Key characteristics and scientific influence of database studies on drug effectiveness in the post-marketing stage

Chihaya Shiragasawa

DP19402

Department of Clinical Medicine (Pharmaceutical Medicine)

Graduate School of Pharmaceutical Sciences

Kitasato University

5-9-1 Shirokane, Minato-ku, Tokyo, 108-8641, Japan

Abstract

In recent years, real-world data (RWD) have been actively used in the field of pharmaceutical research. Database (DB) study, one of the observational studies using RWD, is a very comprehensive, continuous, and rapid research method that plays an important role in the post-marketing stage of drugs, although the interpretation of the results may be limited. DB studies are often focused on drug safety, and previous research reviewing DB studies on drug effectiveness across different disease areas has been limited. Few studies have also reviewed DB studies in terms of scientific impact.

To consider conducting appropriate DB studies on drug effectiveness in the post-marketing stage, first, we tried to reveal the current status of DB studies on drug effectiveness in various therapeutic areas to provide information such as research design and outcome definition [Research 1]. Second, we analyzed which elements (characteristics) of the studies were related to the level of scientific influence, using the number of citations and citation sources, based on the DB studies extracted in the first research[Research 2]. Based on the results, we discussed measures to promote appropriate DB studies to help generate evidence on drug effectiveness with high scientific influence in the post-marketing stage.

We searched the Embase and MEDLINE for DB studies on drug effectiveness published between January 1, 2018, and December 31, 2019. The name and type of the database (administrative claim DB, clinical DB, pharmacy DB, and DB linkage), study design, comparison group, type of outcome, and presence or absence of reference to the outcome definition were extracted and summarized according to disease areas [Research 1]. Next, multivariable linear regression analysis was performed with the number of citations as the response variable and the characteristics of the study as explanatory

variables (impact factor, publication year, DB classification, disease area, sample size, outcome type, outcome validation presence, propensity score use, sensitivity analysis implementation, and funding). In addition, citation sources were reviewed, and reports cited in clinical practice guidelines were characterized [Research 2].

In Research 1, we obtained 225 articles on DB studies on drug effectiveness using DBs that integrate large-scale medical data for secondary use across different disease areas. Among the DB classifications, administrative claim DB (70%, 158/225) were most commonly used, while pharmacy DB was used in only three studies. The largest number of reported studies was associated with cardiovascular, respiratory, and infectious diseases. Outcomes were often inpatient diagnosis, and some ideas included defining effectiveness based on drug use. While various outcomes were uniformly used in studies for the treatment of infectious diseases and respiratory organs, death (overall survival) and drug continuation (progression-free survival) in patients with cancer, laboratory values in endocrine system (mainly diabetes) were used as main outcomes. Outcome validation within the article was limited.

In Research 2, multivariable linear regression analysis showed that the number of citations was significantly associated with impact factor, publication year, outcome type (inpatient diagnosis). The 206 reports had a total of approximately 5000 citations, and among them 31 (15 %) were cited in the 32 clinical practice guidelines. The leading disease areas were circulatory system diseases, followed by respiratory system diseases. Many of these were large-scale, highly comprehensive studies that used national databases. These were cited as supporting information to real-world evidence, mainly for outcomes with little evidence from interventional studies.

This study summarized the status of cross-disease research articles on DB studies

on drug effectiveness. The use of inpatient diagnosis as an outcome for DB studies on drug effectiveness would increase scientific influence. The appropriate setting for effectiveness outcome definitions, such as disease occurrence, as well as the use of large and comprehensive DBs, is important for generating evidence of post-marketing effectiveness with high scientific influence. While considering the strengths and limitations of DB studies, we believe that our comprehensive results would help to promote appropriate DB studies on drug effectiveness in the post-marketing stage.

Table of Contents

List of Tables

List of Figures

Abbreviations

CPRD	Clinical Practice Research Datalink
DB	Database
DPC	Japanese Diagnosis Procedure Combination
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
FDA	US Food and Drug Administration
HTA	health technology assessment
ICD-10	International Classification of Diseases, Tenth Revision
LRx	German longitudinal prescription database
MDV	Medical Data Vision
MHLW	Japanese Ministry of Health, Labour and Welfare
MID-NET	Medical Information Database NETwork
mRS	modified Rankin Scale
OS	overall survival
PFS	progression-free survival
RWD	real-world data
RWE	real-world evidence
SEER	The Surveillance, Epidemiology, and End Results

1. Introduction

In recent years, real-world data (RWD) have been actively used in the field of pharmaceutical research [1-3]. Amid the global movement to strengthen the monitoring system for adverse drug reactions and the increasing importance of health technology assessment (HTA), medical information databases (DB) owned by national organizations, medical institutions, insurance companies, and private companies have been developed, and studies on safety and HTA outcomes are actively conducted. In addition, the 21st Century Cures Act [4] enacted in the United States in 2016 proposed to change the qualification process for drug development tools (e.g., clinical outcome assessment) so that real-world evidence (RWE) can be used as data for new drug applications and post-marketing requirements. Accordingly, pharmaceutical companies are increasingly interested in generating RWE of drug effectiveness.

In addition, regulatory schemes to provide early availability of promising new drugs to patients, such as Accelerated Approval (US Food and Drug Administration; FDA), Conditional Marketing Authorization (European Medicines Agency; EMA), and Conditional Early Approval (Japanese Ministry of Health, Labour and Welfare; MHLW), have been implemented internationally. Under these schemes, to promptly cater to unmet medical needs, some drugs for serious diseases are marketed with less matured data from clinical trials before approval, which requires RWD to validate the effectiveness and safety after approval. Therefore, information on safety and effectiveness in actual clinical practice is of great interest to healthcare professionals and patients.

Efficacy/effectiveness and safety, both of which are essential properties of a drug, should be assessed in a balanced manner in the pre- and post-marketing stage. Important missing information in the post-marketing stage would be information on drug

efficacy/effectiveness obtained from prospective studies with comparison groups. Randomized, placebo-controlled trials in the post-marketing stage are not feasible due to some ethical concerns that, in such trials, patients cannot receive treatments that are supposed to be used in ordinary medical practice. In these cases, although obtaining verifiable results is challenging, proper implementation of outcome studies on effectiveness using RWD may be one solution. However, while various guidelines from regulatory authority for safety outcome studies using DB are available [5-7] and many studies have been conducted in the field of pharmacoepidemiology, there are few guidelines mentioning effectiveness evaluation other than brief descriptions, if any [8]. Although not a regulator, the joint ISPOR-ISPE Special Task Force reports a recommendation paper on real-world data studies of treatment and/or comparative effectiveness [9]. Also, the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) published recommendations on comparative effectiveness research [10]. Needless to say, observational studies produce results with various biases, so appropriate measures must be taken to reduce the bias, such as new user design, propensity score matching, and sensitivity analysis.

A comprehensive evaluation of the current state of DB studies is difficult because there is no standard search term for searching for articles on DB studies. Furthermore, it is rare to find articles summarizing DB studies on drug effectiveness across different disease areas.

Various bibliometric indices have been developed worldwide as measurable indicators to evaluate research outcomes (quality of research), including citation number, journal impact factor, h-index, and field-standardized indices [11-13]. The number of citations is a basic and straightforward indicator for evaluating the scientific influence of

an individual study. In addition, we believe that scientific influence can be evaluated in an additive manner by verifying whether the study has been cited in important documents such as clinical practice guidelines. Few studies have examined the characteristics of DB studies on drug effectiveness with high scientific influence.

Given this background, we tried to reveal the current status of DB studies on drug effectiveness in various therapeutic areas to provide information such as research design and outcome definition [Research 1]. Then, we analyzed which elements (characteristics) of the studies were related to the level of scientific influence, using the number of citations and citation sources, based on the DB studies extracted in Research 1 [Research 2]. Based on the results, we discussed measures to promote appropriate DB studies to help generate evidence on drug effectiveness with high scientific influence in the post-marketing stage.

2. Research 1　(Key characteristics of database studies on drug effectiveness)

2.1. Objectives

The objective of this research was to reveal the current status of DB studies on drug effectiveness in various therapeutic areas and to provide information such as research design and outcome definition that allows researchers to consider conducting appropriate DB studies on drug effectiveness in the post-marketing stage.

2.2. Methods

2.2.1 Search strategy

To extract articles relevant to DB studies on drug effectiveness published between January 1, 2018, and December 31, 2019, we used the following search criteria

in Embase and MEDLINE. We have narrowed down the types of articles to [Journal Article, Article, Comparative Study, Observational Study, Evaluation Studies, Pragmatic Clinical Trial, and Validation Studies].

Search Formula: (S1 OR S2 OR S3 OR S4) AND (S8) NOT (S5 OR S6 OR S7)

S1 : ti(database*) OR ab(database*)

S2 : ti(data PRE/2 base*) OR ab(data PRE/2 base*)

S3 : ti(admin* PRE/2 data*) OR ab(admin* PRE/2 data*)

S4 : ti(claims PRE/2 data*) OR ab(claims PRE/2 data*)

S5 : ti, ab(systematic review OR literature)

S6 : ti(clinical trial)

S7 : ti(randomized)

S8 : ti, ab(efficacy OR effectiveness)

S1 to S4: Extraction condition; articles containing [database], [data base], [admin data], or [claims data] in the title or abstract were extracted. Searches were conducted using asterisk (*) to allow for changes in the endings, and "PRE/2" to prevent the two search terms from being separated by more than two words. S5: Exclusion condition; articles containing [systematic review] or [literature] in the title or abstract were excluded because name of the DB used in literature search is often described in the article of literature search. S6, S7: Exclusion condition; articles containing [clinical trial] or [randomized] were excluded because a dataset obtained in a clinical trial is sometimes referred to as a DB. S8: Extraction condition; articles containing [efficacy] or [effectiveness] were extracted.

2.2.2 Eligibility criteria for study selection

To identify studies on drug effectiveness (with clinical outcomes and with or without comparators) using medical information DBs, we reviewed the title and abstract, and materials and methods if necessary, of each article, and excluded the following studies: non-medical studies (e.g., non-clinical studies, sociological studies), studies on non-drug treatments (e.g., surgery, radiotherapy), and studies on drug safety, actual use, or cost outcomes that do not include any effectiveness clinical outcomes. We also excluded non-DB studies utilizing clinical trial datasets, meta-analyses, genetics, and in vitro and in silico related DB libraries. One researcher screened all articles retrieved and consulted with another researcher in case that data selection needed discussion. This review was not registered.

2.2.3 DB classification and data extraction

For all the selected DB studies on effectiveness (data set I), targeted disease area and type of medical information DB (defined as 1–6 below) utilized in the study were extracted and summarized. Disease areas were classified by the International Classification of Diseases, Tenth Revision (ICD-10) based on the disease for which the effectiveness outcome was intended, not on the patients' primary disease. For example, a study using antimicrobial effect of statins in diabetic patients as an outcome was classified as infectious diseases rather than diabetes. Therapeutic areas of the drugs were classified by the Anatomical Therapeutic Chemical classification.

Medical information DBs were classified into the following six categories:

1. Administrative claim database : a DB based on insurance claim data managed by insurance companies, national or local governments. (e.g., Truven MarketScan)

2. Clinical database : a DB based on medical information such as electronic medical

record including laboratory test results integrated by medical information companies. (e.g., Flatiron Health)

3. Pharmacy database : a DB of prescribing information managed by pharmacy chains. (e.g., German longitudinal prescription database (LRx))

4. Database linkage : a DB in which data from multiple independent DBs are linked.

5. Registry: a DB of disease- and purpose-based patient registry, including the Surveillance, Epidemiology, and End Results (SEER) and the National cancer DB.

6. EMR database : a DB of a single hospital's stand-alone electronic medical record.

Next, for DB studies based on medical information DBs 1, 2, 3, and 4 (data set II), in which large-scale medical data are integrated for secondary use, we extracted the following information: name of the DB, study design, comparison group, type of outcome, and presence or absence of reference to the outcome definition. Subsequently, these studies were summarized in cross tables together with disease areas, DB classifications, comparison groups, and type of outcomes, measures to reduce bias, and funding information to examine the current status. As for the outcome definition, we classified an article as “reference present” if other research reports were cited as references within the article, and we did not verify whether this outcome validation study had been completed.

2.3. Result

2.3.1 Search selection and DB classification

We extracted 3,523 articles from Embase and MEDLINE. Of these, 864 articles were focused on medical research for drugs after excluding articles on non-medical research and research not for drugs based on title and abstract review (Fig. 1). After excluding articles on drugs safety, actual use, and cost; non-DB studies such as prospective observational studies based on primary data, clinical trials, in vitro studies,

and meta-analysis; and non-English articles based on review of the details of the articles, we identified 456 DB studies on drug effectiveness (data set I). Among them, we found 225 articles using administrative claim DB, clinical DB, pharmacy DB, and DB linkage (data set II), which are DBs that integrate large-scale medical data for secondary use. Administrative claim DB accounted for approximately 70% (152/225) of the data set II, while pharm DB had only three cases.

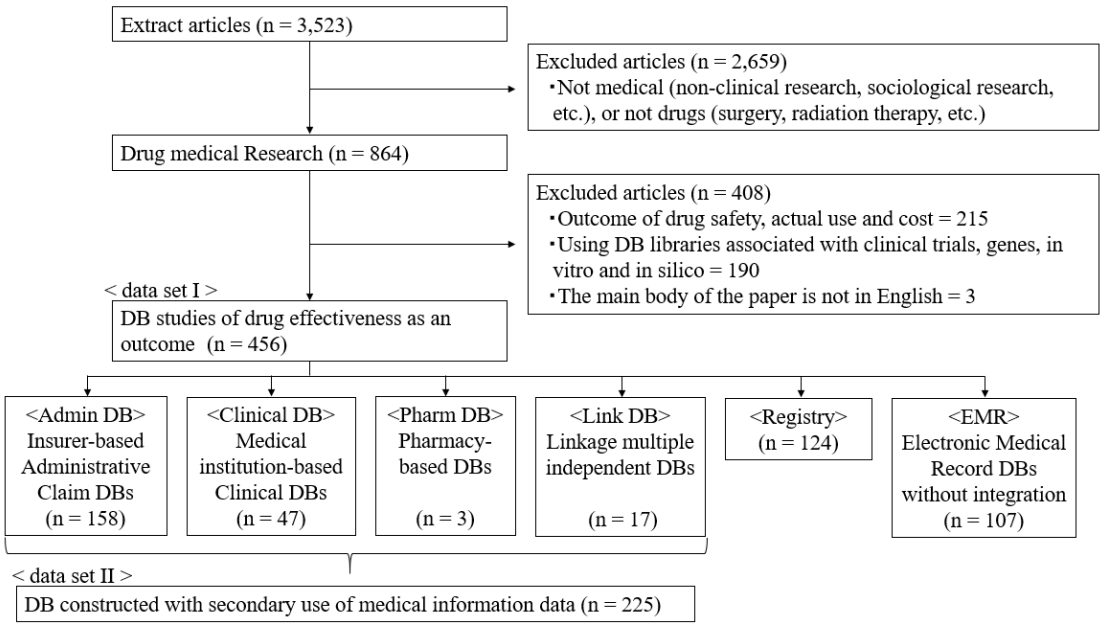

Figure 1 Flow diagram of study selection

2.3.2 DB classification and disease area

DB studies on drug effectiveness were widely available in most disease areas, with the exception of the categories in ICD-10 codes “External causes of morbidity and mortality (V01-Y98),” “Factors influencing health status and contact with health services (Z00-Z99),” and “Codes for special purposes (U00-U85)”, which are rarely used in studies on drug effectiveness. The top three categories of neoplasms (29%, 113/456),

diseases of the cardiovascular system (25%, 112/112), and infectious and parasitic diseases (10%, 46/456) accounted for approximately 64% of the articles (Table 1, data set I). In terms of the DB category, registry and administrative claim database were most commonly used in studies on neoplasms and cardiovascular diseases, respectively. In terms of the target therapeutic area of the drugs based on the Anatomical Therapeutic Chemical classification, we identified the top 10 classifications: L01 Antineoplastic Agents, B01 Antithrombotic Agents, J07 Vaccines, A10 Drugs Used in Diabetes, L04 Immunosuppressants, J05 Antivirals for Systemic Use, M01 Anti-inflammatory and Antirheumatic Products, J Anti-infectives for Systemic Use, R03 Drugs for Obstructive Airway Diseases, and M05 Drugs for Treatment of Bone Diseases.

In data set II, the number of DB studies was higher in cardiovascular, respiratory, and infectious diseases, in this order (Table 1, data set II). Neoplasms, which were the most frequently studied disease area utilizing registry and EMR DB, which were not included in the data set II, ranked fourth. A relatively high percentage of studies in neoplasms and endocrine and metabolic diseases used clinical DBs, while those in mental disorders and digestive system diseases used only administrative claim DBs.

Table 1　DB classification and ICD disease classification in the database studies on drug effectiveness

	DB classification [a]							
	data set I							
	data set II							
ICD-10 codes	Admin	Clinical	Pharm	Link	subtotal	Registry	EMR	total
C00-D48 Neoplasms	6	10		2	18	79	36	133
I00-I99 Diseases of the circulatory system	72	9		7	88	9	15	112
A00-B99 Infectious and parasitic diseases	14	4	1	2	21	10	15	46
J00-J99 Diseases of the respiratory system	12	6	2	4	24	4	9	37
E00-E90 Endocrine, nutritional and metabolic diseases	4	10		1	15	1	8	24
M00-M99 Diseases of the musculoskeletal system and connective tissue	10	7			17	5	1	23
F00-F99 Mental and behavioral disorders	11				11	4	5	20
K00-K93 Diseases of the digestive system	8				8	3	3	14
N00-N99 Diseases of the genitourinary system	3			1	4	1	4	9
G00-G99 Diseases of the nervous system	4				4	2	2	8
R00-R99 Symptoms, signs and abnormal clinical and laboratory findings		3			3	1	3	7
D50-D89 Diseases of the blood and certain disorders involving the immune mechanism	3	1			4	2		6
L00-L99 Diseases of the skin and subcutaneous tissue	1				1	1	4	6
H00-H59 Diseases of the eye and adnexa	1	1			2		1	3
O00-O99 Pregnancy, childbirth and the puerperium	2				2		1	3
H60-H95 Diseases of the ear and mastoid process	1				1			1
S00-T98 Injury, poisoning and certain other consequences of external causes					0	1		1
Multi_ Critical illnesses (respiratory, CNS, cardiovascular, etc.) [b]		1			1			1
P00-P96 Certain conditions originating in the perinatal period		1			1			1
Q00-Q99 Congenital malformations, deformations,					0	1		1

and chromosomal abnormalities								
total	152	53	3	17	225	124	107	456

[a] Admin: administrative claim database, Clinical: clinical database, Pharm: pharmacy database, Link: database linkage, EMR: Electric Medical Record

[b] Because the study involved multiple diseases and could not be classified, we created our own classification name.

2.3.3 Specific DBs used in the studies

For administrative claim DBs, the National Claim DBs of Taiwan, Korea, and France, and the commercial DBs of Truven and Optum were the most commonly used (Table 2). For clinical DBs, several studies using the DB of Medical Data Vision (MDV), a Japanese company, and Flatrion, which has data specialized in the field of oncology, were reported. We found nine studies in which multi-DB was used for similar purposes and their results were analyzed collectively. Specifically, a study used four large-scale commercial DBs in the United States and increased the generalizability of the results based on a significantly large number of cases [14, 15]. Another study observed drug use in an administrative claim DB and examined laboratory data in a small clinical DB to compensate for the robustness of the results [16]. In addition, the use of DB linkage is effective to ensure the robustness of outcomes, as demonstrated by a study that linked a cancer registry and mortality statistics [17].

Table 2 DB Names for each DB classification

DB classification and DB name[a]	country	number of articles
Administrative database		
Taiwan National Health Insurance Research Database(NHIRD)	Taiwan	40
Truven MarketScan	USA	26
Korean Health Insurance Review and Assessment database	Korea	23
French National Health Insurance database (SNIIRAM)	France	7
Japanese Diagnosis Procedure Combination (DPC) database	Japan	7
Optum	USA	6
IBM MarketScan	USA	6
Medicare claims data	USA	6
JMDC	Japan	5
Clinical database		
MDV	Japan	7
Premier Healthcare	USA	6
Flatiron Health	USA	5
Military Health System Data Repository	USA	3
IQVIA EMR database	USA	2
The Health Improvement Network (THIN) UK primary care database	UK	2
Optum's Electronic Health Records database	USA	2
Pharmacy database		
German longitudinal prescription database (LRx)	Germany	2
specialty pharmacy database (northeastern region, United States)	USA	1

[a] Only high frequency DBs for each DB category are shown.
DB : database, EMR : Electronic Medical Records, JMDC : Japan Medical Data Center, MDV : Medical Data Vision

2.3.4 Study design

Most studies were multi-cohort studies with a comparison group(s) (90%, 202/225), followed by single-cohort studies (8%, 18/225). Case-controls studies (1%, 3/225) and self-control studies (1%, 2/225) were rare. There was no clear relationship between the study design and DB type .

2.3.5 Comparison group

Comparison groups commonly set up in the effectiveness studies comprised

patients taking a specific drug with the same indication of the target drug, patients not receiving the target drug, and patients administered a class of drugs with the same indication of the target drug, in this order (Table 3). When excluding studies for cardiovascular system (mainly antithrombotic drugs), patients not receiving the target drug were most commonly used as a comparison group. There was no clear relationship between the comparison group and disease area.

Table 3 Comparison group of database studies on effectiveness

	Category of comparison group [a]				
	A	**B**	**C**	**D**	**E**
total	**59**	**52**	**83**	**25**	**20**
ICD10 codes (Only high frequency disease) [b]	**A**	**B**	**C**	**D**	**E**
I00-I99 Diseases of the circulatory system	11	22	58	10	3
J00-J99 Diseases of the respiratory system	15	3	2	2	2
A00-B99 infectious and parasitic diseases	16	3	1		1
C00-D48 Neoplasms		6	4	5	3
M00-M99 Diseases of the musculoskeletal system and connective tissue	5	4	5	2	1
E00-E90 Endocrine, nutritional and metabolic diseases	1	8	3	1	2
F00-F99 Mental and behavioral disorders	3	1	2	1	4

[a] A: non-users of the target drug, B: class of drugs with the same indication, C: specific drug(s) with the same indication, D: different dose(s) of same drug , E: no comparison or self-control group

[b] Only high frequency disease classifications (10 or more studies) are shown.

2.3.6 Types of outcome

Outcomes used in the studies were categorized in seven: death, laboratory/clinical score, hospitalization/emergency room status, inpatient (admission/discharge) diagnosis, diagnosis, medical procedures, and drug discontinuation/change/addition. Studies using inpatient diagnosis as an outcome were

the most common, followed by diagnosis, death, and drug usage status (Fig. 2). Relatively few studies used laboratory/clinical scores or other medical procedures including surgery.

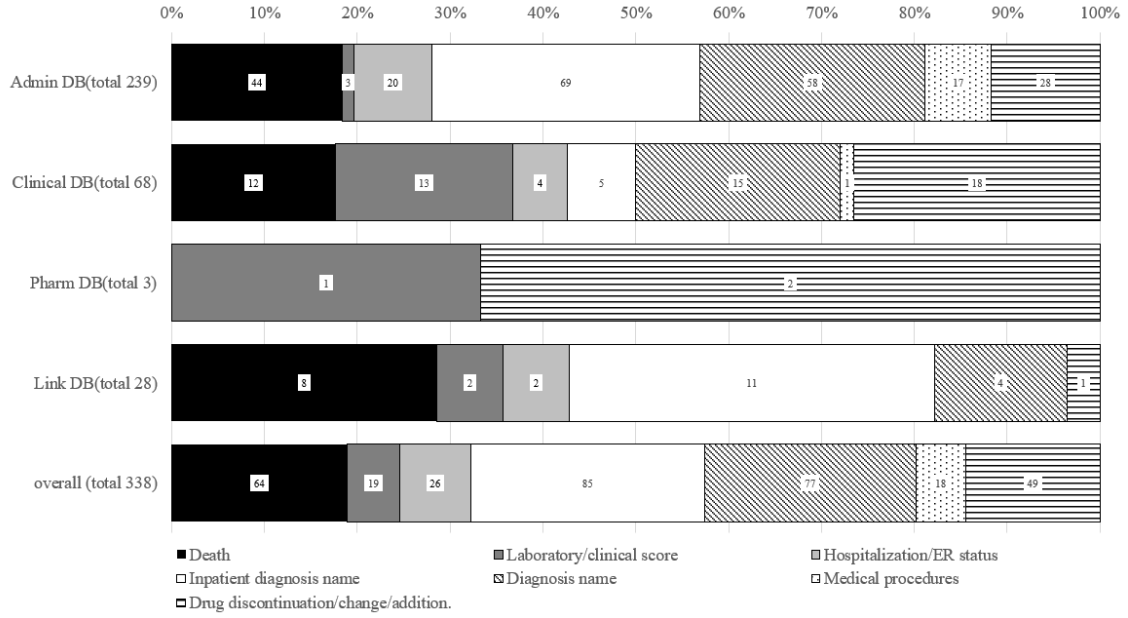

Figure 2 Types of outcomes for each DB classification

If there was more than one outcome within a study, each was counted. The numbers in the bar chart indicate the number of articles.

By the DB classification, clinical DBs often used laboratory/clinical scores. Some administrative claim DBs had laboratory data such as HbA1c used in some studies. The Japanese Diagnosis Procedure Combination (DPC) data [18] include outcome data such as modified Rankin Scale (mRS) to assess neurological functional disability after stroke [19]. Forty-nine studies defined effectiveness by drug discontinuation/change/addition.

By the disease category, while various outcomes were uniformly utilized in infectious diseases and respiratory organs (Fig. 3), death (overall survival, OS) and drug

continuation (progression-free survival, PFS) in patients with cancer, laboratory values involving the endocrine system (mainly diabetes), death and inpatient diagnosis and diagnosis of cardiovascular conditions, medical procedures relating to the digestive system, and diagnosis and drug continuation for the skeletal musculoskeletal system (antirheumatic, drugs etc.) were used as main outcomes.

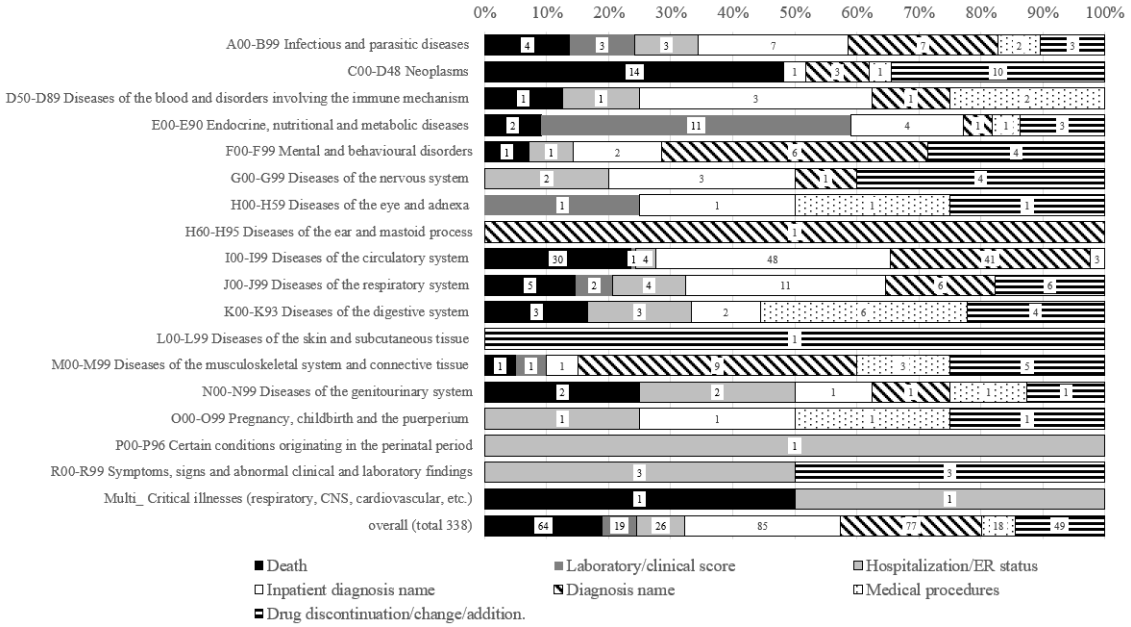

Figure 3 Types of outcomes for each ICD-10 classification

If there was more than one outcome within a study, each was counted. The numbers in the bar chart indicate the number of articles.

2.3.7 Outcome validation

Outcome validation was conducted in only five studies (2%, 5/225) within the present study. In 76 articles (33%, 76/225), the outcome validation and outcome definitions were cited from other articles; outcome validation was not mentioned in 144 studies (64%, 144/225). Of these, 42 studies (18%, 42/225) were for diagnoses (non-inpatient), medical procedures, and drug usage status, which might be considered to

require validation (Fig. 4).

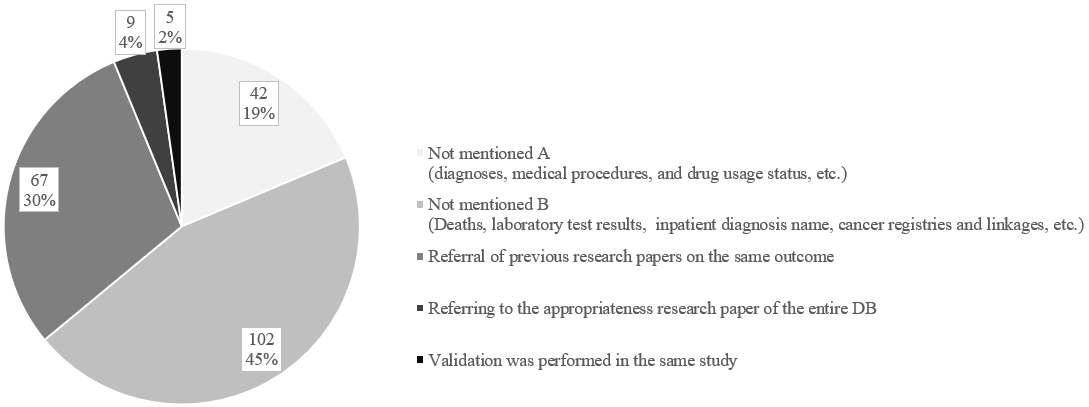

Figure 4 Reference to the outcome validation

The numbers in the pie chart indicate the number of articles and the percentage (%).

2.3.8 Measures to reduce bias

New user design was applied in 73 studies (32%, 73/225), and propensity score analysis (including matching, stratification, and inverse probability weighting) was performed in 131 studies (58%. 131/225) to reduce bias in observational studies. Sensitivity analysis was performed in 90 studies (40%, 90/225) to confirm the robustness of the results.

2.3.9 Funding information

Sixty-eight studies (30%, 68/225) were supported by pharmaceutical companies, and 2 of them were identified as regulatory post-marketing requirement studies.

2.4. Discussion

The present study reported the current status of DB studies on drug effectiveness for all the disease areas. We categorized the types of DB, disease area, study design, and definition of the outcomes used in the studies and provided interesting examples to

encourage researchers to consider conducting DB studies on effectiveness in the future. Three disease areas of neoplasms, cardiovascular diseases, and infectious and parasitic diseases accounted for 63% of DB studies on drug effectiveness in the recent two years. In these disease areas, the outcomes used were death on cancer drugs, stroke on antithrombotic drugs, and diagnosis of influenza on vaccines, which were defined using the data in DB, indicating that several similar studies have been conducted using the same outcome definition that has been validated.

Among several types of DBs, administrative claim DB and clinical DB were used in many studies. We understand that the most appropriate type of DB was selected in light of the research question. At the same time, the amount of data, the ease of handling as an analysis data set, and the support system to enhance the researchers' accessibility may have influenced the choice of DB. Although there were a few articles of effectiveness studies using pharmacy DB, they were informative in defining effectiveness by drug usage status. There were few articles using DB linkage, in which data from multiple independent DBs are linked, and several others defined outcomes within a single DB despite the limitations of case traceability. However, among DBs classified as a single DB in the present study, there might be some DBs that were linked to multiple DBs such as the Clinical Practice Research Datalink (CPRD) [20]. In DBs in the data set II, the largest number of studies focused on cardiovascular, respiratory, and infectious diseases, followed by oncology because many studies in oncology area were conducted using the registry, which was not included in the data set II. Registries such as the SEER [21] have been enriched by the contributions of regional governments and societies, which are actively used in studies on drug effectiveness.

Concerning the specific DBs used in DB studies, the National Claim DBs of each

country is the most frequently used, and the National DBs of Taiwan and Korea in East Asia had a strong presence. Also, several Japanese DBs such as MDV [22]; JMDC, an insurance claims DB linked to annual health checkup data provided by JMDC Inc. [23]; and DPC DB [18], comprising administrative claims data and discharge information for acute-care inpatients in Japan, were used in several studies. It is also necessary to evaluate the quality of the data contained in the DB and the quality control process, and select a DB.

Most of the study designs were multi-cohort studies with a comparison group; drugs with the same indication, no drug use, or a class of drugs with the same indication were often used. Of course, the comparison group should be set up appropriately according to the research question, but comparisons with individual drugs may result in a smaller number of cases and lower statistical power. We understand that studies using a class of drugs and no drug use as a comparator were conducted to ensure a larger number of cases.

Studies using inpatient diagnosis as an outcome were the most common. Although outcome definitions may be difficult to establish for some diseases, data on impatient diagnosis may be highly reliable as a medical information DB entry. Studies that used laboratory/clinical scores or medical procedures such as surgery as outcomes were limited possibly due to the difficulty in standardizing laboratory values and hence in compiling data in a unified manner, and the lack of regular testing.

A total of 61% (138/225) of the studies used a single type of source to define the outcome. There were studies in which four different outcomes were used, but unexpectedly, few studies combined multiple different outcomes into a single outcome definition (11 studies) . The combination of diagnosis and drug administration was used often regardless

of the disease, thus indicating that although simpler definitions of effectiveness were common, a small number of studies added multiple different outcomes to create more robust outcome definitions based on DB characteristics.

We found some characteristics in the outcome data used in each disease area. Specifically, studies in infectious diseases and respiratory diseases were defined by a variety of outcomes. Death (OS) and drug discontinuation were used as outcomes in the field of oncology; we understand that the uniform operation of disease status examination and drug administration regimens in oncology makes drug discontinuation acceptable as a proxy clinical outcome for PFS. Drug usage status has relatively high reliability in medical information DBs and can be expected to be used as an outcome definition for effectiveness. In our research, 16 studies were defined only by discontinuation/change/addition of drugs associated with antirheumatic, anticancer, antithrombotic, antipsychotic, and allergic diseases. Because medical information DBs (excluding registries) do not usually include data such as disease improvement scores to assess effectiveness, it may be possible to more easily define drug use as an outcome for effectiveness when outcomes cannot be determined by diagnosis. In one study in rheumatology, the relationship between the DAS28 score and six drug use indices such as adherence, bio-formulations, DMARDs, and steroid drug escalation followed up to1 year was validated and used as an index of effectiveness [24]. Another study evaluated the effectiveness of sublingual immunotherapy for allergic rhinitis by scores determined based on the types and frequency of concomitantly prescribed drugs as proxy clinical data [25]. Outcomes for effectiveness would then be possible to be defined using these available data, and there is room for consideration in each disease area, though it may be limited to disease areas where the drug administration regimen is unified to some extent

and there is a wealth of information available.

In recent years, with the use of regulatory schemes such as Accelerated Approval (FDA), an increasing number of drugs are being marketed with less matured safety and efficacy information prior to the approval. As a result, it is becoming increasingly important for pharmaceutical companies to generate post-marketing evidence of effectiveness and safety for their products. In general, post-marketing studies with safety as an outcome are highly necessary and are being conducted actively. However, there have not been many studies on effectiveness using DBs.

Japanese pharmaceutical companies have planned 59 post-marketing DB studies for 39 products (published in the website of the Pharmaceuticals and Medical Devices Agency (PMDA) between July 2017 and September 2020) [26, 27], but there were only two studies that used effectiveness as the outcome. Obviously, appropriate research methods should be selected depending on the research question. However, in Japan, most effectiveness data obtained in post-marketing surveys conducted by the pharmaceutical industry are those from prospective, observational studies with no comparison group regardless of research questions. On the other hand, the difficulty in defining outcomes for drug effectiveness and the lack of data may be one of the reasons why DB studies have not been conducted. The MDV contains DPC data with a discharge summary (and death) and is considered as a usable DB. DB studies on effectiveness using MDV have been conducted in osteoporosis, diabetes, cancer, and antithrombosis areas. It is expected that industry, government, and academia can make active use of Japan's National DB, including DPC data.

In general, the strengths of DB studies are that they are comprehensive in terms of diseases, drugs, tests, and patient populations (comprehensiveness), that data are

collected continuously (continuity), that data from large populations can be collected quickly (rapidity), and that they can be conducted at relatively low cost. In contrast to data from studies that are collected prospectively in accordance with their own objectives, however, RWD, which are collected for other purposes, should be used with the recognition that they have limitations in terms of data quality, clinical relevance, and reliability.

A prospective approach to DB studies by pharmaceutical companies should be considered. For example, post-marketing requirement study for an antithrombotic drug Dabigatran prospectively confirmed its stroke prevention effect using two large US administrative claim DBs for three years (six interim analyses) [28, 29]. The results of retrospective studies are often considered unreliable because their publication was attributed to good results. To dispel this concern, a protocol should be developed prior to the creation of data in real-world setting and prospective data collection from DBs should be planned. Of course, the analysis plan can be fine-tuned after confirming the details of the data, but all changes should be transparent.

Although validation studies are recommended for safety outcome assessment [5, 6], they are often more labor-intensive than the DB study itself, and this is a bottleneck in the implementation of a DB study. Outcome validation was not mentioned in 144 (64.0%) of the studies included in the present study, almost half of which were deaths, laboratory values, and hospitalizations. These outcomes are relatively robust indicators reported by the CPRD in the United Kingdom [30]. Outcome validation is being enriched in the major Japanese DBs: mortality outcomes [31, 32] in JMDC, Charlson and Elixhauser Comorbidity Indices [33], atrial fibrillation [34] in MDV, and even PMDA-managed validation studies in the Medical Information Database NETwork (MID-NET)

[35] and DPC database [36, 37]. When companies conduct validation studies using methods such as chart review, they should consider collaborating with academia.

Among the 225 studies in the present study, 68 studies (30%) were supported by pharmaceutical companies and only two of them were identified as regulatory post-marketing requirement studies. Unfortunately, at this point, DB studies on efficacy/effectiveness do not seem to be sufficiently useful for the reappraisal of marketing authorization. There might be a hurdle in conducting studies on drug effectiveness with a comparison group in company-initiated studies presumably due to the limited certainty of the expected result. Understandably, observational studies produce results with various biases, so appropriate measures need to be taken to reduce the bias, such as new user design, propensity score analysis, and sensitivity analysis. About 75% of the studies took some of these measures. We believe that pharmaceutical companies must generate information that will benefit patients by making fair use of RWD wisely and turn new technologies into value for patients.

This study summarized the status of cross-disease research articles from a large number of articles on DB studies. A limitation of our study is that it is unclear whether the search strategy encompasses all the DB studies on drug effectiveness. It may be possible to make a more detailed analysis if the DBs could be classified from different perspectives (e.g., classification by primary care / secondary care), but detailed information could not be obtained. We did not review the history of previous DB studies, but instead identified studies published in the most recent two years from January 2018 to December 2019. We did not check in detail whether the outcomes of each study were validated.

In conclusion, administrative claim DB and clinical DB were most commonly used in DB studies on drug effectiveness, while pharmacy DB was used in a few studies. The largest number of reported studies was associated with cardiovascular, respiratory, and infectious diseases. Outcomes were often inpatient diagnosis, and some ideas included defining effectiveness based on drug use. Our study also confirmed the limited outcome validation within the articles. Appropriate measures were taken to reduce bias in most of the studies, including new user designs, propensity score analysis, and sensitivity analysis.

3. Research 2　(Key scientific influence of database studies on drug effectiveness)

3.1. Objectives

We analyzed which elements (characteristics) of the studies were related to the level of scientific influence, using the number of citations and citation sources, based on the DB studies extracted in Research 1.

3.2. Methods

Of the 225 DB studies on drug effectiveness identified in Research 1, 19 reports that were published in journals without an impact factor were excluded, leaving 206 reports for the present analysis. Multivariable linear regression analysis was performed, with the number of citations as the response variable and the characteristics of the study as explanatory variables. The explanatory variables were the impact factor, publication year, DB classification, disease area, sample size, outcome type, outcome validation presence, propensity score use, sensitivity analysis implementation, and funding. In addition, we reviewed the citation sources and characterized the reports cited in the clinical practice guidelines.

3.2.1 Characteristics of the study

The number of citations were determined from data from Google Scholar as of January 10, 2022. The impact factors were obtained from the Journal Citation Reports 2020, provided by Clarivate. DBs were classified into the administrative claim database and others (clinical database, pharmacy database, and database linkage). Disease areas were classified according to the International Classification of Diseases tenth revision

(ICD-10), into diseases of the cardiovascular system, which was the most common, and other diseases. The sample size was calculated based on the number of cases analyzed (including the control group). The types of outcomes were categorized into inpatient (admission/discharge) diagnosis, which was the most common, and others (death, laboratory/clinical score, hospitalization/emergency room status, diagnosis, medical procedures, and drug discontinuation/change/addition). Outcome validation was grouped into three categories: implementation of outcome validation mentioned in the report, not mentioned in the report A (the outcome was death, laboratory/clinical score, inpatient diagnosis, or hospitalization/emergency room status), and not mentioned in the report B (the outcome was diagnoses [non-inpatient], medical procedures, or drug usage status). The presence or absence of propensity score usage and sensitivity analysis were confirmed based on the description in the report. Funding information was categorized as funding from pharmaceutical companies or other (including none).

3.2.2 Analyses

We conducted a multivariable linear regression analysis using the “number of citations” as a response variable and the 11 factors mentioned above as exploratory variables. All variables were incorporated into the multivariable model, and the association between explanatory variables was examined using Cramér's V, Spearman's rank correlation coefficient, or the correlation ratio before conducting the regression analysis. A subgroup analysis was conducted for studies on cardiovascular diseases. In this analysis, due to a small number of studies, univariate analyses were conducted in advance, and associated variables ($p<0.1$) were incorporated into the multivariable model. The correlation between the impact factor and the number of citations was examined

using Spearman's rank correlation coefficient. All analyses were performed using StatsDirect (StatsDirect Ltd., Cheshire, UK). A statistically significant association was defined as a p-value < 0.05, if not otherwise specified.

3.3. Result

The characteristics of the 206 reports included in the analysis are presented in Table 4. The median number of citations and the impact factor were 11 and 4.2, respectively. The most common disease area was cardiovascular disease, outcome type was inpatient diagnosis, and bias reduction measures were the use of the propensity score.

Table 4. Description of the study characteristics

Characteristic of the research	Category	Number of studies (%)
Number of citations	Average [median]: 19.9 [11.0]	
Impact factor	Average [median]: 6.2 [4.2]	
Publication year	2018	87 (42.2)
	2019	101 (49.0)
	2020	17 (8.3)
	2021	1 (0.5)
DB classification	Administrative claim database	141 (68.4)
	Clinical database	49 (23.8)
	Pharmacy database	3 (1.5)
	Database linkage	13 (6.3)
Disease area (ICD-10)	I00-I99 Diseases of the circulatory system	82 (39.8)
	J00-J99 Diseases of the respiratory system	23 (11.2)
	A00-B99 Infectious and parasitic diseases	21 (10.2)
	C00-D48 Neoplasms	17 (8.3)
	M00-M99 Diseases of the musculoskeletal system and connective tissue	14 (6.8)
	E00-E90 Endocrine, nutritional, and metabolic diseases	13 (6.3)
	F00-F99 Mental and behavioral disorders	11 (5.8)
	Other (11 disease areas)	25 (12.1)
Sample size	Average [median]: 94 163 [11 873]	
Study design	Cohort study	189 (91.7)
	Single cohort study	12 (5.8)
	Case-control study	3 (1.5)

	Self-control study	2 (1.0)
Comparison group (If there was more than one comparison group within a study, each was counted.)	Specific drug(s) with the same indication	79
	Non-users of the target drug	56
	Class of drugs with the same indication	49
	Different dose(s) of drug	23
	Self-control	2
	No comparison group	13
Type of outcome (If there was more than one outcome within a study, each was counted.)	Inpatient diagnosis name	81
	Diagnosis name	73
	Death	61
	Drug discontinuation/change/addition	45
	Hospitalization/ER status	24
	Laboratory/clinical score	16
	Medical procedures	14
Outcome validation	Not mentioned A (diagnoses, medical procedures, drug usage status, etc.)	36 (17.5)
	Not mentioned B (deaths, laboratory test results, inpatient diagnosis name, cancer registries, linkages, etc.)	93 (45.1)
	Referral of previous research reports on the same outcome	64 (31.1)
	Referring to the appropriateness research paper of the entire DB	8 (3.9)
	Validation was performed in the same study	5 (2.4)
Measures to reduce bias (If there was more than one measure within a study, each was counted.)	Propensity score analysis (matching/stratification/inverse probability weighting)	127
	Sensitivity analysis	87
	New user design	68
	Instrumental variable methods	4
	The above measures have not been taken	47
Funding	Pharmaceutical companies	57 (27.7)
	Other(Government, University, other than pharmaceutical companies)	99 (48.1)
	No funding	44 (21.4)
	Not listed	6 (2.9)

Multivariable linear regression analysis showed that the number of citations was significantly associated with impact factor, publication year, and outcome type (inpatient diagnosis) (Table 5). Similar results were obtained in the subgroup analysis of studies of cardiovascular diseases (Table 5). A positive correlation (correlation coefficient = 0.59)

was found between the number of citations and the impact factor (Figure 5). However, there was no association between the number of citations and the presence or absence of outcome validation, use of propensity scores, or implementation of a sensitivity analysis.

Table 5. Characteristics of studies related to the number of citations (overall/subgroup)

	Overall		Subgroup[a] (Diseases of the circulatory system)	
Explanatory variable	Coefficient	p-value	Coefficient	p-value
Impact factor	2.256	< 0.001	1.865	< 0.001
Publication year	-6.800	0.001	-9.869	0.007
DB classification (others vs. admin DB)	2.352	0.454	-	-
Disease area (ICD-10) (others vs. diseases of the circulatory system)	1.141	0.728		
Sample size	< 0.001	0.097	-	-
Type of outcome (others vs inpatient diagnosis name)	6.255	0.046	11.783	0.018
Outcome validation (Not mentioned A vs. Not mentioned B) (Not mentioned A vs. mentioned)	-6.191 -2.021	0.124 0.633	- 5.003	- 0.291
Propensity score (non-use vs. usage)	0.716	0.814	3.283	0.623
Sensitivity analysis (not conducted vs. conducted)	0.127	0.965	-	-
Funding (others vs. pharmaceutical companies)	4.266	0.175	-	-

[a]Five variables were selected based on univariate analyses

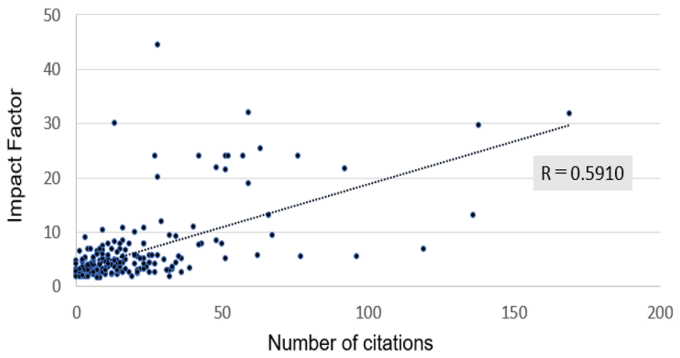

Figure 5　Distribution of the number of citations and impact factor

The 206 reports included in the analysis had approximately 5000 citations (approximately 4000 cited reports and books with English or Japanese titles). Among them, 31 (15 %) reports were cited in 32 clinical practice guidelines and within those, 11 were cited in multiple guidelines. The leading disease areas were circulatory system diseases (18 reports), followed by respiratory system diseases (6 reports). Many of these were large-scale and highly comprehensive studies using national databases (from Korea, Germany, the United States, Taiwan, Finland, etc.) (Table 6). Notable examples of reports cited in clinical practice guidelines are shown in Table 7. These were cited as supporting information, mainly to outcomes with little evidence from interventional studies, based on real-world evidence (RWE).

Table 6　Disease areas and DB names of studies cited in clinical practice guidelines

Disease area (ICD-10 code)/DB name	Number of reports
I00-I99 Diseases of the circulatory system	18
HIRA database (Korea, 11), Truven MarketScan (US, 3), NHIRD (Chinese Taipei, 2), Medicare claims data (US), National Records of Scotland (Scotland)	
J00-J99 Diseases of the respiratory system	6
LRx (German, 2), NHIRD (Chinese Taipei, 2), HIRA database (Korea), Optum (US)	
＜Other disease areas＞ A00-B99 Infectious and parasitic diseases/NHIRD (Chinese Taipei,) F00-F99 Mental and behavioral disorders/Finnish nationwide DB H00-H59 Diseases of the eye and adnexa/JMDC (Japan) K00-K93 Diseases of the digestive system/DPC database (Japan) M00-M99 Diseases of the musculoskeletal system and connective tissue/SIDIAP (Spain) O00-O99 Pregnancy, childbirth, and the puerperium/SNIIRAM (France) R00-R99 Symptoms, signs, and abnormal clinical and laboratory findings/Premier Healthcare (US)	1 (each)

HIRA: Korean Health Insurance Review and Assessment database, NHIRD: Taiwan National Health Insurance Research Database, LRx: German longitudinal prescription database, JMDC: Japan Medical Data Center, SIDIAP: System for the Development of Research in Primary Care, SNIIRAM: French National Health Insurance database

Table 7　Notable examples of reports cited in clinical practice guidelines

Author/Journal	Title/DB name	Clinical practice guidelines (Organization)	Characteristics of the cited evidence
Amgad et al., JAMA 2018	Assessment of outcomes of treatment with oral anticoagulants in patients with atrial fibrillation and multiple chronic conditions: a comparative effectiveness analysis DB: Medicare claims data	Guidelines for outpatient management (American Academy of Family Physicians)	Rationale for recommending the use of oral anticoagulants in patients with moderate to severe disease
Kim et al., Journal of the American College of Cardiology 2019	Outcomes of direct oral anticoagulants in patients with mitral stenosis DB: Korean Health Insurance Review and Assessment (HIRA) database	Guidelines for the diagnosis and management of atrial fibrillation (European Society of Cardiology)	Evidence for patients with mitral stenosis for which there is no large-scale evidence
Zielen et al., Allergy 2018	Sublingual immunotherapy provides long-term relief in allergic rhinitis and reduces the risk of asthma: A retrospective, real-world database analysis DB: German longitudinal prescription database (LRx)	Allergen immunotherapy in children user‘s guide (European Academy of Allergy and Clinical Immunology)	Evidence for sublingual immunotherapy with concomitant drug use defined as a proxy effectiveness clinical outcome
Markku et al., JAMA Psychiatry 2018	Real-world effectiveness of pharmacologic treatments for the prevention of rehospitalization in a Finnish nationwide cohort of patients with bipolar disorder DB: Finnish nationwide databases	The Maudsley guidelines on advanced prescribing in psychosis (book)	Evidence of readmission of patients with moderate or more severe disease who have been hospitalized once

3.4. Discussion

DB studies that used inpatient diagnoses as outcomes have been shown to have high scientific influence. Although clinicians with expertise in disease coding systems and disease areas should be consulted when defining outcomes [9], the use of inpatient diagnosis should first be considered when the outcome is a disease that is likely to result in hospitalization (e.g., myocardial infarction). Although there are regional differences in the coding of clinical data, the U.S. guidelines recommend using disease codes from hospitalization data whenever possible when defining outcomes using claims data [5]. The rationale for this is that ICD codes in claims data at the time of hospitalization are generally more reliable and likely to reflect more serious diseases. As expected, because the impact factor is the ratio of citations to journals, older publications have a greater chance of citation.

The presence or absence of outcome validation, which was thought to enhance the quality of DB studies, was not significantly associated with the number of citations. Outcome validation studies are often time-consuming and costly, and, as shown in Research 1 results, there are few cases. In Japan, access to hospital medical record data without patient consent may violate the Personal Information Protection Law, which is a major obstacle in conducting validation studies [38]. For medical databases under the administration of the Pharmaceuticals and Medical Devices Agency (MID-NET) [35], multiple outcome validation studies have been conducted to increase user convenience [39]. Similar to the results of the outcome validation study, the use of propensity scores was not significantly associated. However, most reports used measures to reduce bias, such as incorporating background factors and other variables into the multivariable model. The use of propensity scores is controversial; propensity score matching is likely to

reduce the number of subjects, and there is no consensus on the balance between accuracy of results and the adjustment for confounders [40, 41].

Thirty-one studies (15 %) were cited in the clinical practice guidelines, and among them, 11 were cited in multiple guidelines. Evidence from DB studies of effectiveness using RWD was considered important in actual practice and was cited primarily as RWE to support the results of interventional studies.

It would be useful to review DB studies across diseases for effectiveness outcome definitions. We identified some cases in which disease occurrence (e.g., confirmation of recurrence) was used as the outcome, and proxy effectiveness evaluations were conducted based on drug usage. In addition, many of the DBs used in such studies were large and comprehensive on a national scale, suggesting the importance of ensuring a sufficient number of cases and covering various patient backgrounds, even after narrowing the study population by disease and drug exposure.

The present study is the first to confirm the scientific influence of a DB study on drug effectiveness by analyzing its citation numbers and sources. This study had some limitations. First, we analyzed reports published in 2018 and 2019, while the number and source of citations change over time. Second, although the multivariable regression analysis confirmed that the disease area (cardiovascular vs. others) did not significantly affect the number of citations, it did not rule out the possibility that the large number of reports on cardiovascular disease, approximately 40 % (82/206), may have influenced the results.

In conclusion, the use of inpatient diagnosis as an outcome of DB studies on drug effectiveness may increase scientific influence (number of citations). Conducting validation studies, using propensity scores, and conducting sensitivity analyses were not

associated with scientific influence but are important to increase the validity of the results of the DB studies. An appropriate setting for effectiveness outcome definitions, such as inpatient diagnosis, disease occurrence (e.g., confirmation of recurrence), and proxy effectiveness evaluations based on drug use, as well as the use of large and comprehensive DBs, are important for generating evidence of post-marketing effectiveness with high scientific influence.

4. Overall discussion

This research is the first to summarize the characteristics and scientific influence of DB studies on drug effectiveness across diseases.

It was indicated that DB studies on efficacy were conducted worldwide for a variety of diseases, and that many of these studies utilized large-scale and comprehensive national DBs. In terms of Japan, we expect that private pharmaceutical companies and others will be able to utilize national DBs in a more user-friendly and timely manner, as with the case overseas. The most common outcome definition of effectiveness was using inpatient diagnoses, and the characteristics of the DBs used and outcome definitions were found to be distinctive by disease area. It is important to increase the internal validity of outcome definitions through various innovations that take into account medical practices.

Some interesting studies that could be recognized through cross-disease investigation were reports that defined drug use as a proxy effectiveness outcome (Machado et al. 2018, Zielen et al. 2018). Note that these two reports were cited in clinical practice guidelines for relevant areas (rheumatoid arthritis and allergic rhinitis) and were considered important evidence. When outcomes cannot be defined by disease name, it is important to devise and set outcome definitions because DBs do not include data such as efficacy scores.

It was suggested that studies using inpatient diagnosis as outcomes have a high scientific influence. If outcome can be defined as a disease with a high probability of hospitalization (e.g., myocardial infarction), the use of inpatient diagnosis should be considered first. On the other hand, the presence or absence of validation studies for the outcomes, which are thought to enhance the quality of DB studies, was not significantly associated with the number of citations. Validation studies are often time-consuming and

costly, and as shown in the results of Research 1, there are only a few examples. We hope that outcome validation studies for each DB will be more active in the future. In the medical DB under the PMDA (MID-NET), multiple outcome validation studies have been conducted to enhance user convenience. As with the implementation of validation studies, the use of propensity scores was not found to be significantly associated, but most reports took measures to reduce bias, such as incorporating background factors and other variables into multivariable analysis models. The application of propensity scores is controversial; for example, propensity score matching is likely to reduce the number of subjects, and no consensus has been reached on its appropriateness.

There were 15% (31/206) of studies cited in clinical practice guidelines, and evidence from DB studies of effectiveness utilizing some RWDs was considered important as information for actual practice. Most of the reports cited in the clinical practice guidelines used large and comprehensive DBs, such as the national DB, suggesting the importance of having a sufficient number of subjects and coverage of patient backgrounds, even after narrowing the study population by disease and drug exposure.

We believe that, with due attention to the strengths and limitations of DB studies, the information on the key characteristics of DB studies on drug effectiveness with high scientific influence obtained in our study would help to promote appropriate DB studies in the post-marketing stage.

5. Conclusion

This study summarized the status of cross-disease research articles on DB studies on drug effectiveness. The use of inpatient diagnosis as an outcome for DB studies on drug effectiveness would increase scientific influence. The appropriate setting for effectiveness outcome definitions, such as disease occurrence, as well as the use of large and comprehensive DBs, is important for generating evidence of post-marketing effectiveness with high scientific influence. While considering the strengths and limitations of DB studies, we believe that our comprehensive results would help to promote appropriate DB studies on drug effectiveness in the post-marketing stage.

6. References

[1] Schad F, Thronicke A. Real-world evidence-current developments and perspectives. Int J Environ Res Public Health. 2022;19(16):10159.

[2] Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol. 2005;58(4):323-337.

[3] U.S. Food & Drug Administration. Real-world data (RWD) and real-world evidence (RWE) are playing an increasing role in health care decisions. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence. Accessed 7 Oct 2022.

[4] U.S. Food & Drug Administration. Food & Drug Administration Work Plan and Proposed Funding Allocations of FDA Innovation Account –Required by Section 1002 of the 21st Century Cures Act 6 (Public Law 114-225). 2017. https://www.fda.gov/files/Proposed-FDA-Work-Plan-for-21st-Century-Cures-Act-Innovation-Account-Activities-as-Submitted-to-Science-Board.pdf Accessed 25 Sep 2021.

[5] U.S. Food & Drug Administration. Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data. 2013. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/best-practices-conducting-and-reporting-pharmacoepidemiologic-safety-studies-using-electronic Accessed 25 Sep 2021.

[6] European Medicines Agency. Guideline on good pharmacovigilance practices(GVP):Module VIII-Post-authorisation safety studies(Rev 3). 2017. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf Accessed 25 Sep 2021.

[7] Pharmaceuticals and Medical Devices Agency. Guidelines for Conducting Drug Epidemiology Research in Evaluating Drug Safety Using Medical Information Databases. 2014 (In Japanese). https://www.pmda.go.jp/safety/surveillance-analysis/0032.html Accessed 25 Sep 2021.

[8] European Medicines Agency. Scientific Guidance on Post-authorisation Efficacy Studies. 2016. https://www.ema.europa.eu/en/documents/scientific-guideline/scientific-guidance-post-authorisation-efficacy-studies-first-version_en.pdf Accessed 26 Sep 2021.

[9] The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) guide on methodological standards in pharmacoepidemiology (Revision 10). 2022. https://www.encepp.eu/standards_and_guidances/documents/01.ENCePPMethodsGuideRev.10_Final.pdf. Accessed 7 Oct 2022.

[10] Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, Madigan D, Makady A, Schneeweiss S, Tarricone R, Wang SV, Watkins J, Daniel Mullins C. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. Pharmacoepidemiol Drug Saf. 2017 Sep;26(9):1033-1039.

[11] Moed HF. New developments in the use of citation analysis in research evaluation.

Arch Immunol Ther Exp(Warsz). 2009;57(1):13-18.

[12] Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl

Acad SciUSA. 2005;102(46):16569-16572.

[13] Santangelo GM. Article-level assessment of influence and translation in biomedical

research. Mol Biol Cell. 2017;28(11):1401-1408.

[14] Weycker D, Li X, Wygant GD, Lee T, Hamilton M, Luo X, Vo L, Mardekian J, Pan X, Burns L, Atwood M, Hanau A, Cohen AT. Effectiveness and Safety of Apixaban versus Warfarin as Outpatient Treatment of Venous Thromboembolism in U.S. Clinical Practice. Thromb Haemost. 2018 Nov;118(11):1951-1961.

[15] Li X, Keshishian A, Hamilton M, Horblyuk R, Gupta K, Luo X, Mardekian J, Friend K, Nadkarni A, Pan X, Lip GYH, Deitelzweig S. Apixaban 5 and 2.5 mg twice-daily versus warfarin for stroke prevention in nonvalvular atrial fibrillation patients: Comparative effectiveness and safety evaluated using a propensity-score-matched approach. PLoS One. 2018 Jan 26;13(1):e0191722.

[16] Ito Y, Van Schyndle J, Nishimura T, Sugitani T, Kimura T. Real-World Effectiveness of Sodium Glucose Co-Transporter-2 Inhibitors in Japanese Patients with Diabetes Mellitus. Diabetes Ther. 2019 Dec;10(6):2219-2231.

[17] Ho JC, Hsieh ML, Chuang PH, Hsieh VC. Cost-Effectiveness of Sorafenib Monotherapy and Selected Combination Therapy with Sorafenib in Patients with Advanced Hepatocellular Carcinoma. Value Health Reg Issues. 2018 May;15:120-126.

[18] Nakamura K. Diagnosis Procedure Combination Database Would Develop

Nationwide Clinical Research in Japan. Circ J. 2016 Oct 25;80(11):2289-2290.

[19] Enomoto M, Endo A, Yatsushige H, Fushimi K, Otomo Y. Clinical Effects of Early Edaravone Use in Acute Ischemic Stroke Patients Treated by Endovascular Reperfusion Therapy. Stroke. 2019 Mar;50(3):652-658.

[20] Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, Smeeth L. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015 Jun;44(3):827-36.

[21] Duggan MA, Anderson WF, Altekruse S, Penberthy L, Sherman ME. The Surveillance, Epidemiology, and End Results (SEER) Program and Pathology: Toward Strengthening the Critical Relationship. Am J Surg Pathol. 2016 Dec;40(12):e94-e102.

[22] Introducing MDV Database: Medical Data Vision Co., Ltd. Available from:

https://www.mdv.co.jp/solution/pharmaceutical/english/ Accessed 25 Sep 2021.

[23] Introducing JMDC Database: JMDC Inc. Available from:

https://www.jmdc.co.jp/en/jmdc-claims-database/ Accessed 25 Sep 2021.

[24] Machado MAÁ, Moura CS, Guerra SF, Curtis JR, Abrahamowicz M, Bernatsky S. Effectiveness and safety of tofacitinib in rheumatoid arthritis: a cohort study. Arthritis Res Ther. 2018; https://doi.org/10.1186/s13075-018-1539-6

[25] Zielen S, Devillier P, Heinrich J, Richter H, Wahn U. Sublingual immunotherapy provides long-term relief in allergic rhinitis and reduces the risk of asthma: A retrospective, real-world database analysis. Allergy. 2018 Jan;73(1):165-177.

[26] Pharmaceuticals and Medical Devices Agency. Homepage for List of RMP Submissions (In Japanese). https://www.pmda.go.jp/safety/info-services/drugs/items-information/rmp/0001.html. Accessed 7 Oct 2022.

[27] EPS Corporation. In: Post-marketing Database Survey and Product List (In Japanese).

https://www.eps.co.jp/ja/service_ermp-update.php. Accessed 7 Oct 2022.

[28] Schneeweiss S, Gopalakrishnan C, Bartels DB, Franklin JM, Zint K, Kulldorff M, Huybrechts KF. Sequential Monitoring of the Comparative Effectiveness and Safety of Dabigatran in Routine Care. Circ Cardiovasc Qual Outcomes. 2019; https://doi.org/10.1161/CIRCOUTCOMES.118.005173

[29] Huybrechts KF, Gopalakrishnan C, Bartels DB, Zint K, Gurusamy VK, Landon J, Schneeweiss S. Safety and Effectiveness of Dabigatran and Other Direct Oral

Anticoagulants Compared With Warfarin in Patients With Atrial Fibrillation. Clin Pharmacol Ther. 2020;107(6):1405-1419.

[30] Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. Br J Gen Pract. 2010;60(572):e128-36.

[31] Ooba N, Setoguchi S, Ando T, Sato T, Yamaguchi T, Mochizuki M, Kubota K. Claims-based definition of death in Japanese claims database: validity and implications. PLoS One. 2013;31.8(5):e66116.

[32] Sakai M, Ohtera S, Iwao T, Neff Y, Kato G, Takahashi Y, Nakayama T; BiDAME (Big Data Analysis of Medical care for the Elderly in Kyoto). Validation of claims data to identify death among aged persons utilizing enrollment data from health insurance unions. Environ Health Prev Med. 2019;24(1):63.

[33] Kimura T, Sugitani T, Nishimura T, Ito M, Validation and Recalibration of Charlson and Elixhauser Comorbidity Indices Based on Data From a Japanese Insurance Claims Database. Japanese Journal of Pharmacoepidemiology/Yakuzai ekigaku. 2019;24(2):53-64.

[34] Koretsune Y, Yamashita T, Yasaka M, Oda E, Matsubayashi D, Ota K, Kobayashi M, Matsushita Y, Kaburagi J, Ibusuki K, Takita A, Iwashita M, Yamaguchi T. Usefulness of a healthcare database for epidemiological research in atrial fibrillation. J Cardiol. 2017 Aug;70(2):169-179.

[35] Pharmaceuticals and Medical Devices Agency. In: Basic information for MID-NET users: List of outcomes to be considered (In Japanese). https://www.pmda.go.jp/files/000225081.pdf. Accessed 7 Oct 2022.

[36] Yamana H, Moriwaki M, Horiguchi H, Kodan M, Fushimi K, Yasunaga H. Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. J Epidemiol. 2017;27(10):476-482.

[37] Iwamoto M, Higashi T, Miura H, Kawaguchi T, Tanaka S, Yamashita I, Yoshimoto T, Yoshida S, Matoba M. Accuracy of using Diagnosis Procedure Combination administrative claims data for estimating the amount of opioid consumption among cancer patients in Japan. Jpn J Clin Oncol. 2015;45(11):1036-41.

[38] Iwagami M et al. Task force report on the validation of diagnosis codes and other outcome definitions in the Japanese receipt data (In Japanese). Japanese Journal of Pharmacoepidemiology/Yakuzai ekigaku. 2018;23(2):95-123.

[39] Pharmaceuticals and Medical Devices Agency. Homepage for List of outcome validation. https://www.pmda.go.jp/files/000225081.pdf. Accessed 7 Oct 2022.

[40] Joffe MM, Rosenbaum PR. Invited commentary: propensity scores. Am J Epidemiol. 199915;150(4):327-333

[41] Williamson E, Morley R, Lucas A, Carpenter J. Propensity scores: from naive enthusiasm to intuitive understanding. Stat Methods Med Res. 2012;21(3):273-393

7. Acknowledgement

I would like to express my deep and sincere gratitude to Professor Mamoru Narukawa for giving me the opportunity to do my research and providing invaluable guidance throughout this study. I am also grateful to Assistant Professor Saki Aino for her assistance on my research.

I would like to thank Professor Rie Kubota, Professor Nobuhiko Okada, and Lecturer Makiko Iwasawa, for their reviews and helpful discussion on this study.
I would like to offer my special thanks to Ms. Yukiko Minami for her generous support on my student life at Kitasato university. Also, my special appreciation goes to the classmates at department of Clinical Medicine (Pharmaceutical Medicine). I have learned a lot from them and they have been motivating and encouraging me. Their support and inspiration made my days exceptional moment. I wish the best of luck for their research and career.

Lastly, I would like to express my profound gratitude to my family who have been supporting and encouraging me throughout the years at Kitasato university. Their moral and continuing support considerably helped me keep myself motivated.