ~1.02 million new cases and 380,400 cancer-related deaths in 2022[4]
- RWE complements clinical trials by capturing **real-world effectiveness, safety, treatment patterns, and outcomes** across diverse populations[5]
- **Database suitability varies** by clinical detail, population coverage, longitudinality, and outcome ascertainment[6,7]
- **Evidence gap:** No comprehensive review has systematically assessed the **oncology applications, study designs, strengths, and limitations** of Japanese claims databases

OBJECTIVE

The study aimed to systematically characterize oncology research using large-scale Japanese healthcare databases to guide appropriate database selection and improve oncology RWE quality

METHODS

Study Design and Data Sources

- Systematic literature review of research using Japanese administrative claims databases in oncology
- The PubMed literature search details are provided in **Figure 1**

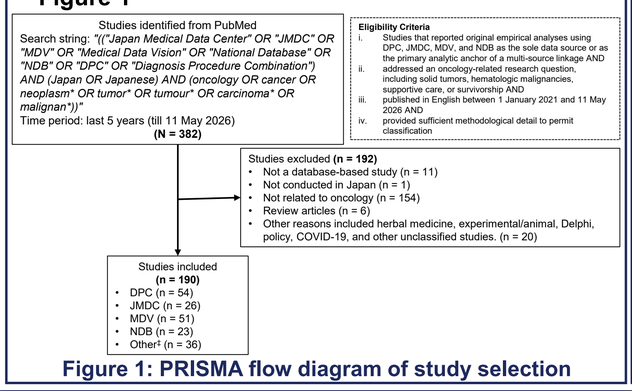

Figure 1: PRISMA flow diagram of study selection

Data Extraction and Analysis

- The information collected from eligible studies included bibliographic identifiers, primary database(s), cancer type, research objective, study design, statistical methods, and primary and secondary endpoints
- Research objectives were classified into seven prespecified domains: treatment patterns; safety/pharmacovigilance; healthcare resource utilization (HCRU)/cost; comparative effectiveness; survival; precision oncology/biomarkers; and machine learning/predictive modeling
- Data were summarized descriptively by study design and endpoint type

RESULTS

Distribution of Publications by Year and Database

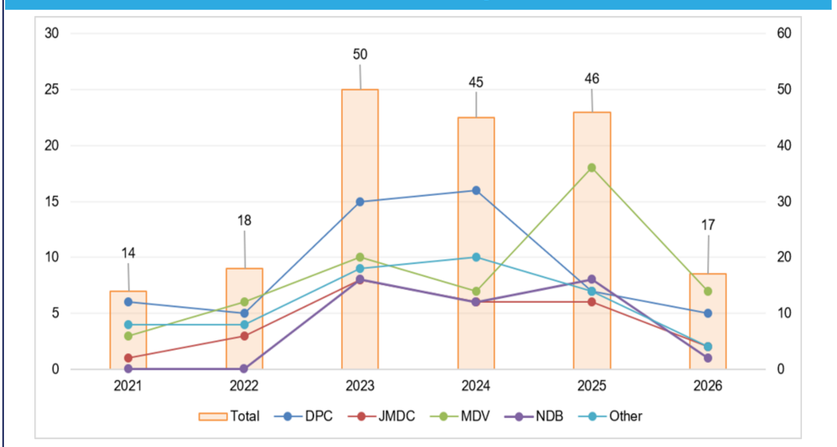

Figure 2: Distribution of publications by year and database source

- **Publication activity peaked in 2023 and remained high through 2025:** it rose from 14 studies in 2021 to **50 in 2023**, with **141/190 studies (74.2%) published during 2023–2025**, indicating a sustained period of intensive oncology RWE generation (**Figure 2**)

Types of Research Studies by Database

Table 1: Study designs across Japanese health insurance claim databases

Study designs, n (%)	DPC n = 54	JMDC n = 26	MDV n = 51	NDB n = 23	Other‡ n = 36	Total N = 190
Type of study						
Case-control	0 (0.0)	1 (3.8)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.5)
Cohort	27 (50.0)	19 (73.1)	31 (60.8)	9 (39.1)	15 (41.7)	101 (53.2)
Cross-sectional	1 (1.9)	2 (7.7)	1 (2.0)	3 (13.0)	1 (2.8)	8 (4.2)
Insufficient details	26 (48.1)	4 (15.4)	19 (37.3)	11 (47.8)	20 (55.6)	80 (42.1)
Type of study design						
Prospective	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (8.3)	3 (1.6)
Retrospective	54 (100)	26 (100)	51 (100)	23 (100)	33 (91.7)	187 (98.4)

- **Cohort** study design predominated across databases, **JMDC (73.1%), MDV (60.8%), DPC (50.0%), Others (41.7%),** and **NDB (39.1%);** followed by **cross-sectional** study design (**Table 1**)
- Most studies using **DPC, JMDC, MDV,** and **NDB** were **retrospective** designs

Cancer Types by Database

Table 2: Cancer types represented in studies using Japanese administrative claims and healthcare databases, by database source

Cancer group, n (%)	DPC n = 54	JMDC n = 26	MDV n = 51	NDB n = 23	Other‡ n = 36	Total N = 190
Hematologic malignancies	3 (5.6)	0 (0.0)	13 (25.5)	0 (0.0)	0 (0.0)	16 (8.4)
Gastrointestinal cancers	4 (7.4)	2 (7.7)	6 (11.8)	2 (8.7)	4 (11.1)	18 (9.5)
Hepatobiliary and pancreatic cancers	5 (9.3)	1 (3.8)	3 (5.9)	3 (13.0)	3 (8.3)	15 (7.9)
Genitourinary cancers	5 (9.3)	1 (3.8)	9 (17.6)	1 (4.3)	0 (0.0)	16 (8.4)
Gynaecologic cancers	3 (5.6)	1 (3.8)	1 (2.0)	0 (0.0)	0 (0.0)	5 (2.6)
Thoracic cancers	10 (18.5)	2 (7.7)	7 (13.7)	1 (4.3)	6 (16.7)	26 (13.7)
Breast cancer	2 (3.7)	4 (15.4)	4 (7.8)	6 (26.1)	2 (5.6)	18 (9.5)
Head and neck cancer	1 (1.9)	0 (0.0)	1 (2.0)	0 (0.0)	0 (0.0)	2 (1.1)
Central nervous system tumours	2 (3.7)	0 (0.0)	0 (0.0)	0 (0.0)	1 (2.8)	3 (1.6)
Sarcomas and skin cancers	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (5.6)	2 (1.1)
Other/rare cancers	3 (5.6)	2 (7.7)	3 (5.9)	0 (0.0)	5 (13.9)	13 (6.8)
More than one cancer type	14 (25.9)	12 (46.2)	4 (7.8)	10 (43.5)	13 (36.1)	53 (27.9)
Not specified	2 (3.7)	1 (3.8)	0 (0.0)	0 (0.0)	0 (0.0)	3 (1.6)

- Cancer-type specialization differed by database: **DPC** - thoracic cancers (18.5%; n=10), **MDV** - hematologic malignancies (25.5%; n=13) and genitourinary cancers (17.6%; n=9) (**Table 2**)
- **NDB** showed a strong focus on breast (26.1%; n=6) and hepatobiliary/pancreatic cancers (13.0%; n=3)
- Multi-cancer studies were particularly common in **JMDC** (46.2%), **NDB** (43.5%), and **other** (36.1%) databases

Research Outcomes by Database

Table 3: Distribution of target research outcome by database source

Outcome type, n (%)	DPC n = 54	JMDC n = 26	MDV n = 51	NDB n = 23	Other‡ n = 36	Row Total
OS	11 (25.0)	4 (9.1)	18 (40.9)	0 (0.0)	11 (25. 0)	44 (7.0)
Mortality	33 (37.5)	8 (9.1)	29 (33.0)	3 (3.4)	15 (17.0)	88 (14.0)
PFS	1 (20.0)	0 (0.0)	1 (20.0)	0 (0.0)	3 (60.0)	5 (0.8)
Real-world PFS	1 (10.0)	0 (0.0)	5 (50.0)	0 (0.0)	4 (40.0)	10 (1.6)
Hospitalization	48 (43.6)	10 (9.1)	31 (28.2)	9 (8.2)	12 (10.9)	110 (17.5)
Economic evaluation	11 (39.3)	1 (3.6)	10 (35.7)	4 (14.3)	2 (7.1)	28 (4.5)
Cost outcomes	13 (35.1)	1 (2.7)	13 (35.1)	5 (13.5)	5 (13.5)	37 (5.9)
Safety	42 (37.2)	17 (15.0)	37 (32.7)	6 (5.3)	11 (9.7)	113 (18.0)
Response rates	2 (16.7)	1 (8.3)	1 (8.3)	0 (0.0)	8 (66.7)	12 (1.9)
Quality of life	7 (38.9)	2 (11.1)	3 (16.7)	2 (11.1)	4 (22.2)	18 (2.9)
Dose modification	2 (10.0)	1 (5.0)	11 (55.0)	3 (15.0)	3 (15.0)	20 (3.2)
Treatment switching	4 (7.3)	5 (9.1)	36 (65.5)	2 (3.6)	8 (14.5)	55 (8.8)
Persistence/adherence	4 (9.5)	6 (14.3)	26 (61.9)	2 (4.8)	4 (9.5)	42 (6.7)
Biomarker outcomes	4 (9.5)	9 (21.4)	11 (26.2)	2 (4.8)	16 (38.1)	42 (6.7)
Machine learning used	1 (33.3)	0 (0.0)	1 (33.3)	0 (0.0)	1 (33.3)	3 (0.5)

- **DPC** was commonly used for **healthcare utilization and survival-related outcomes**: hospitalization (43.6%; n=48) and mortality (37.5%; n=33) (**Table 3**)
- **MDV** showed a strong focus on **treatment-related outcomes:** treatment switching (65.5%; n= 36), persistence/adherence (61.9%; n=26), dose modification (55.0%; n=11), and overall survival (40.9%; n=18)
- **JMDC** was suitable for **safety/pharmacovigilance** (15.0%; n=17), **biomarker/precision oncology** (21.4%; n= 9) and multi-cancer studies
- **NDB** was suitable in multi-cancer and breast cancer research, with broader population coverage
- **Other databases captured specialized outcomes:** response rates (66.7%; n=8) and biomarker outcomes (38.1%; n=16)

Key Characteristics of Databases for Oncology Applications

Table 4: Comparative overview of key strengths, limitations, and suitability observed in oncology research using Japanese health insurance claim databases

Database	Key strengths in oncology	Limitations	Focus area for oncology research
DPC	Hospitalization, adverse-event, and survival/outcome studies; detailed inpatient/procedural data	Primarily captures acute inpatient care; limited cross-hospital traceability and long-term follow-up	Surgical/procedural oncology; hospitalization and inpatient safety; acute outcomes; inpatient survival
JMDC	Longitudinal insurer-based claims; multi-facility follow-up; safety/pharmacovigilance; precision oncology/biomarker research	Population coverage excludes the late-stage elderly care system; limited laboratory data	Treatment patterns; safety/pharmacovigilance; persistence/adherence; biomarker analyses; multi-cancer studies
MDV	Treatment-pattern studies; treatment switching, persistence/adherence and dose modification	Hospital-based; limited tracking after transfer to other hospitals	Treatment sequencing; switching; adherence; dose modification
NDB	Nationwide coverage; population-level treatment patterns; multi-cancer research	Limited clinical severity and mortality data	Population oncology; treatment patterns; HCRU/costs
Other‡	Precision oncology/biomarkers; response outcomes	Heterogeneous data sources with non-generalizable characteristics	Precision oncology; biomarkers; response assessment; specialized oncology research

Algorithm for Selection of Database

Database Selection Algorithm

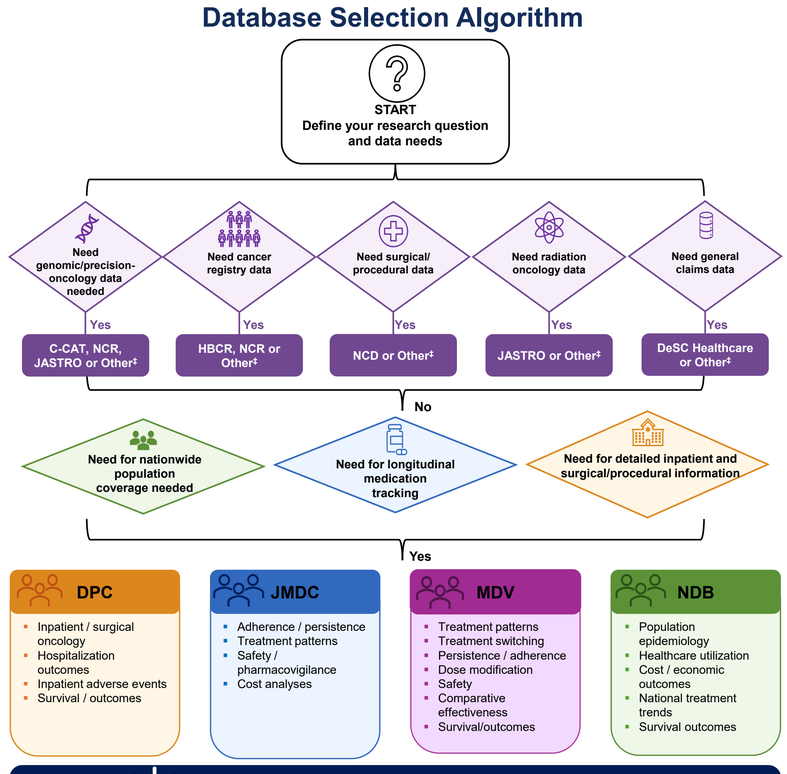

Figure 3: Database selection algorithm (choosing the right Japanese health insurance claim database for oncology research)

- A **research-question–driven database selection approach** can optimize the use of Japanese health insurance claims databases for oncology RWE and generate **fit-for-purpose evidence for patient care** (**Figure 3**)

Benefits of RWE using Claims Database

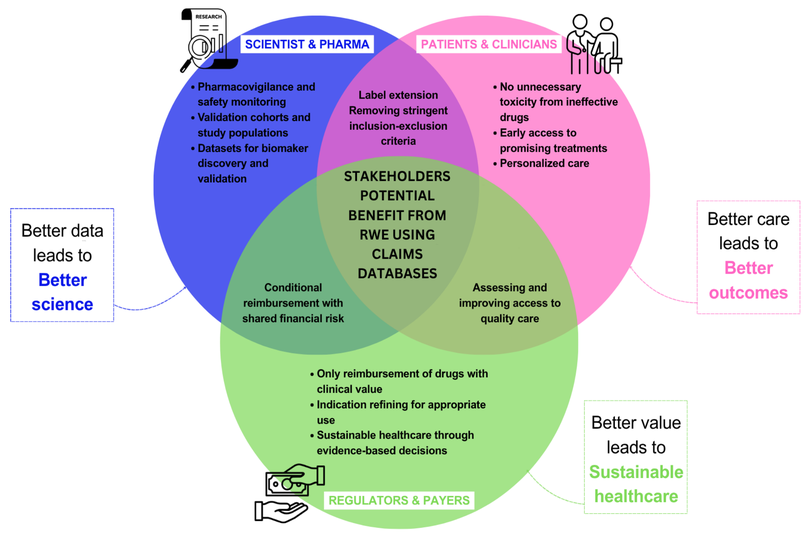

Figure 4: Stakeholder benefits of real-world evidence using claims databases

Claims-based RWE can align the needs of **scientists, clinicians, patients, regulators, and payers** to drive better science, better care, and sustainable healthcare (**Figure 4**)

DECLARATIONS

- **Acknowledgments:** Medical writing support was provided by Mansi Patel of MedPro Clinical Research
- **Funding:** This study received no external funding
- **Disclosures:** All authors are employees of MedPro Clinical Research and declare that they have no conflicts of interest
- **Ethical considerations:** Only previously published studies were analyzed. No institutional review board approval was required

References: 1. Bando, H. *et al.* (2023) ESMO Real World Data and Digital Oncology, 2, pp. 100005–100005. **2.** Wakabayashi, Y., *et al.* (2021) *BMC Medical Informatics and Decision Making*, 21(1), pp. 167–167. doi:10.1186/s12911-021-01526-6. **3.** Fujinaga, J. and Fukuoka, T. (2022) Drugs - Real World Outcomes, 9(4), pp. 543–550. doi:10.1007/s40801-022-00331-5. **4.** Nguyen, P. T. et al. (2023) Cancer Epidemiology, Biomarkers & Prevention, 32(12), pp. 1756–1770. doi:10.1158/1055-9965.EPI-23-0754. **5.** Azoulay, L. (2022) The Oncologist, 27(9). doi:10.1093/oncolo/oyac114. **6.** Baudrier, C. et al. (2022) Health Informatics Journal, 28(2). doi:10.1177/14604582221101526. **7.** Sato, S. and Yasunaga, H. (2023) Annals of Clinical Epidemiology, 5(2), pp. 58–64. doi:10.37737/ace.23008.

Abbreviations: C-CAT, Center for Cancer Genomics and Advanced Therapeutics; DPC, Diagnosis Procedure Combination; EMR, Electronic Medical Record; HBCR, Hospital-Based Cancer Registry; JJCLCR, Japanese Joint Committee of Lung Cancer Registry; LC-SCRUM-Asia, Lung Cancer Genomic Screening Project for Individualized Medicine in Asia; MDV, Medical Data Vision; NCD, National Clinical Database; NDB, National Database of Health Insurance Claims and Specific Health Checkups of Japan; OCR, Osaka Cancer Registry, OS; overall survival; PFS, progression free survival.
‡Other databases included Comprehensive Cancer Genome Profiling and Clinical Annotation Database (C-CAT; n = 9), Hospital-Based Cancer Registries (HBCR; n = 9), Japanese Registry of All Cardiac and Vascular Diseases and Diagnosis Procedure Combination database (JROAD-DPC; n = 2), National Clinical Database (NCD; n = 2), National Cancer Registry (NCR; n = 1), National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB) and Japanese Society for Radiation Oncology (JASTRO) Structure Survey (n = 1), database provided by DeSC Healthcare Inc. (n = 1), LC-SCRUM-Asia (n = 1), Hospital-Based Cancer Registries, Department of Oncology and Lung Cancer Center of China–Japan Friendship Hospital (n = 1), HBCR National Cancer Center Hospital (n = 1), national database compiled by the Ministry of Health, Labour and Welfare (MHLW) (n = 1), National Insurance Claims Database (n = 1), Osaka Cancer Registry (OCR; n = 1), Japanese Lung Cancer Registry Study (n = 1), Tokushukai REAl-World Data Project (TREAD; n = 1), Japanese Specific Health Check and Guidance System (n = 1), and the Report on Regional Public Health Services and Health Promotion Services 2017 (n = 1).

Poster presented at ISPOR Asia Pacific Summit, Bangkok, Thailand, 6-8 September 2026