PUBLIC CONSULTATION VERSION

2020.10.01

Use of Real-World Data and Real-World Evidence to Support Drug Reimbursement Decision-Making in Asia.

A non-binding guidance document prepared by the **REAL** World Data In A**S**ia for H**E**alth Technology Assessment in Reimbursement (**REALISE**) working group

NUS National University of Singapore | Saw Swee Hock School of Public Health

HITAP Health Intervention and Technology Assessment Program

HTAsiaLink

iDSI Better decisions. Better health.

GUIDANCE DOCUMENT

USE OF REAL-WORLD DATA AND REAL-WORLD EVIDENCE TO SUPPORT DRUG REIMBURSEMENT DECISION-MAKING IN ASIA

CONTRIBUTING AUTHORS

This version of the guidance has undergone review by:

Jeonghoon AHN
Ewha Womans University, Seoul, South Korea

Diana Beatriz S. BAYANI
Saw Swee Hock School of Public Health (SSHSPH), National University of Singapore, Singapore

Kelvin CHAN
Sunnybrook Odette Cancer Centre, Canada
Sunnybrook Research Institute, Canada
Canadian Centre for Applied Research in Cancer Control, Canada

Dechen CHOIPHEL
Essential Medicine and Technology Division, Department of Medical Services, Ministry of Health, Bhutan

Wanrudee ISARANUWATCHAI
Health Intervention and Technology Assessment Program (HITAP), Ministry of Health, Thailand
Centre for Excellence in Economic Analysis Research, St. Michael's Hospital, Canada
Institute of Health Policy, Management and Evaluation, University of Toronto, Canada

Sarin KC
HITAP, Ministry of Health, Thailand

Brendon KEARNEY
Faculty of Medicine, University of Adelaide, Australia

Lydia Wenxin LIN
Saw Swee Hock School of Public Health, National University of Singapore, Singapore

Jing LOU
Saw Swee Hock School of Public Health, National University of Singapore, Singapore

Ryota NAKAMURA
Hitotsubashi Institute for Advanced Study, Hitotsubashi University, Japan

Fiona PEARCE
Agency for Care Effectiveness, Ministry of Health, Singapore

Shankar PRINJA
School of Public Health, Post Graduate Institute of Medical Education and Research, India

Raoh-Fang PWU
National Hepatitis C Program Office, Ministry of Health and Welfare, Taiwan

Asrul Akmal SHAFIE
Discipline of Social and Administrative Pharmacy, School of Pharmaceutical Sciences, Universiti Sains Malaysia, Malaysia

Yot TEERAWATTANANON
HITAP, Ministry of Health, Thailand

Sean TUNIS
Center for Medical Technology Policy (CMTP), United States

Hwee-Lin WEE
SSHSPH, National University of Singapore, Singapore

Hui-Min WU
National Hepatitis C Program Office, Ministry of Health and Welfare, Taiwan

John ZALCBERG
Cancer Research Program, School of Public Health and Preventive Medicine, Monash University, Australia
Cancer Research at the Alfred Hospital, Australia

NUS National University of Singapore | Saw Swee Hock School of Public Health

HITAP Health Intervention and Technology Assessment Program

HTAsiaLink

iDSI Better decisions. Better health.

GUIDANCE DOCUMENT

USE OF REAL-WORLD DATA AND REAL-WORLD EVIDENCE TO SUPPORT DRUG REIMBURSEMENT DECISION-MAKING IN ASIA

REALISE PROJECT MEMBERS

REALISE working group members

Jeonghoon AHN, Dechen CHOIPHEL, Anne Julienne GENUINO, Anna Melissa GUERRERO, Budi HIDAYAT, Yuehua LIU, Mardiati NADJIB, Ryota NAKAMURA (Involved in first draft of Theme 3), Fiona PEARCE, Shankar PRINJA, Raoh-Fang PWU, Asrul Akmal SHAFIE, Binyan SUI, Auliya SUWANTIKA, Hui-Min WU, Kun ZHAO

REALISE international advisory panel

Amanda ADLER, Kelvin CHAN (Involved in first draft of Theme 3), Brendon KEARNEY, Sean TUNIS, John ZALCBERG

REALISE core team members

Diana Beatriz S. BAYANI (Led Theme 3), Brandon CHUA, Sarin KC (Co-led Theme 1), Lydia Wenxin LIN (Co-Led Theme 1, Led Theme 2), Jing LOU (Co-led Theme 1);

Wanrudee ISARANUWATCHAI, Yot TEERAWATTANANON, Hwee-Lin WEE (Eds.)

NUS National University of Singapore | Saw Swee Hock School of Public Health

HITAP Health Intervention and Technology Assessment Program

HTAsiaLink

iDSI Better decisions. Better health.

Table of Contents

Table of Figures and Tables

Table of Boxes

1 Introduction

1.1. Purpose: What do we want to achieve?

A collaboration between global experts and leaders from health technology assessment (HTA) agencies in Asia, the **REAL** World Data **I**n A**S**ia for H**E**alth Technology Assessment in Reimbursement (**REALISE**) working group seeks to develop non-binding guidance that will provide a framework to generate and use real-world data (RWD) / real-world evidence (RWE) in a consistent and efficient manner for decision-making in Asia. The acronym REALISE signifies our desire to realize ('to cause to happen or to facilitate') the potential of RWD/RWE while realizing ('being aware of') its strengths and limitations. The issues to be addressed in the guidance document will include but are not limited to: (a) When is it appropriate to consider RWD/RWE for reimbursement decisions?; (b) What types of RWD should we collect?; (c) What are the data sources for collecting RWD?; (d) How should we collect RWD?; (e) Who should collect RWD?; (f) How will RWD be used to generate RWE?; (g) How should we use RWE in decision making?; (h) What are the potential biases and how to deal with these biases?; and (i) What are the ethical considerations in collecting RWD and generating RWE?

It is our goal that the proposed guidance document will increase the quality of RWD/RWE collected and used in HTA. However, we recognize that the actual implementation of this guidance document will vary from country to country due to many reasons including capacity constraints, lack of political support, and local legislation. That is, each health system will have its own practical barriers in utilizing RWD and hence, we propose that all recommendations in the guidance document are non-binding in nature to ensure that users can adapt the contents to their local needs.

1.2. Background

1.2.1. RWD and RWE: Definitions

There is growing interest globally in using real-world data (RWD) and real-world evidence (RWE) for regulatory and reimbursement decision-making for health technologies. This is because RWD, defined as *data collected during routine delivery of health care (e.g. from observational studies, electronic medical records (EMR), claims and billing activities, product and disease registries, patient-generated data),*[1 2] and RWE, defined as *evidence that is derived from the analysis of RWD*,[2 3] have shown several potential benefits in informing health-related decision-making. We adopt these definitions from the HTA glossary (htaglossary.net), a collaboration between the International Network of Agencies for Health Technology Assessment (INAHTA), HTA international (HTAi), and other partners to develop a common vocabulary for work in HTA (Box 1.1).[2 4] Benefits of RWE in decision-making include, but are not limited to, reducing time and cost to source relevant information to inform an HTA if population-specific data are required and sufficient local evidence is lacking from available trials,[5] providing evidence with higher external validity compared to randomized controlled trials (RCTs) (see Box 1.2), giving decision makers more certainty of the safety, effectiveness, and cost-effectiveness of technologies in the local setting,[6] and filling the information gap in the absence of clinical trials (e.g. when it is not feasible or ethical to conduct a trial, or there is significant unmet need).[7]

Box 1.1. Differentiating RWD and RWE[8]

In a report by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Task Force created to make recommendations on RWD studies, they state, "The notion was that *data* conjures the idea of simple factual information, whereas *evidence* connotes the organization of the information to inform a conclusion or judgment. Evidence is generated according to a research plan and interpreted accordingly, whereas data is but one component of the research plan. Evidence is shaped, while data simply are raw materials and alone are noninformative."

Box 1.2. Importance of external validity

Constrained study designs with strict inclusion and exclusion criteria for the study population are typically adopted for RCTs, based on severity of illness, comorbidities, use of other medications, and adherence to protocols.[9] This means that the drug is tested for safety and efficacy on a small, non-representative segment of the population.[10] RCTs generally produce more favorable outcomes than observed in real world settings due to lower rates of discontinuation, more frequent protocol driven visits, and exclusion of patients with comorbidities.[9] External validity matters because safety, and effectiveness of a drug over the longer term is what ultimately counts toward drug cost for payers, and in the assessment of its value for money for the payer's specific population.

1.2.2. RWD and RWE: The global context

RWE use in health care decision-making is not new. Regulators have been using routine data to monitor safety in Europe and the US for many years. As an example, the European Medicines Agency used registry and claims data from Denmark and the UK from 2001-2011 to quantify the risk of lactic acidosis following metformin use among patients according to renal function.[11] The contraindications on the product label were consequently modified based on this study, rather than requiring the manufacturers to conduct an expensive post-marketing trial. The United States (US) Food and Drug Administration's (FDA) Sentinel initiative, launched in May 2008, is another example of a pharmacovigilance program that assesses personal health data of over 223 million US residents to monitor the safety of approved drugs.[12] More recently in 2016, the 21st Century Cures Act required the FDA to develop a framework and guidance for evaluating RWE in the US for regulatory purposes, to standardize the use of RWE to inform regulatory approval of new indications for drugs, and to support post-approval requirements.[13]

In reimbursement and coverage decisions, RWE is increasingly recognized as a tool for accelerated access programs in several European countries. United Kingdom's (UK) National Institute for Health and Care Excellence (NICE) has conditional reimbursement schemes using RWE.[14] US payers (e.g. insurance providers) often use epidemiological data based on claims data, to estimate the proportion of patients that are likely to claim for treatment. Health Canada accepts the use of all relevant data, including RWE, as evidence for a drug's efficacy and safety and does not limit its study design.[15]

1.2.3. RWD and RWE: Early versus late reimbursement in the Asian context

HTA agencies in Asia have used RWE (or sometimes referred to as 'local evidence') to inform coverage decisions in the past without explicit and formal methodological guidelines in place. However, optimal collection, analysis and use of RWD/RWE to inform HTA requires a conceptual framework to standardize

processes and ensure consistency. Such a framework is currently lacking in Asia, a region that is likely to benefit from RWD/RWE.

RWD are particularly relevant in Asia where there is often a greater reliance on clinical effectiveness data from routine healthcare data sources (such as observational studies or disease registries) for regulatory and reimbursement purposes than in the United States or Europe for two reasons. First, only around 17% of the clinical trials are conducted in Asia[16] due to barriers related to financial and human capacity, ethical and regulatory systems, lack of research environment, and operational issues.[17] Second, there could be an under-representation of Asian populations in pivotal clinical trials.[16 18] In some Asian health systems such as Singapore, Taiwan and South Korea, electronic medical records (EMR) may also be used to generate local clinical effectiveness data. These data are important to demonstrate the efficacy and safety of medical treatments despite biological variations (e.g. because of differences in body weight or pharmacokinetics and/or pharmacodynamics due to different genetic makeups between Caucasians and Asians),[19] and non-biological variations (e.g. clinical trial findings among Caucasians may not be readily generalizable to Asians) seen in the patient populations. At the same time, in view of under-representation of Asians in clinical trials, it is usually unfeasible for most Asian health systems to replicate the RCTs in their local contexts, due to financial, capacity, and resource limitations, thus increasing the potential value of RWD/RWE in estimating the benefits and risks of therapies in Asian populations. Furthermore, there may be differences in local clinical practice guidelines driven by budget and resource constraints. For example, in health systems with larger budgets such as the UK,[20] the use of high cost biologic agents as first- or second-line therapies for rheumatoid arthritis is recommended in line with their registered indications, supported by clinical trial data. However, in Thailand, due to concerns over the sustainability of reimbursing these high cost drugs, biologic agents are only recommended as third line therapies for rheumatoid arthritis.[21] Therefore, the results from trials conducted in other health systems may not be easily generalizable to countries where these agents are used in a different line of therapy in local clinical practice.

In addition, in many Asian health systems (e.g. China, India, Indonesia, Malaysia, Philippines, Singapore, and Thailand), reimbursement decisions are currently made up to several years after market entry. In this time, drugs can be prescribed by physicians and are paid for like any other non-subsidized drugs, out of pocket or through private insurance coverage. The delay to reimbursement provides these health systems with an opportunity to accumulate local clinical effectiveness data from other RWD sources to inform subsequent decision-making. This not only provides more certainty around the likely effect of the technology in the local population, but has the additional benefit of allowing longer-term effectiveness and safety data to be collected beyond the initial clinical trial period, which is particularly relevant for technologies where adverse events may take time to develop or are so rare that they are not detected until a sufficiently large number of patients have used the technology. In other Asian health systems (e.g. Japan, Taiwan and South Korea), reimbursement decisions coincide with or closely follow the timing of market entry shortly after regulatory approval. In these health systems, RWD and RWE are usually considered when re-assessing initial funding decisions or for price adjustment. Regardless of the timing, RWD and RWE have important roles to play in reimbursement decisions. Hence, RWD collected in these instances need to be carefully managed and analyzed. An alignment of RWD/RWE policies across Asia would equip decision makers with context-relevant evidence and improve timely patient access to new technologies.

1.3. The REALISE Working Group: Who are we?

1.3.1. Scope and position

The REALISE working group regards RWD and RWE as complementary to RCT, the current gold standard for generating evidence on treatment efficacy.

The approach for this guidance document, given the interest area and experience of the REALISE working group, is to focus on the use of RWD and RWE to inform drug assessments. Other technologies where HTA is applicable, such as medical devices or companion diagnostics, are not covered by this guidance. This document, which is the beginning of a series of projects the REALISE working group will be undertaking is intended to be a living document that will be updated over time as new approaches to optimize the generation and use of RWD and RWE emerge.

1.3.2. Organizational structure

The REALISE working group comprises three subgroups: the (a) International Advisory Panel (IAP), (b) HTAsiaLink working group, and (c) Core Team. The IAP are prominent experts from leading HTA organizations in Australia, Canada, the UK and the US, where the use of RWD/RWE in HTA is already established. They provide guidance on how RWD/RWE are collected, analyzed and assessed in their countries. The HTAsiaLink working group includes representatives from 11 Asian health systems (Bhutan, China, India, Indonesia, Japan, Malaysia, Philippines, Singapore, South Korea, Taiwan and Thailand), who share their experiences and perspectives on opportunities and challenges in using RWD/RWE in their local contexts. The core team comprises staff from Saw Swee Hock School of Public Health (SSHSPH), National University of Singapore (NUS), and Health Intervention and Technology Assessment Program (HITAP), Ministry of Health, Thailand. Figure 1.1 shows the organizational chart of the WG.

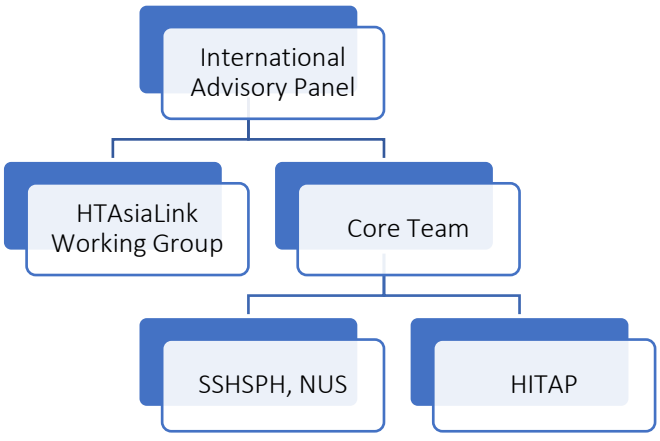

Figure 1.1. REALISE organizational chart

1.3.3. Grant information

This work is supported by an unrestricted grant from The International Decision Support Initiative (iDSI, www.idsihealth.org), a global network of health, policy and economic expertise, working to achieve Universal Health Coverage and the health Sustainable Development Goal (SDG 3), and which supports countries to get the best value for money from healthy spending. iDSI receives funding support from the Bill & Melinda Gates Foundation, the UK Department for International Development, and the Rockefeller Foundation. The funders had no role in study design, data collection and analysis, decision

to publish, or preparation of the manuscript. The findings, interpretations and conclusions expressed in this document do not necessarily reflect the views of the aforementioned funding agencies.

1.4. Content development: How were the contents of this guidance document developed?

In preparing this guidance document, we supplemented a review of the literature with in-person meetings, stakeholder surveys, interviews with country representatives, and teleconferences with the working group representatives.

The literature review was designed to be pragmatic rather than exhaustive, and was used to identify key papers and examples for illustration, rather than to identify all papers on RWD/RWE. The working group had two in-person meetings to deliberate on the scope and content for the document; the first meeting was in April 2019 following the 8th HTAsiaLink conference and the second meeting was a 2-day symposium in Singapore in October 2019. A survey circulated to REALISE members focused on: (a) background of respondent; (b) current practice with regards to the use of RWD/RWE for HTA for reimbursement decisions; (c) current practice with regards to pragmatic clinical trials; (d) challenges encountered in RWD/RWE generation; and (e) availability of a local guidance document on RWD/RWE generation. Working group members were also invited for an hour-long interview to understand the health care context in their individual countries and how RWD is collected and used, with 8 countries interviewed in 2019. Regular teleconferences for the collective group were held to gather opinions on the document and also individually on select topics (e.g. to obtain country examples, consult on specific themes of the document, etc.).

1.5. Structure of document: How is this guidance document structured?

We have organized this report into three themes, using examples from Asia whenever possible throughout the document. Theme 1 describes the scenarios under which the use of RWD and RWE are appropriate to inform HTA. Theme 2 continues with how RWD and RWE may be collected in the Asian context including the RWD types, data sources, and study design; while Theme 3 describes how to translate RWD to RWE while also accounting for biases, confounding, missing data and medication non-adherence.

2 Theme one: Scenarios to use RWD/RWE

When is it appropriate to consider RWD/RWE for reimbursement decisions?

2.1. Lack of data

More innovative drugs are entering the market using a combination of data from RCTs and observational studies.[14] In some cases, study sizes decrease as certain patient populations are small to begin with, or as medicine becomes more targeted and personalized. This theme is concerned not only with the absence of data but also when the available data is of low quality and therefore unreliable. RWD in such cases can be utilized as supplementary evidence to RCTs and can enhance decision-making.

2.1.1. RCT data

RWD and RWE may be considered if current RCT and systematic reviews of RCTs are lacking or of insufficient quality to inform decision-making. Factors that lead to poor quality RCTs and systematic reviews include small numbers of patients involved, relatively short follow-up, outcomes that were incomplete or poorly captured, studies that were underpowered, studies with limited external validity (especially for high risk patient groups who are excluded from RCTs such as pediatrics and geriatrics), and inappropriate synthesis of data in systematic reviews. An example of the use of RWD to compensate for poor quality pivotal trial evidence may be found in the NICE technology appraisal guidance of percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral compression fractures (TA279; Box 2.1).[22]

Box 2.1. Using RWD when RCT data is lacking: Treating osteoporotic vertebral compression fractures

Percutaneous vertebroplasty and percutaneous balloon kyphoplasty are minimally invasive procedures used to treat spinal compression fractures. They were evaluated with RCTs which measured the efficacy of intervention in reducing mortality from osteoporotic vertebral fractures. Based on a meta-analysis of 3 RCTs over 12 months, the benefit on treatment on mortality had a hazard ratio of 0.68, with no statistical significance.[22] However, the RCT result was initially considered to be uncertain given the 3 RCTs were very small studies, with only 276 patients included in total. RWD, with a much larger sample size, is accepted by NICE to support data from small trials. To confirm the efficacy results, the hazard ratio from the trial data was compared to available RWD and was found to be close to the findings from two large observational datasets. First, the US Medicare Registry, which included 858,979 patients with newly diagnosed vertebral fracture with 4 years follow up, reported a hazard ratio of 0.63 in the treatment group. Second, the German Health Insurance Fund, which recorded data for3,607 patients with vertebral fractures, reported a hazard ratio of 0.57. Both registry data were not designed to test for mortality, but the results were nonetheless in concordance with the trial data.[22] However, it should be noted that there is still a possibility that the agreement between the trial data and the observational studies may be due to chance.

Nonetheless, it is essential that RWD and RWE used for national reimbursement decisions represent the target clinical population reflective of the local context, so that the RWD and RWE can be generalized to routine clinical practice. Unfortunately, there are few data sources, including EMRs and

claims data, that truly reflect the local population and researchers need to consider the systematic biases that may be present in using certain databases. Using multiple datasets reduces the likelihood of biases. In the Asian context, a list of some of the national-level databases and other sources of RWD available in each country can be found in Appendix 8.1. The representativeness of each source with regards to the local patient population and local clinical practice is variable in each country. An example of RWD that is representative of 99.9% of the population can be found in Taiwan from the Health and Welfare Data Center (HWDC)[23] which centralizes most health-related databases in Taiwan, including comprehensive cross-linkage to claims data, registries, and national surveys.

2.1.2. Rare diseases

Rare diseases are most frequently cited as an area where RWD and RWE need to be collected in the absence of sufficient trial evidence. Conducting RCTs for rare diseases is particularly challenging because of the small numbers of patients available for recruitment, the high variability in clinical presentation and prognosis across patients with the same condition, difficulties in accurately diagnosing patients with specific conditions, and a lack of a consistent definition for what constitutes a rare disease in each country. In the Asian context, South Korea considers a rare disease as a condition where there are fewer than 20,000 patients, or for which the prevalence is unknown owing to difficulties in diagnosing the condition, or that are designated by the procedures and standards set by the Ministry of Health and Welfare.[24] In Singapore, rare disease is defined as <4 in 10,000 people, and ultra-rare is <2 in 50,000 people.[25] Thailand defines a rare disease by the "size of population affected by disease", where a prevalence of 10,000 is considered as the threshold.[26] There is no official definition in Malaysia but the Malaysia Rare Disease Society defines it as 1 in 4000.

Typically, there is no standard of care or treatment for most rare diseases, therefore, comparative trials are typically not feasible. RWD/RWE from patients' medical records and rare disease registries are becoming crucial to demonstrate orphan drugs' long-term safety and efficacy for regulatory and HTA purposes.[14] Several countries and interest groups have advocated for creating a centralized national registry of rare diseases to serve investigators conducting rare disease research, as well as other stakeholder groups (families, clinicians, manufacturers) who may stand to benefit from the repository of information available.[27] To assist with the formation of such a database, the definition of what constitutes a rare disease has to be clearly specified. Individual groups can then deposit de-identified data using a standardized template for diseases that meet the agreed definition.

2.1.3. Surrogate versus final end-points

This working group recognizes that clinical trials cannot continue indefinitely, and it may not always be feasible to capture ultimate endpoints such as overall survival or long-term health-related quality of life (HRQoL). If the manufacturer intends to make a claim regarding the long-term efficacy of the drug based on a surrogate endpoint, the link between the surrogate endpoint and the long-term outcome (e.g. overall survival or long-term HRQoL) would need to be demonstrated. Under such a scenario, well-conducted observational studies that provide convincing evidence for the link between surrogate endpoints and longer term endpoints should be considered (Box 2.2).

Box 2.2. Using RWD in linking surrogate and final end-points: Gastrointestinal stromal tumors

A meta-analysis of 14 RCTs and 5 observational studies of sufficient methodologic quality in patients with unresectable and/or metastatic gastrointestinal stromal tumors and found a strong positive relationship between overall survival (OS) and progression free survival (PFS; served as a surrogate endpoint), especially in later lines of therapy.[28] These findings suggest that PFS could serve as a surrogate marker for OS for this cancer type, although further patient-level data analyses are needed to strengthen its validity. In rare diseases, and for many cancer types, PFS is already reported to be a commonly accepted primary end-point and proxy for longer-term survival benefits of treatment, given the paucity of long-term data.[29]

Surrogate endpoints however, should not be regarded as a replacement for a final endpoint. We advise caution when using RWD/RWE to demonstrate the link between a surrogate and a specific endpoint. A strong sensitivity analysis is recommended to validate the causal link and to omit spurious effects between the surrogate and the endpoint, otherwise, this may lead to false positive conclusions. One ideal method is the analysis of multiple studies of known effective drugs, which assess both the direct and surrogate endpoints, in order to establish and quantitate the relationship.[30] This need not be performed for all surrogate endpoints as regulators such as the US FDA have already established a list of surrogate endpoints that are considered valid for use in drug approval.

2.2. Contextualizing

This section discusses the importance of ensuring that data retrieved is relevant to the local context and situation that it is being used for. Common examples where RWD may be useful include validating the choice of survival curve drawn from studies conducted in other settings, utilizing local data for parameters into economic models, and re-evaluating reimbursement decisions based on data collected retrospectively that have been drawn from the local population/users of the reimbursed drug.

2.2.1. Extrapolating beyond RCTs

In HTAs conducted for drugs, it is common to extrapolate efficacy estimates from short-term trials over the course of the patient's lifetime. The difference in survival between the treatment and control groups is an important measure of clinical efficacy. However, most trials are too short to include accurate information on how long all patients are likely to survive (with and without treatment). There are a variety of statistical methods that can be applied to extrapolate survival beyond a trial's duration, and the predicted differences in survival depend on which statistical extrapolation is used. When survival curves need to be extrapolated, RWD can be used to provide "validation" and determine if the extrapolation method used was appropriate and if results are likely to be clinically plausible (Boxes 2.3-2.4).[31 32]

The biggest caveat in validating the extrapolation of survival is that RWD comprises largely of data from people who have lived a sufficiently long time (survivor bias). The RWD would tend to validate the curve that looks the best when it is in fact invalid to use the RWD to validate the RCT, because the survivor group is different from the trial population and also different from the patients in routine clinical practice.[33] Therefore, despite the use of RWD in practice for such validation, the HTA community agrees that RWD can be used to supplement and not validate clinical trial data as most trials overestimate the

true effect.[33 34] The very premise for using RWD is that the real-world and trial populations are different, contributing to what is known as the 'efficacy-effectiveness gap'[34] and there is no easy way to judge whether the drug will work better, as well, or worse in the real world.

Box 2.3. Extrapolating beyond RCT with RWD. Example 1: Sorafenib for advanced hepatocellular carcinoma

The UK Cancer Drugs Fund reconsidered (TA474) a previously published NICE technology appraisal guidance of sorafenib for treating advanced hepatocellular carcinoma (TA189)[31] folowing the availability of new data from the manufacturer. The Appraisal Committee reviewed data from three longitudinal observational studies, Palmer et al. 2013, the GIDEON study, and King et al. 2016 and decided that the GIDEON study with a sample matched to the participants of the original RCT (SHARP) was most appropriate to validate the manufacturer's choice of extrapolation method for the survival curve in the original appraisal.[31] Hence, an important prerequisite of using RWD to validate extrapolated curves from RCTs is that the samples should represent the target population.

Box 2.4. Extrapolating beyond RCT with RWD. Example 2: Azacitidine for myelodysplastic syndrome

In another example of using RWD to extrapolate beyond RCT, a prospective observational study was conducted in Ontario from 2010 to 2016 to compare different dosing schedules of azacitidine. Azacitidine is an anti-cancer drug for the Myelodysplastic syndrome (MDS).[32] In 2010, the drug was approved based on an RCT, showing overall survival benefit. The registered dosing regimen required azacitidine to be initially taken 7 days in a row. However, in Canada, clinics usually only open from Monday to Friday, not on Saturdays and Sundays. From a logistical perspective, the intended 7-day regimen was not implementable in Canada. Repeating the RCT with a modified dosing regimen was also not feasible. Hence, a prospective observational study with 3 dosing schedules was proposed: (a) give 5 consecutive doses during the weekdays, skip the weekend, and then give the remaining 2 doses over the next two weekdays, (b) get 6 consecutive doses, by opening the clinics on Saturday mornings to allow for the additional 6th dose and (c) get 7 consecutive doses. The Ontario government provided temporary funding for all 3 regimens from 2010-2016 in order to facilitate the collection of RWD for evaluation. After 6 years, it was shown that the survival curves based on the 3 regimens were similar, suggesting that there was no significant difference in survival. The provincial HTA committee, the Ontario Steering Committee for Cancer Drugs (OSCCD), discussed the RWE with the Ministry of Health, who subsequently converted the temporary funding to permanent funding for all three regimens.

2.2.2. Localizing economic models

RCTs or observational studies conducted in a foreign-country context are unlikely to inform policy making in the local Asian contexts, considering differences in current clinical practices, healthcare financing systems, ethics and judicial systems. Local RWD can help to close the gap and this may be one of the most important applications of RWD and RWE in Asia (Box 2.5).

Box 2.5. Examples of localization of economic models using RWD (Taiwan and Malaysia)

Taiwan: With a well-developed national database of registry and claims data, Taiwan was able to utilize RWD to localize their studies across diseases. One example is the evaluation of the long-term cost-effectiveness of different cervical cancer screening strategies in Taiwan.[6] Chow et al. used a natural history model for cervical cancer adopted from the literature, and estimated survival rates for cervical cancer over different time horizons from the Taiwan Cancer Registry. Age-specific mortality was obtained from the Department of Statistics for Taiwan's female population; local direct medical costs from the Bureau of National Health Insurance (NHI) and another local publication. With these findings, the authors recommended a screening strategy for combined human papillomavirus (HPV)-Pap smear every 5 years for the publicly financed healthcare system, over the other 8 strategies evaluated.

Malaysia: Epidemiology and resource utilization evidence generated from the Malaysian Dialysis & Transplant Registry was used in an HTA which compared single use vs reusable dialyzers in hemodialysis.[35]

2.2.3. Re-evaluation of initial reimbursement decisions

As new technologies or policies are introduced into the health system, the opportunity cost and marginal effectiveness of some existing technologies might change, calling for a re-evaluation of the existing technologies. Under such circumstances, the original RCTs might not be able to represent the updated real-world settings. Thus, if the relevant health data system is developed, countries can utilize the rich information from RWD to inform the re-evaluation (Box 2.6).

Box 2.6. Example of reassessment using RWE: Australia's National Cervical Cancer Screening Program

RWE were used to inform revisions to the National Cervical Cancer Screening Program (NCSP) in Australia.[36] NCSP was established in 1991, providing bi-yearly conventional Pap tests for 18- to 69-year-old women. Registers were established within each jurisdiction. NCSP significantly reduced cervical cancer incidence and mortality rates in 1990s. However, in recent years, evidence from RCTs has shown that HPV DNA testing is more effective than traditional cytology-based screening.[37] [38] The former might also save costs by allowing patients' self-collection of testing samples. Meanwhile, a nationwide free HPV vaccination program, introduced in 2007, has high coverage across Australia and has significantly reduced cervical abnormalities for vaccinated women, especially for youngest women.[39] The development in new testing technologies, together with the success of HPV vaccination program, inspired a revision of NCSP, proposed as: a 5-yearly testing with a HPV test (with partial genotyping) and reflex liquid-based cytology, for 25- to 74-year-old women. Based on registry and immunization data, the Medical Services Advisory Committee (MSAC) in Australia evaluated the original NCSP and the revised NCSP, taking into account the effect of vaccination. While the default position of the MSAC is that RCT data remains the gold standard for evidence generation, MSAC agreed that the registry data were useful to demonstrate that with vaccination offered, the revised NCSP, compared to the original NCSP, saved both costs and life-years. As a result, the revised NCSP was implemented from December 2017.

2.2.4. Leveraging RWD/RWE for price negotiations or managed entry schemes

RWD/RWE can be leveraged by pharmaceutical manufacturers to support flexible subsidy arrangements with payers, and can help to balance the need for early access to innovative drugs with the need for evidence-based decision making (see Box 2.7 for an example). One payment model involves paying the manufacturer in annual installments over several years, with the annual payment contingent on the real-world performance of the product. For example, if the treatment efficacy is expected to last for 10 years, then the reimbursement is divided into 10 annual installments, with each installment being paid out contingent on the patient still being alive and responding to treatment. This payment model needs to be supported with the development of a registry to collect patient outcomes so that the manufacturer can be duly reimbursed. Economic models that informed the original cost effectiveness analysis using trial data, can then be updated with RWD collected for a more accurate assessment of the ICER in the local context. In other situations, price negotiations between payers and pharmaceutical manufacturers based on cost-effectiveness analyses may fail and require RWE for outcomes-based agreements. This is because, for treatments with a small market size or which address a high unmet need (such as in the case of treatments for rare diseases), manufacturers may try to justify setting a higher price irrespective of the ICER. Another example of leveraging RWE, although less common, is in price re-negotiations or managed exit (disinvestment) of drugs, especially in settings where the evidence to support initial market entry is very weak.

Box 2.7. Example of a managed entry scheme: Australia

The Australian government introduced the managed entry scheme (MES) in 2010 to accelerate patient access to innovative drugs.[40] Conditions for an MES in Australia include high and unmet need for the drug and evidence that can be gathered within a suitably short time frame to resolve any initial uncertainties in the evidence base. One product that went through this process in 2013 was crizotinib for the treatment of ALK positive non-small cell lung cancer (NSCLC). The Pharmaceutical Benefits Advisory Committee (PBAC) initially deferred the reimbursement decision due to uncertainty with the incremental 12-month OS proposed. A resubmission was subsequently made by the manufacturer in March 2014 with a MES proposal. To address the uncertainty surrounding the survival benefit of crizotinib, the manufacturer agreed to collect 12-month survival data for the first 50 patients receiving crizotinib after it was listed on the Pharmaceutical Benefits Scheme (PBS). A price reduction was agreed if the claimed survival benefit was not realized, and the manufacturer agreed to rebate the government a prespecified (confidential) percentage of the cost of treatment depending on the OS outcomes. In 2017, the manufacturer successfully provided survival outcomes collected from patients receiving treatment that were consistent with their original survival claims. The PBAC subsequently allowed crizotinib to continue to be listed on the PBS at the initial MES entry price and further data collection was no longer required to support the listing.[40]

2.3. Using RWE with caveats

While many see value in RWE and are exploring ways to utilize routine health data sources, there are inherent limitations associated with RWE application. We conclude the chapter with certain situations in which caution is needed when using RWD/RWE.

2.3.1. Biases arising from RWD/RWE

Biases may be introduced by confounding and/or selection bias in the RWD that may not have been adequately dealt with. Confounding represents a mixing of effects between the treatment group and external factors that may also influence the outcome, potentially obscuring or distorting the relationship that can be inferred.[41] These factors that influence the association between a treatment and the effect may either be known or unknown.[42] The most common concern in observational studies and real-world sources, like patient registries, is confounding by indication. Selection bias occurs when the observed subgroup of patients is not representative of the broader population of interest,[43] when using patient-level data from real-world sources and is a threat to both the internal and external validity of the study and its generalizability to a larger population. It is important to note the difference between confounding and selection bias and that methods to control for the former may not address the latter.

Box 2.8. Biases arising from RWD: Lesinurad for treatment of chronic hyperuricemia

An example where bias was introduced into an HTA through the use of RWD was in NICE's technology appraisal of lesinurad for treating chronic hyperuricemia in people with gout (TA506)[44]. In RCTs, lesinurad was found to improve serum uric acid level, without any evidence in reducing flares, increasing tophi healing, or delaying death. However, the manufacturer presented a meta-analysis of 6 observational studies showing that people who took uric acid lowering therapies lived longer than those who did not, which conflicted with findings in the RCTs. The RWE was not accepted by the NICE, which noted that (a) no evidence from RCTs validated the relationship between lowering serum uric acid levels and life expectancy, even with drugs other than lesinurad, (b) the observational studies from the UK did not suggest that uric acid-lowering treatment extended life and (c) known and unknown confounders, e.g. renal function and socioeconomic status, were not well controlled for in the observational studies.

A discussion of common biases in RWD/RWE and approaches to mitigate them is covered in Theme 3 ('From RWD to RWE'). It details the different biases that may affect both external and internal validity of a study such as selection bias, confounding, misclassification, as well as missing data. An overview of the most common statistical approaches to address the limitations of RWD are reported, together with examples from mostly Asian studies that have used the mentioned approaches.

2.3.2. RWD data quality

Fit-for-use RWD for HTA is a challenge because the data is not originally intended for research. Noise in routinely collected data can be caused by coding inaccuracies and inconsistent naming conventions over time and across sites.[45] Study sites may lack data management protocols, are subject to human errors in data entry, and omit important variables needed for HTA. To overcome this issue, a study in Malaysia for example, required extensive primary data collection to supplement data from the Asian-Heart Failure (HF) Registry Data in order to estimate the cost of heart failure in Malaysia.[46] Validation of the collected RWD can be addressed with quality management/assurance plans, e.g. one that periodically checks a subset of the extracted data for accuracy, consistency, completeness and plausibility.

The following chapter, Theme 2 ('Collecting RWD') will discuss more of these data quality and validation issues by each type of RWD source, methods of collection, and introduce suggestions for best practices in data collection in Asia.

2.4. Conclusion

RWE is already utilized in many countries as supplementary evidence to inform reimbursement decisions. Because of limitations and/or the lack of RCT-generated efficacy data, HTA agencies have been exploring the benefits and limitations of using RWD to supplement and enrich primary evidence to demonstrate the cost-effectiveness of drugs in each local context. Examples when RWE may be useful to inform decision making includes disease areas where RCTs are limited and/or of poor quality, or impossible to conduct for ethical reasons or due to small numbers (e.g. rare diseases). RWE can also be used to contextualize and localize economic models, extrapolate RCT data beyond trials, and for price setting and negotiations with manufacturers based on real-world outcomes.

While RWD and RWE are useful in the stipulated scenarios, it is important to remain cautious in their application as they can be subject to various forms of bias and generate misleading conclusions. The recommendations in this chapter relate to use of RWD and RWE with the following caveats:

Box 2.9. Recommended caveats while using RWD and RWE

1. RWD and RWE are generalizable to routine clinical practice only if the data represents the target clinical population reflective of the local context;
2. Trial and real-world populations are, by definition, different and any comparisons made, even 'validations', should be cognizant of the efficacy-effectiveness gap;
3. Observational studies can link surrogate and final end points, but sensitivity analyses should be used to validate the causal link and avoid spurious conclusions; and lastly,
4. Limitations and threats to validity from confounding, missing data, and overall low-quality data should be noted.

Evidence should only be accepted to inform decision making if it is considered robust and generalizable to the local context.

3 Theme two: Collecting RWD

3.1. Introduction

Many concerns raised about the value of RWD to inform reimbursement decisions relate to the perceived quality and validity of the RWD collected, which is heightened by the lack of, and difficulties establishing, universally accepted methodological standards or principles for the design, conduct, and/or reporting of RWD/RWE.[47] Despite growing interest from stakeholders involved in HTA, these concerns reduce the incentive to generate and use it.

However, RWD users can exercise caution over the potential quality concerns, and introduce validation processes for data collected from different sources and through different study designs. To do that, this theme begins with the question of characterizing 'What RWD to collect?' for reimbursement decisions in the Asian setting; followed by the common sources for RWD ('Where to collect?'), their pros and cons, and the good practices associated with using them. Study designs, for example observational studies, pragmatic trials, and single arm trials are introduced in 'How to collect?', with a summary of methodological standards for each study type that are available in the literature.

The chapter concludes with a case study illustrating the importance of contextualizing the ethical and legislative issues associated with collecting RWD to each local setting ('Who to collect?') and a set of general recommendations.

3.2. What RWD to collect?

3.2.1. What RWD are needed to inform drug reimbursement decisions in Asia?

The type of RWD needed will vary depending on local HTA processes, and the perceived value that RWD may add to the evaluation to address any areas of uncertainty. Characterization of the different types of RWD that can be collected is usually guided by the PICO (Population, Intervention, Comparators and Outcomes) framework,[48] and how the RWD is intended to be used to inform reimbursement decisions (e.g. data collection to give decision-makers more confidence in the value of the technology before making a reimbursement decision; or afterwards post-reimbursement to inform reassessments). Additional RWD that do not fit within the PICO framework (e.g. epidemiological data, prescribing trends and treatment adherence rates etc.) may also need to be collected to support decision-making.

Population characteristics

Although population characteristics do not often directly inform the treatment effect or economic modelling parameters required for HTAs, RWD such as person-level demographic and socio-economic information as well as medical history are important to collect in order to establish balanced groups for relative effectiveness comparisons (see Box 3.1). Approaches to generate balanced groups are described in Theme 3 ('From RWD to RWE'). Variables that may be confounders (e.g. prognostic factor and/or effect-modifiers) need to be considered. Patient characteristics are also important for understanding how the drug or medical device works in an "extended" patient population compared to the clinical trial. Patient-level variables include:

- Demographics (e.g. age, sex, ethnicity)
- Socio-economic factors (e.g. geographical location (urban vs rural), income, education, insurance)

- Medical history or pre-existing conditions (e.g. Charlson comorbidity index, family history, genotype, biomarkers, prognostic factors, laboratory data, disease staging, current treatment, line of therapy).

Among these, treatment and line of therapy are especially relevant in the Asian context as medical practice can be heterogeneous in countries with large populations or different geographic regions, and often does not reflect practices used in clinical trials. This is particularly true for decentralized healthcare systems such as China, India, Indonesia and the Philippines. RWD that capture how individual patients are treated in different Asian countries will help inform whether the effectiveness outcomes of evaluated drugs are likely to be comparable across Asian settings, and will be especially useful for some countries that do not have sufficient resources to collect local effectiveness data.

> **Recommendation**: In addition to patient level demographic and socioeconomic RWD, variables that describe medical history/condition and practice variation across Asian countries are important, especially for decentralized health systems, to inform if patient groups and findings are comparable across different settings.

Box 3.1. Example of using patient-level RWD for HTA in Asia: Taiwan

Cheng et al. evaluated the real-world cost effectiveness of using drug eluting stents (DES) compared to bare-metal stents (BMS) for coronary heart disease in Taiwan.[23] BMS are included in the National Health Insurance coverage but not DES; if patients choose DES, the price differential beyond BMS has to be borne by them. A retrospective claims data analysis was conducted to inform the evaluation and used linked data from the National Health Insurance (NHI) Longitudinal Health Insurance Database to identify patients with stable coronary heart disease who underwent a coronary stent from 2007-2008 and follow them for five years to capture patient-relevant outcomes. The BMS cohort was 2:1 propensity score matched by *gender*, *age*, *stent number*, and the *Charlson comorbidity index (CCI)* to reduce confounding, resulting in 852 patients being included in the study: 568 in the BMS group, and 284 in the DES group. The study demonstrated that DES was a more cost-effective strategy than BMS and made a strong recommendation for the National Health Insurance to consider fully reimbursing DES instead of the current policy.

Intervention and control

Beyond population characteristics, there are other differences between RCT data and RWD. For example, dosing in clinical trials is usually fixed but can in fact be highly variable in the real-world setting.[49 50] The tail end of the duration of treatment in the real-world also cannot be informed by clinical trials due to their limited length of follow up.[14] Treatment continuation is considered as a meaningful outcome of therapeutic efficiency over time, but treatment in RCTs is more intensively monitored than in routine practice leading to artificially high protocol-driven continuation rates. Hence, understanding various real-world aspects of the intervention in terms of the actual duration that patients stay on treatment, dosing, waning of effect, discontinuation rates, and the reasons for

discontinuation will impact the effectiveness, cost and financial implications of the intervention and give decision-makers more certainty of its overall cost-effectiveness in the local context.

> **Recommendation**: We should harness real-world data to better understand the optimal dosing, duration of treatment, waning of effect, and rate of discontinuation in different patient populations. Patient heterogeneity in the real world means that trial data may not be readily generalizable to all local contexts.

We recommend that Asian countries with late or delayed reimbursement decisions after market approval, which therefore have more time to collect RWD, plan to collect RWD on the intervention of interest beyond the follow up period of its pivotal RCT conducted for regulatory approval to capture longer-term outcomes. The choice of comparator or control should be relevant to the policy question being addressed, to ensure that the incremental impact of the intervention is evaluated compared to the existing standard of care in the local context and is directly relevant to decision-making. Unlike RCTs, the comparator is not necessarily fixed, as there could be differences in clinical practices across different settings. Additionally, there may be more than one appropriate comparator depending on the variability of routine clinical practice.

> **Recommendation:** Timeframe of RWD collection for the intervention should be long enough to allow longer-term safety and efficacy outcomes to be captured.

Outcomes – Effectiveness

RWD enables estimates of effectiveness rather than efficacy to be collected in a variety of real-world clinical practice settings. The diverse study population in a real-world setting may more accurately reflect the range and distribution of patients likely to receive treatment. Clinical measures of effectiveness typically include biological measures of morbidity and mortality and may be surrogate and/or long-term measures.

Estimates of relative effectiveness (difference in effect between intervention and control) are used in the denominator of a cost-effectiveness ratio in economic evaluations. Guidelines for health economic models to inform cost-effectiveness considerations frequently specify a preference for a lifetime time horizon or, one that is sufficiently long to reflect all important differences in costs and outcomes during the course of a disease.[51] Given the reality that it is often infeasible to conduct RCTs with an indefinite time horizon, RWD is useful to supplement the RCT data and provide estimates of clinical effectiveness, particularly among a heterogeneous, unselected population.[52]

Often, interventions are deemed to have a beneficial impact based on RCTs that have been designed to detect incremental differences in *surrogate* outcomes (e.g. progression-free survival). Although these endpoints may be sufficient for regulatory purposes, they are often highly uncertain, especially when used to extrapolate treatment benefits over a long time horizon and inform cost-effectiveness assessments.[53] It is highly recommended that the impact of interventions on *final longer-term* outcomes such as mortality are re-evaluated based on person-level data using appropriate methods (including microsimulation or the net benefit framework) to compare actual versus estimated

outcomes and determine the true value of the intervention. Some of these methods are described in Theme 3 ('From RWD to RWE').

> **Recommendation:** Since drugs are typically evaluated using model-based estimations of cost-effectiveness and surrogate outcomes from RCTs, the impact of interventions on final outcomes should be revisited using person-level RWD once it is available to assess if the outcomes in the trial actually translate to clinically meaningful improvements for patients.

Patient reported outcomes (PROs), directly obtained from patients rather than being clinically measured, are another example of useful person-level data that can be important measures of effectiveness (see next section).

Outcomes – Patient-reported outcomes (PROs)

PROs encompass patients' own assessment of their health condition and treatment, including symptoms, functional status, and health related quality of life. They can be general or disease specific measures. For HTA purposes, PROs to elicit patient preferences and generate health utilities, allows for estimation of quality-adjusted life years (QALYs) that are used in the denominator of an incremental cost-utility ratio.[54] Our focus on EQ-5D is due to its use in most of the pivotal trials conducted by manufacturers, and it may be reasonable to make use of the same instrument in comparisons between trial-based health utilities and real-world based health utilities. Population-specific value sets for standardized PRO instruments such as the EQ-5D have been developed in China, Hong Kong, Indonesia, Japan, Malaysia, Philippines, Singapore, South Korea, Taiwan, and Thailand,[55-59] and a study is currently underway to determine the Indian value set using the EQ-5D-5L. HTA agencies prefer PRO instruments that have been validated in their countries. However, researchers often reference published utilities from other countries when conducting HTAs and then highlight lack of local estimates as a limitation to their study, because collecting local utility data is typically resource intensive in terms of time and cost and is not feasible to conduct. Where possible, other patient-reported outcome measures (PROMs) and patient-reported experience measures (PREMs) can also be collected to support reimbursement.

The routine collection of PROs is challenging and numerous barriers to real-world collection of PROs include: (a) clinician perception that PRO completion takes up valuable time during the patient visit and consultation times are not sufficiently long to enable PROs to be adequately captured; (b) limited ability to deliver PROs in user-friendly electronic formats for patients; and (c) clinical workflow demands that hamper the integration of data collection into clinical practice.[60] In countries where manufacturer-led submissions are accepted, it may be easier to collect PROs at the point of initial reimbursement by encouraging manufacturers to submit local PRO evidence as a condition of reimbursement. PROs collected by the manufacturers could then be reviewed periodically as part of reassessment processes for their technologies. However, most Asian HTA agencies currently do not accept manufacturer-led submissions and in-house technical staff are responsible for evidence generation.[61] Local collection of PROs is strongly encouraged but requires fostering collaborations between HTA agencies and academic units or clinicians to overcome potential logistical barriers associated with data collection. Tools like the

International Society for Quality of Life Research's (ISOQOL) *User's Guide for Implementing Patient-Reported Outcomes Assessment in Clinical Practice* can be a useful resource.[62]

Recommendation: Local collection of PROs is strongly encouraged but requires fostering collaborations between HTA agencies and academic units or clinicians to overcome potential logistical barriers associated with data collection.

Outcomes – Safety

RCT safety data usually lacks generalizability to real-world settings due to selective recruitment of patients and consideration of few predefined adverse drug events (ADEs) over a limited follow-up period.[65] For example, patients with comorbidities are often excluded from clinical trials, but they may experience greater toxicity than what is observed in clinical trial participants. RWD provides information on the evolving long-term risk-benefit profile of a new drug beyond the time frame of an RCT, in a much larger patient sample. Real-world studies are important to confirm that the ADE profiles of the drugs when used in routine clinical practice are similar to those observed in clinical trials (Box 3.2).[63] Occasionally, new adverse events that were not reported in the pivotal trials may be detected using real-world data.[66]

Box 3.2. Using RWD in post-marketing surveillance: Crizotinib in Japan

In Japan, crizotinib was approved and made available for the treatment of non-small cell lung cancer (NSCLC) positive for ALK and ROS1 fusion genes, despite limited representation of Japanese patients in the clinical trial program (only 15 Japanese patients participated in the Phase I clinical trial of crizotinib). In view of the limited information available on its efficacy and safety in real-world settings, Ueno et al. conducted a post-marketing surveillance study with 2,028 patients to determine the safety of crizotinib in a Japanese clinical setting.[63] The observation period was 52 weeks (vs median study duration of 31 weeks for crizotinib's PROFILE 1007[64] phase 3 trial). The authors concluded that no new safety concerns were observed.

Additionally, RWD is a readily available resource for improving pharmacovigilance. Safety monitoring and reporting of ADEs are already part of the post-marketing surveillance required by Asian regulatory bodies. The majority of Asian countries are members of the World Health Organization (WHO) Program for International Drug Monitoring and adopt elements of the ADE case safety report forms in the WHO INTDIS (International Drug Information System) format, or the ICH-E2B standard native to the WHO central database, VigiBase.[65,67] However, surveillance of drugs following marketing authorization in Asia is a largely passive process that relies on spontaneous reporting systems, which means that the RWD generated is limited to reporter submissions (patient or clinician). These systems can be hampered by incomplete information, such as on exposure and outcomes, and underreporting and reporting biases exist.[68] An improvement that supplements such reporting is the routine use of RWD for active surveillance, whereby safety concerns can be more proactively identified by interrogating existing real-world datasets such as administrative claims and electronic medical records. This approach is more resource intensive but allows the study of causal relationships between an intervention and its effects, by ascertaining the number of adverse events (numerator) in patient populations exposed and unexposed to an intervention (denominators), followed by using observational methods for signal detection.[69]

> **Recommendation:** Routine use of RWD such as administrative claims and electronic medical records should be adopted to identify important safety signals more proactively.

Outcomes – Cost

A review by Sculpher et al. identified that the unit costs associated with health care resources are the most common contributor to variability in economic results between different geographical locations.[70] This is perhaps why cost is one of the most common model parameters to be localized using RWD. There are many examples in Asia where RWD studies' effectiveness estimates were based on overseas studies while cost data were localized. One such example is a cost effectiveness analysis (CEA) of pembrolizumab compared to standard of care for patients with advanced melanoma in Hong Kong.[71] Overall survival, quality of life and adverse event data were derived from the final analysis of the multi-center KEY-NOTE-006 trial[72] involving 16 countries outside Asia, while costing data such as costs for drug acquisition and treatment administration were obtained from the Hong Kong Hospital Authority.

Cost outcomes are defined to include estimates of direct medical and non-medical resource utilization and their associated costs in the real-world, which are used in the numerator of an incremental cost-effectiveness ratio (ICER) in economic evaluations. RWD sources like claims databases (3.3.2) and electronic medical records (EMRs) (3.3.3) are useful sources for local cost information. Key data elements are itemized unit cost (distinguished from 'charge') and resource use ('number of units') used in costing.

We note several issues with RWD collection involving direct medical costs that are relevant to Asia. If EMRs or hospital registries are the main source for this information, cost data relating to resource use outside of the hospital setting (e.g. transfer of discharged patients to rehabilitation centers or hospices) can be extremely fragmented.[73] Furthermore, concise economic evaluations should only concern costs related to the intervention and not costs induced by unrelated comorbid diseases.[73] The heterogeneity of the population that has informed the RWD does not preclude an older population for which these comorbidities are present, therefore, interpretation of the true costs is less straightforward. It is not always easy to establish such a relation in databases, and care should be taken when determining which cost items should be included in the study.

> **Recommendation:** Determining which costs in RWD are related to the intervention is not straightforward when the patient population may have comorbidities. In Asian settings, the relevance of a cost item should be considered in relation to the burden of data collection and the perspective of the study (e.g. societal, public healthcare payer). Not all costs have to be included but sufficient justification should be given for including or excluding specific study cost items.

The decision of which cost items to include should depend on the population, treatment, disease, and perspective of the study. RWD costs elicited should also take into account the financing mix of public and private sources in any country context, which may have important implications in terms of deciding

which costs to include depending on the perspective of the HTA study or research that it is informing. These factors are becoming increasingly important in Asian countries given the greater reliance on private health expenditures and patient co-payments to provide access to treatments where government subsidy is not available or insufficient to cover the majority of the treatment cost. Examples where the financing mix is particularly important include India where there are increasing out-of-pocket payments, and Singapore's multi-payer system (a mix of government subsidies, insurance and patient co-payments).

In Asian settings where data can be fragmented and incomplete, balance should be established between the relevance of a cost item in relation to the burden of collection. It can be guided by main cost drivers identified from previous research from RCTs and studies on burden of disease, or by using the 'value of information' approach if there is available expertise.[74] It may be not feasible, or efficient to collect data on services provided outside of the hospital if they do not contribute to a large proportion of direct medical cost.

Epidemiological data

Epidemiological data such as disease incidence or prevalence are important for conducting budget impact analyses and health economic analyses. Incidence, prevalence, and transitional probabilities extracted from RWD can inform the model parameters for pharmacoeconomic models. In a study in Malaysia, the authors used surveillance data from countries which have implemented the pneumococcal conjugate vaccine (PCV) to estimate the real-world impact of vaccination.[75] With RWD, epidemiological estimates of the indirect effects of vaccination (herd immunity) and serotype replacement could be elicited directly to include in cost-effectiveness models rather than relying on assumptions, which are often highly uncertain. The study provided strong evidence for the Ministry's decision to introduce PCV into its national immunization program, as it was found to be cost-saving compared to no vaccination and was expected to have a beneficial public health and economic impact on the Malaysian population.

Adherence

Adherence to treatment and persistence are considerations in health economic analysis that can be collected from RWD sources such as registries, EMRs, or surveys. An understanding of adherence levels matters as adherence rates affect comparative effectiveness estimates. Costs and effects incurred from a patient who did not adhere may not reflect the true cost-effectiveness of the therapy being assessed for reimbursement.

Methods of adherence measurement vary. Commonly used are direct measures of drug levels or its metabolite in body fluids (e.g. laboratory test data in EMRs), prescription refill rates (claims data, EMRs), and clinician assessments / patient self-reports (surveys). While direct measures of clinical evidence are most accurate, it is intrusive and expensive, requiring staff and techniques to perform, and biases can occur when patients know test schedules. Measures involving secondary claims database analyses of refill rates like the medication possession ratio (MPR) are used in some Asian countries (South Korea, Taiwan).

MPR is the sum of the days' supply for all fills of a given drug for a specified time period divided by the number of days in the period,[76] estimated from patient refill records in the EMR. It can be useful for research with large populations but assumes that medication taking behavior corresponds to refills, which is not always true. Survey methods on medication taking behavior are influenced by patient

desirability biases and has relatively poor sensitivity and specificity.[77] Given that there is no ideal measure of medication adherence, a multi-measure approach for (a) measurements of adherence directly and (b) using MPR as a proxy of adherence is recommended.[77]

> **Recommendation:** Adherence is important RWD that should be collected. A multi-measure approach is likely to provide a better estimate of adherence as existing methods of assessing adherence have their inherent limitations.

3.3. Where to collect? Sources of RWD and good practice guidelines

RWD can be collected from various sources including product or disease registries, routine administrative data sets such as claims databases, electronic medical records (EMRs), health surveys, or from daily wearables and personal tracking devices. Researchers and users of RWD should deliberate the benefits and limitations of various sources in the local context. It is critical to understand the potential reimbursement questions that may or may not be answerable because of the availability, access, and quality of the RWD sources. The strengths and limitations of these common RWD sources, as well as recommendations in the literature will be discussed below. In many cases the best solution to a policy question requires integrating sources and leveraging the strengths of each. The working group does not prefer one RWD source over another.

3.3.1. Disease registries

Disease-specific, and other public health-relevant registries such as for births, deaths, immunization records etc., consist of structured datasets that can be made available for analysis. They can be used for understanding natural history, assessing or monitoring real-world safety and effectiveness, assessing quality of care and provider performance, and as inputs for cost-effectiveness analyses. Disease registries involve prospective data collection that reflect everyday clinical decision making. Rare disease registries in particular play a key role in providing RWD to inform decisions regarding clinical effectiveness because evidence cannot be easily obtained through clinical trials due to limited patient numbers typically recruited. A non-exhaustive list of registries in Asia can be found in Appendix 8.1.

Data quality issues

Data quality is a major issue with disease registries when study sites are not experienced in data collection intended for research.[8] One example was the use of the ASIAN-Heart Failure (HF) registry to estimate the cost-of-illness of heart failure in Malaysia. The authors raised challenges in obtaining sufficient information on resource utilization because medication profiles were overwritten each time the registry electronic data was updated, deleting the information on medications previously prescribed.[46] In other cases poor data management, compounded by constraints in manpower and funding, can lead to valuable data loss.

Registries do not always include all patients in the target population. For example, some of the existing cancer patient registries in Japan only include patient information that is registered voluntarily, making it unclear whether the data are entirely representative of the populations under evaluation. Hence, it is sometimes difficult to justify the use of such data for reimbursement decisions at national level.

Table 3.1. Pros and cons of registries

Registries: Pros	Registries: Cons
• Patients are often followed over a longer time frame, allowing for an assessment of longer-term outcomes • Most registries have very few required visits, evaluations, or procedures. Treatment patterns reflect the everyday clinical decision-making that is most relevant to providers and payers • Rare disease registries record and increase understanding of specific diseases among a very limited patient population	• Registries sometimes include study sites that are not as experienced in data collection intended for research, affecting data quality • Selection bias can occur, especially for patients who did not or could not provide consent to enter registry,[78] i.e. data does not include the entire patient population • May lack control group within the same registry, increasing risk of bias because of systematic differences in the sources for selecting cases and controls

How are registry data used in drug reimbursement decisions in Asia?

A successful example of RWD from registries used in initial reimbursement decisions is the use of Malaysia's National Obstetric Registry to evaluate the cost-effectiveness of carbetocin compared to oxytocin as prophylaxis against post-partum hemorrhage during cesarean deliveries.[79] Data from the registry contributed by 14 major hospitals confirmed that compared to oxytocin, administration of carbetocin was simpler and its longer-acting nature reduced the need for additional medications. Taking this evidence into consideration, carbetocin was found to be cost effective.

Taiwan conducts post-market reassessment of reimbursement, and registries are frequently used to collect outcomes data especially for high cost oncology drugs with high uncertainty of clinical effectiveness, e.g. in managed entry agreements for direct-acting antiviral (DAA) medications for hepatitis C and immunological drugs for various cancers. Payments and claims data from the National Health Insurance Administration (NHIA) are directly contingent on patient's clinical response. Among Asian countries, Taiwan is exceptional in its accessibility policy for disease registries and other national databases. Access to the Health and Welfare Data Center (HWDC) that links many sources of national health data can be granted to anyone conditional upon prior approval of the research or industry-sponsored project by an Institutional Review Board (IRB). Safeguards to data privacy include deidentified datasets only accessible on-site, having statistical analysis syntax reviewed before access, and analyzed results examined before data export.

Despite these examples, access to disease registries is a commonly cited challenge in Malaysia, South Korea, Japan, and Singapore. The National Institute of Health (NIH) Centers for Disease Control and Prevention (KCDC) in Korea accept registration for clinical research in their CRIS system (cris.nih.go.kr). Publicly funded research, including many disease registries funded by KCDC are available, but most of them are considered investigator-initiated rather than public as researchers view the registries as their own. Public access to disease registries is therefore limited and potentially requires the individual to know the researchers or have contacts in the NIH/KCDC. A similar situation occurs in Japan whose registries have been implemented by medical societies and parties and are hence researcher-owned. Japanese registries are not allowed for HTA use. National registries in Singapore produce standard public reports of aggregated outcomes periodically, but they often have limited use as evidence to inform reimbursement decisions, given the data is not disaggregated and may not be relevant to inform estimates for particular subgroups or disease types. Several disease registries are also established in

different public healthcare institutions, but they are often investigator-led and data are not readily shared between hospitals or with the public.

Health-relevant registries like for births and deaths are sometimes separately curated outside the Ministries of Health and are typically requested for separately (e.g. from the Internal Ministry in Taiwan and the Immigration Checkpoint Authority in Singapore). Data linkage between such data and patient registries can be an enormous challenge.

Box 3.3. Good practices when collecting data for registries

How a system collects, cleans, monitors, and reports registry data determines whether the data can be useful toward the registry's goals. Critical factors relating to the quality of the data collection include how data variables are defined, whether personnel are adequately trained to enter the data, and verification checks targeting errors during collection resulting in out-of-range and logically inconsistent values.[73]

Registries should be carefully planned with clear objectives and extensive clinical input. These data collection, management, and quality assurance procedures should be defined in a detailed manual when establishing the registry and not after.[80] Registries may also be required to comply with local guidelines or legislation (e.g. Singapore's National Registry of Diseases Act). Quality assurance ensures that data are collected in accordance with the pre-defined procedures and that they meet the requisite standards of quality to meet the registry's intended purpose.

Importantly, as certain requirements may have significant cost implications, a risk-based approach to developing a quality assurance plan is recommended.[80] It should be based on identifying the most likely sources of error or potential lapses that affect the quality of the registry in its intended use of the data.

For more recommendations on data collection and quality assurance for registries, see:

- Gliklich RE, Dreyer NA, Leavy MB: Registries for Evaluating Patient Outcomes: A User's Guide[80]
- Blommestein HM, Franken MG, Uyl-de Groot CA: A Practical Guide for Using Registry Data to Inform Decisions about the Cost Effectiveness of New Cancer Drugs[73]
- China Real world Evidence (ChinaREAL) Consortium: Technical Guidance for Developing Patient Registry Databases (in Mandarin Chinese)[81]
- Singapore: National Registry of Diseases Act[82]
- Japan: Guidelines for setting up registries in Japan are now under development and planned to be published in 1 to 2 years. These guidelines relate more to privacy and data sharing than how to scientifically use the data. The government is planning to make RWD from registries available for research purposes in 5 to 10 years' time.

Box 3.4. Good practices when collecting data for rare disease registries

Rare disease registries can be a valuable tool for increasing understanding of the disease and supporting the development of orphan drug therapies but have their unique challenges due to limited patient population. Expert opinion from stakeholders such as patient advocacy groups, payers, patients, and their caregivers/families should be collected.[80] This requires the registry administrators to also effectively educate stakeholders about how they can meaningfully contribute to the data capture and the type of experiences or information that they should share.

Unique features of rare disease registries include:

- Limited number of patients with conditions of interest (either due to low incidence or large number of undiagnosed patients due to clinicians' lack of familiarity with the condition in their local context);
- Limited information available on the disease to guide development of a research and data collection plan. Diagnostic criteria may be complex or evolving;
- Disease-specific patient reported outcome measures may not be available. Long-term and even lifelong follow-up may be needed;
- Need to adapt and change over time as knowledge increases or treatments become available; and
- Limited treatment options (95% of rare diseases do not have effective treatments) and many of the available treatments are unaffordable for patients.

Thus, a key focus is on engagement and retention of patients/providers over the duration of the registry, closely monitoring follow-up rates over time to identify potential issues that may shape the data structure and definitions. Developing clear policies are recommended for governance of the registry and data access if multiple stakeholders are involved.

3.3.2. Claims databases

Compared to disease registries that are disease-centric, claims databases are focused on data that is generated from payment activities. Claims databases are also called billing and administrative databases. They consist of bills that health care providers submit to public (e.g. National Health Insurance Administration, Taiwan) or private insurance entities for reimbursement of covered services. Claims data are especially rich as an RWD source for countries/regions like Japan, South Korea, and Taiwan with national payers because of the breadth and comprehensiveness of all patient encounters across the full continuum of care. Variables and data types recorded can be diagnoses and procedure codes, dates of service and lengths of stay, pharmacy dispensing data, clinical data, and patient demographics. The purpose of claims data is for payment, making it convenient for researchers to establish the cost for certain diagnoses by consulting fee schedules and reimbursement data for CEAs.[83]

The data has the benefit of relatively structured data fields and lend themselves well to retrospective analyses of clinical and economic outcomes, which can be conducted in a relatively short period of time and at lower cost compared to prospective data collection.[8] In terms of scope, claims databases tend to capture a more holistic view of information from all providers caring for a single patient as long as they have made claims submissions, with the exception of healthcare systems that do not consolidate claims by patient, e.g. in Japan. An additional benefit given the large volume of historical data is being

able to identify outcomes of patients with rare events more easily in order to assess the economic impact of various interventions.

Data quality issues

Retrospective billing and claims data, however, often face issues of data quality (missing data, unintentional miscoding, and intentional miscoding or 'upcoding'), and may not collect all variables of interest (e.g. symptoms, health status). There is also often a lack of distinction between cost and charges,[8] which can be an issue if claims data does not represent the economic value of resources used to provide services, and is influenced by monopolies or monopsony, then its utility for costing as part of a CEA is limited. Claims-based reporting also has latency of data refresh at varying intervals depending on the provider and are hence not updated in real time.

Table 3.2. Pros and cons of claims databases

Claims databases: Pros	Claims databases: Cons

• Comprehensive billing record covering all medical claims of a population, useful in measuring and estimating resource use and costs for economic evaluation
• More structured, and standardized format than EMRs
• Holistic view of all interactions of patient with health care services (Vs EMRs, which are provider specific)
• Analyses can be performed at low overall cost and in a short period of time
• Claims databases lend themselves well to retrospective longitudinal and cross-sectional analyses of clinical and economic outcomes at patient, group, or population levels
• Researchers can identify outcomes of patients with rare events given the large number of people captured in the database
	• Latency of data refresh for claims data (versus EMRs that are updated in real-time)
	• Depending on the comprehensiveness of coverage, the claims database may have limitations in terms of utility
	• May not allow granular assessment of cost components, which may be relevant for the CEA
	• Validity of retrospective claims database analyses in terms of:
	  ○ limited clinical information on health outcomes, health status, and symptoms (only captures data relevant to billing such as diagnosis and procedures)
	  ○ data quality (missing data, coding errors)
	• Availability of claims data for analysis is subject administrative approvals, and generally may not be available in public domain
	• Privacy of patient data is a priority in many countries. Potential difficulties in access and obtaining consent if data privacy is not adequately enforced

How are claims data used in drug reimbursement decisions in Asia?

As part of an economic evaluation in Indonesia to determine whether pneumococcal conjugate vaccine (PCV)10 or PCV13 should be included in the national benefits package, all country-specific health care costs used in the CEA were obtained from local clinical and billing databases.[84] In Korea, a cost utility analysis (CUA) of the Prostate Specific Antigen (PSA) test which included local cost estimates from Health Insurance Review and Assessment (HIRA) claims data, and effectiveness data from linked data from HIRA claims, NCI central cancer registry, vital statistics, and a tertiary hospital laboratory, provided evidence against the inclusion of PSA in the national screening program compared to the current

opportunistic screening. . The study results showed the PSA test was cost-ineffective as part of the National Cancer Screening program in Korea.[85]

Box 3.5. Good practices when collecting data for claims databases

Claims databases are also integral to providing local resource use and costs as long as the research population and comparison groups are clearly defined, identified, and addressed by an appropriate study design.[8 9] Recommendations to minimize threats to validity include:

- *Consistency of study design and conclusions with the claims database*. Before conducting a study, the degree to which the required data elements are captured by the claims database should be investigated.[9] Designs that require very specific diagnostic codes or lab data cannot be fulfilled with claims databases. Conclusions should not go beyond the capabilities of the database.
- *Use of a study design that includes comparisons.* Comparison groups are constructed to be as similar as possible except for the treatment of interest. Also, conducting both pre-post (in same group of patients) and cohort (different groups) comparisons is a stronger design than either alone.[9]
- *Use of appropriate constructs*. This relates to the translation of concepts into variables that are captured by the claims database, dependent on diagnostic codes and other criteria.[86] These decisions should be documented carefully.

For further reading, see:

- Motheral BR, Fairman KA: The Use of Claims Databases for Outcomes Research: Rationale, Challenges, and Strategies[9]
- Birnbaum H, Cremieux P, Greenberg P, LeLorier J, Ostrander J, Venditti L: Using Healthcare Claims Data for Outcomes Research and Pharmacoeconomic Analyses[86]

3.3.3. Electronic medical records (EMRs)

Claims data are designed to hold only pieces of information relevant to facilitate payment. It requires diagnosis, services, and cost data but is otherwise limited in clinical information on the actual patient health status and outcomes. While claims data have broad scope and coverage, EMRs provide a much richer dataset that is generated in real time (no latency), allowing for rapid response. EMRs contain structured and unstructured data fields that include real-time patient demographic and health information from clinical encounters, including diagnoses, symptoms, treatments, patient habits and surveys, lab results, and prescriptions. In addition to these core data elements, EMRs include peripheral documents such as imaging data, pathology reports, and patient history documents. That said, while detailed longitudinal data in EMRs and at a patient-level is a rich RWD source, critical data may not be stored in a readable format and transforming the said information that is originally intended for clinical purposes to RWE can be challenging. In contrast to claims databases, EMRs also tend to be provider specific and are limited in scope to patients who access the provider's services.

Data quality issues

EMR data, collected in routine clinical care for non-research purposes, is often dirty and incomplete due to the collection being done over a long period of time, and the data must be cleaned before it can be used. Noise in the data is caused by coding inaccuracies and inconsistent naming conventions over

time and across sites, etc.[45] Biased sampling may occur when the same patient visits different providers for medical help who do not share information between each other.

Table 3.3. Pros and cons of EMRs

EMRs: Pros	EMRs: Cons
• Generates real time data about clinical treatment and outcomes. Evaluating real time EMR data can allow for a more rapid response (no latency) • Contains rich, longitudinal information if patient stays with the same provider, including disease-specific symptoms, patient vital signs, habits, etc. at the person-level	• EMR may be hospital/provider specific, and does not capture patients who access health services outside of the provider's health system • Logistical challenges even in accessing own data • Critical unstructured data may be stored in non-machine-readable formats (like handwritten notes) • Transforming the information meant for a clinical workflow to a format for research purposes requires further statistical analysis

How are EMR data used in drug reimbursement decisions in Asia?

RWD from the EMR has also been used in Bhutan to determine whether adding the rotavirus (RV) vaccine in the National Immunization Program would be cost effective, with relevant hospitalization and death statistics extracted from Bhutan's Health Management and Information System (HMIS).[87] In the near future, Bhutan also has plans to roll out an electronic patient information system (ePIS) to all levels of health facilities, from sub-posts, primary health care centers, to district, regional, and national referral hospitals. Policy makers and researchers are expected to be able to use integrated EMR data for future evidence-based decision making on health benefit package development.

Box 3.6. Good practices when collecting data for EMRs

There are practical challenges to the interoperability of EMR systems, which have to do with diverse systems used across providers, as well as diverse clinical data standards used by the health care and research communities. Some of these can be addressed if there are existing local requirements for open data standards or EMR data standardization. For data extraction, a protocol-defined data collection plan should define the 'when' and 'how' of measurement (timing and method of measurement).[88] Extra attention should be paid to the reliability and quality of unstructured EMR data and how they are translated into formats ready for analysis. Like registries, a quality management plan (e.g. standard operating procedures) should address validation of collected data, for example to periodically check a subset of the extracted data for accuracy, consistency, completeness and plausibility with the EMR source data.[88 89] The quality of EMR data must be ascertained in order to ensure their appropriateness for use to inform reimbursement decision making.

For further reading on data collection and quality assurance for EMRs, see:

- U.S. FDA Guidance for Industry: Use of Electronic Health Record Data in Clinical Investigations[88]
- ChinaREAL Consortium: Technical Guidance for Developing Research Databases Using Existing Health and Medical Data (in Mandarin Chinese)[90]

Box 3.6. (continued)

For further reading on specific verification checks for EMR data, see:

- Duke Margolis RWE Collaborative RWD Quality Working Group, for verification recommendations adapted from database level checks used in PCORnet and Sentinel: Determining Real-World Data's Fitness for Use and the Role of Reliability[89]

3.3.4. Health surveys

Population, household, and health surveys are designed to collect health related information from a target sample of the community, which can inform disease epidemiology, health status and well-being, health care utilization, health care expenditures, and treatment patterns. With rigorous study designs, surveys can provide information on a large group of the population, unlike data collection from limited participants in a RCT. Some examples from Asia include India's National Sample Survey Office (NSSO) that collects population-wide data on morbidity rates, health seeking behavior, and social consumption related to health, such as expenditure on healthcare.[91] This can be used to estimate costs in HTA. In Japan, health check-up data has been routinely collected for the population and includes information about basic health indices and non-communicable diseases, which can be used to derive epidemiological estimates for specific conditions. The National (Population) Health Surveys and National Health Surveillance Surveys have been collected on a regular basis in Singapore which, among other things, track the health and risk factors, as well as lifestyle practices of Singapore residents. Other similar surveys include Thailand's National Health Examination Survey[92] and Taiwan's National Health Interview Survey. Health surveys may also include patient reported outcome instruments that are administered to patients in order to gather data on their quality of life. However, population health surveys as stand-alone sources of data can be challenging to use without the ability to link them to other RWD databases. Other limitations of health survey data for initial coverage and reimbursement decisions include the lack of a representative population and lack of relevant data on specific products.

Data quality issues

Surveys are subject to issues of subjectivity and recall bias from respondents.[93]

Table 3.4. Pros and cons of health surveys

Health surveys: Pros	Health surveys: Cons
• Health surveys typically collect information on representative individuals in the target population • Can be methodologically rigorous • With well-designed sample surveys, can provide information about all members of the target population, not just those who are participating in a given RCT • Can make unique contributions about generalizability of treatments and their impacts and about use of and expenditures for health services	• Lack of relevant data on specific treatments/products to guide reimbursement decisions • Relevance of health surveys is dependent on periodicity of survey. If done infrequently, the data may not be relevant • Surveys are subject to issues of subjectivity and recall bias • Can be more resource intensive to distribute survey, follow up on responses and collate data

How are survey data used in drug reimbursement decisions in Asia?

A national level population-based survey in India has been used to answer a policy question of interest, whether to reimburse a new pan-genotypic direct antiviral, sofosbuvir/velpatasvir for the treatment of hepatitis C virus (HCV) instead of standard treatment.[94] This is motivated by the effectiveness of sofosbuvir/velpatasvir regardless of genotype, unlike in standard of care where drug treatment was previously dependent on HCV genotype. Removing the need for genotyping prior to treatment initiation had potential to improve patient outcomes as HCV patients in some districts could not access genotyping facilities and did not continue treatment. The demographic data in this study was obtained from census data, and out-of-pocket costs estimated using the National Sample Survey Office's population-based measure of social consumption in health, across 65,932 sampled households.[91] Despite the utility of readily available statistics, one limitation was the lack of control over degree of data disaggregation and unable to be broken down further, which is a common challenge in extracting RWD relevant to reimbursement from health surveys. The effectiveness of treatment among HCV patients was derived based on analysis of the routine program data. The authors showed that treating HCV patients was cost saving and strongly recommended that schemes targeting Universal Health Coverage should include the treatment of HCV in their benefit packages. Secondly, the authors also showed that while the budget impact of a universal application of the new drug maybe be very high, it could be initially introduced for HCV patients with cirrhosis with the greatest clinical need. This led to a change in the standard treatment guidelines for treatment of HCV, not just in the state but also in the national HCV control program.

Box 3.7. Good practices when collecting data using health surveys

Key factors found in the literature to promote survey data collection from diverse populations are: (a) awareness of the importance of the research, (b) acceptability of participation through social support and community, and (c) access to participation through transportation provision, translation for multilingual populations, and financial incentives.[95] In addition, documentation of the following information improves rigor of health survey data collection and ensures that data is gathered in an ethical manner:[96]

- *How, where, how many times, and by whom were potential respondents contacted?*
- *How many people were approached and how many of those agreed to participate?*
- *How did those who agreed to participate differ from those who refused with regard to characteristics of interest in the study?* For example, their gender, age, and features of their illness or treatment (if any); how they were identified?; where they were approached? Sufficient information on demographics and characteristics of groups and individuals should be available, that would provide fair estimations that address the cost-effectiveness question of interest
- *How the survey was administered*? Self-administered (by post, internet, in person), face-to-face (computer assisted or paper-and-pencil), and telephone interview all are associated with their own pros and cons[95]
- *What the response rate was?* The number of usable responses as a proportion of the number of people approached is the best indicator to measure how much confidence can be placed in the results for the specific instrument, which reduces potential for bias as the response rate increases.

Box 3.7. (continued)

For further reading, see:

- Eurostat: Guidelines for the Development and Criteria for the Adoption of Health Survey Instruments. Eurostat 2005. Luxembourg[95]
- Kelley K, Clark B, Brown V, Sitzia J: Good Practice in the Conduct and Reporting of Survey Research[96]

3.3.5. Wearables and personal tracking devices

Wearables and personal tracking devices (including mobile technologies, health apps) capture person-generated health data (PGHD), which are defined as wellness and/or health-related data created, recorded, or gathered by or from patients to help address a health concern.[89] PGHD offers a rich RWD source of patient characteristics and outcomes collected during the course of individuals' normal routines and daily life. Examples range from patches for electrocardiogram monitoring, wrist-worn devices for activity monitoring and sleep assessment, to sensors with subcutaneous probes for continuous glucose monitoring. The different types of PGHD can be grouped into person-reported data, task-based measures, active sensor data, and passive sensor data.[89] The sources of RWD continue to expand and pose new possibilities for use in reimbursement and reassessment.

An example of the untapped potential of RWD from PGHD was the finding that Fitbits, wearable devices that measure resting heart rate and sleep time, hold promise in real-time flu surveillance at the US state level.[97] The weeks during which de-identified Fitbit users in five states had elevated heart rates and more sleep time tended to be those when influenza-like illnesses were most common in those states. When the Fitbit data were included in flu-intensity prediction models, correlations of the final models with the actual Centers for Disease Control and Prevention (CDC) influenza rates were excellent (0.97).

In the ongoing novel coronavirus (COVID-19) pandemic, human mobility studies have shown that aggregate and anonymized mobile phone location data can assist the modeling of the geographical spread of epidemics.[98] Digital contact-tracing technologies have also been deployed, such as Korea's smartphone app Corona 100m and Singapore's TraceTogether.[98]

Data quality issues

Despite the promise of these evolving technologies, the accuracy, usability, and robustness of these relatively novel sources of RWD need to be established, especially for acceptance of data collected in this way for regulatory or reimbursement purposes.

Table 3.5. Pros and cons of wearables and person-generated health data

Wearables and personal trackers: Pros	Wearables and personal trackers: Cons
• Routine collection of objective real-world data on the impact of an intervention, from individuals during their everyday life • Scalable data collection and extensive reach • Reduced barriers to participation • Lower costs than manual data collection	• Lack of completeness of data when patients fail to consistently wear, charge or sync a device • Accuracy, usability, and robustness needs to be established • Various law implemented by national and state governments to protect data collected through apps and sensors, which could affect information on data provenance

How are PGHD used in drug reimbursement decisions in Asia?

The use of PGHD to inform reimbursement decision remains uncommon worldwide. Effort is required to explore ways to protect patient privacy and need for additional regulatory approval for wearables data collection. In Japan, wearable devices for fall prevention among elderly people have been introduced.[99] While still a private sector service, it has potential to be reimbursed in the public through Long-term Care Insurance due to the care needs of a rapidly aging Japanese population.

Box 3.8. Good practices when collecting data using personal tracking devices

PGHD is an emerging field with massive data volume collected through a constantly increasing number of devices, apps, and websites. Lack of standard data definitions or formats and data validation are key challenges. Several sets of recommendations have been made for quality PGHD data collection. For instance, biostatisticians and relevant data scientists should be involved in all decisions involving protocol design and collection.[100] Only a minimum set of necessary data that can address the study endpoints should be collected, and the protocol should include strategies to monitor and optimize data quality. Devices and wearables should have suitable measurement properties that can measure the concept of interest in the target population. Hence, verification and validation checks of tools related to PGHD collection focus on establishing content validity, intra-device and inter-device reliability, concurrent validity, responsiveness of data, usability of device, and interpretability.[101]

For further reading, see:

- The Clinical Trials Transformation Initiative (CTTI) Mobile Clinical Trials Program: Advancing the Use of Mobile Technologies for Data Capture and Improved Clinical Trials[100]
- Duke Margolis RWE Collaborative RWD Quality Working Group: Determining Real-World Data's Fitness for Use and the Role of Reliability, Chapter 3[89]
- Critical Path Institute's Electronic Patient-Reported Outcome (ePRO) Consortium: Selection of and Evidentiary Considerations for Wearable Devices and their Measurements for Use in Regulatory Decision Making[101]

3.3.6. Prioritization of RWD variables in the local setting

Many health systems in Asia face challenges in collecting RWD due to the lack of infrastructure and human capacity to support data collection; lack of clinician, institutional or legislative support for data

collection; and lack of experience in its quality assessment and assurance. The working group agrees that RWE should be considered as supplementary evidence and is unlikely to replace evidence generated from clinical trials for reimbursement decisions. While the preceding sub-sections have discussed what RWD to collect and where to collect, not all RWD is feasible to collect in each local context, therefore, data collection efforts are likely to be more targeted to address specific research needs in each country. The REALISE working group was surveyed on the top locally-collected variables that their countries would prioritize over regional or international data. Overall, variables with high uncertainty or that are key drivers of cost effectiveness (such as costs and epidemiological data) are typically collected in the local context. Survey findings are presented in Table 3.6.

Table 3.6. Prioritization of RWD variables in reimbursement decisions for REALISE members

RWD variables	Q1. Top 5 variables preferred not to be taken from Europe, US, and outside of Asia (We would like to know which *region-specific* data are acceptable to you)	Q2. Top 3 variables preferred not to be taken from Europe, US, and Asia (We would like to know which *country-specific* data are most important to you)
Population characteristics	IN5, KR	KR, MY, SG, TW3
Intervention and control	IN4, MY, TW(I)	TW3(C)
Outcomes – Effectiveness	MY, SG, TW	
Outcomes – Patient reported outcomes (PROs)	KR, MY, SG, TW	KR
Outcomes – Safety	MY, SG, TW	
Outcomes – Cost	IN1, KR	IN1, KR, MY, SG, TW1
Epidemiological	IN2, KR, SG	IN2, KR, MY, SG, TW2
Adherence	IN3, KR, MY	IN3, KR, TW3

Notes:

- IN: India, KR: South Korea, MY: Malaysia, SG: Singapore, TW: Taiwan. Some countries responded with variables in ranked order, indicated by numbers after their country code.
- In Korea, Top 5 cannot be reduced to Top 3 as all five are necessary.
- In Singapore, regional/international data is acceptable to inform Singapore's HTA if it is considered generalizable to the local context. Collection of local data is not mandatory and is not required in most instances except for local costs because Singapore is a small country and there is limited incentive for companies to collect local data when patient populations are small. Local comparators may differ from comparisons in the trial, therefore indirect comparison may be required. Local epidemiological data/drug utilization patterns etc. may be collected by the HTA agency (ACE) to validate clinical trials and demonstrate whether they are generalizable to the local setting.

3.3.7. RWD sources for key RWD types

Table 3.7 summarizes the link between common RWD types (3.2) and their sources (3.3). For the interested reader, a list of real-world data sources available in Asia is provided in Appendix 8.1.

Table 3.7. RWD sources for common RWD types. Shaded areas indicate sources that could be used for given types of RWD.

What? RWD type (3.2)	Where? Source (3.3)				
	Disease and other registries	Claims databases	Health surveys	Electronic medical records	Wearables, personal tracking
Population characteristics	IN, JP, MY, SG, TW	IN, JP, KR, MY, SG, TW	IN, JP, KR, MY, SG, TW	IN, JP, MY, SG, TW	IN, JP, MY, SG, TW
Intervention and control	IN, JP, KR, MY, SG, TW	IN, JP, KR, MY, SG, TW		IN, JP, KR, MY, SG, TW	IN, JP, MY, SG, TW
Outcomes – Effectiveness	IN, JP, KR, MY, SG, TW	IN, JP, KR, MY, SG, TW		IN, JP, KR, MY, SG, TW	IN, JP, MY, SG, TW
Outcomes – Patient reported outcomes (PROs)	IN, JP, KR, MY, SG, TW		IN, JP, KR, MY, SG, TW		IN, JP, MY, SG, TW
Outcomes – Safety	IN, JP, KR, MY, SG, TW	IN, KR, MY, SG, TW		IN, JP, KR, MY, SG, TW	TW
Outcomes – Cost		IN, JP, KR, MY, SG, TW	KR	IN, JP, KR, MY, SG, TW	
Epidemiological	IN, JP, KR, MY, SG, TW	IN, JP, KR, MY, SG, TW	IN, JP, KR, MY, SG, TW	IN, JP, KR, MY, SG, TW	TW
Adherence	TW	IN, JP, KR, MY, SG, TW	IN, JP, MY, SG, KR	IN, JP, KR, MY, SG, TW	IN, JP, MY, SG, TW

Notes: IN: Indonesia, JP: Japan, KR: South Korea's non-covered services, MY: Malaysia, SG: Singapore, TW: Taiwan

3.4. How to collect? Study designs and good practice guidelines

A variety of study designs may be used to collect RWD and generate RWE. Researchers in Asia must consider which study design, is best suited to generate sufficient evidence for their reimbursement question, in view of the available data and the resources on hand. Fundamental design elements include, for example, the choice for prospective or retrospective studies, which hinges on considerations of access and completeness of existing RWD sources, and the time available to collect the data to inform the research question. This working group does not recommend any specific study design and advises that the choice of study design should be determined based on the data required.

3.4.1. Observational studies (cohort, case control, case series)

Observational studies are currently the most common sources of RWE and will likely remain so for the next few years until alternative data collection methods such as pragmatic clinical trials and single arm trials become more acceptable to inform HTA decisions. Observational designs observe the effect of an intervention through the natural relationship between exposure and outcome variables of interest. Prospective observational designs follow subjects forward in time and are suited for RWE development when researchers need greater certainty around the temporal relationship between the exposure and outcome variables. Retrospective designs rely on existing databases of routinely collected data where the study duration is already complete. This feature makes retrospective studies well-suited for investigations that require longer timeframes for the relationship between the exposure and outcome

variables to become observable. They are also typically less expensive and faster to conduct but may lack the variable control that characterizes prospective studies.

Box 3.9. Guidelines for good conduct and reporting of observational studies

The appropriateness of observational studies to obtain local RWE depends on the research question, current state of theory and knowledge, availability of valid measurement tools, and the proposed use of the data. Standards for observational studies focus on the design and conduct of observational research, by (a) having adequate sample size, (b) avoiding selection biases, (c) measuring the exposures and outcomes accurately and reliably, (d) ensuring that controlling confounders are considered in the design, and (e) planning the appropriate analyses.[47 102] Beyond conducting the research, better and more transparent reporting of observational studies is also encouraged make published evidence available for decision making. It is crucial to be transparent about confounding, bias, missing data, and generalizability to the local context as these may lead to research findings being misapplied.

For further reading on published standards on *design* and *conduct* of observational research:

- STRATOS Initiative: STRengthening Analytical Thinking for Observational Studies[102]
- Agency for Healthcare Research Quality (AHRQ): Developing a Protocol for Observational Comparative Effectiveness Research[103]
- Comparative Effectiveness Research Collaborative: Observational Study Assessment Questionnaire[104]
- European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP): Checklist for Study Protocols[105]
- ENCePP: Guide on Methodological Standards in Pharmacoepidemiology[106]
- U.S. FDA: Guidance for Industry Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment[107]
- Good ReseArch for Comparative Effectiveness (GRACE) Checklist and Principles[108 109]
- ISPOR Good Research Practices for Retrospective Database Analysis Task Force: Report (Parts I, II, and III),[110-112] Checklist for Retrospective Database Studies,[113] Prospective Observational Studies to Assess Comparative Effectiveness[114]
- Patient-Centered Outcomes Research Institute (PCORI): Methodology Standards[115]
- ChinaREAL Consortium: Technical Guidance for Designing Observational Studies to Assess Therapeutic Outcomes Using Real World Data (in Mandarin Chinese)[116]

For further reading on good *reporting* of observational studies, see:

- STROBE: Strengthening the Reporting of Observational Studies in Epidemiology[117]
- RECORD: REporting of studies Conducted using Observational Routinely-collected health Data.[118] RECORD was created as an extension to STROBE to include reporting items specific to routinely collected health data.

3.4.2. Pragmatic clinical trials

If RWD required for local evaluations is not accessible, and conducting a traditional RCT is anticipated to be either infeasible or less optimal for generating the desired evidence, subsequent decisions around study design may point to pragmatic clinical trials (PrCTs) as an option. In these designs, randomization

is maintained as a critical design element of the study that takes place within routine clinical care. Unlike observational studies, PrCTs are interventional studies, although their protocols are typically less restrictive than conventional trials in terms of the inclusion criteria for participants so that the study population is more representative of the target population likely to use the intervention in local practice. HTA of the Da Vinci robotic surgical system in China is an evaluation that has captured public interest and demonstrated the potential advancements that PrCTs can make in reimbursement decisions (Box 3.10).

There are limitations with generating RWD using PrCT despite the ongoing interest. Some difficulties include randomization and blinding, especially when it may be infeasible to blind the physician managing the care of the patient. Furthermore, blinding comes with operational complexities, which may make PrCTs less pragmatic and more expensive.

Box 3.10. Example of PrCT in Asia: HTA of the Da Vinci robotic surgical systems in China

Between 2010 to 2015, 34 Da Vinci surgical robots were purchased and installed in 30 tertiary public hospitals across China. In order to generate context-specific evidence and support further investment, HTA of the Da Vinci surgical robots was commissioned by the National Health and Family Planning Commission with a focus on real-world use of the technology in public hospitals.[119] A full HTA was conducted based on RWD from nine public hospitals in the central and eastern region. The authors designed a cohort study to assess the cost-effectiveness of Da Vinci-assisted prostatectomy (427 vs 421) and hysterectomy (247 vs 105) compared to standard laparoscopic procedures and concluded that the Da Vinci robots should not be procured in large numbers before requiring public hospitals to collect more evidence to demonstrate if they are cost effective in the local context.

Box 3.11. Good practices for the conduct of PrCTs

Pragmatic randomized clinical trials are conducted to answer the important question of how a treatment works in a 'real-world setting' among a heterogeneous 'real-world population'. As such, consideration during the design of the study should be given to how a heterogeneous patient population should be identified, in order to optimize the generalizability of the PrCT findings. There should be minimal inclusion and exclusion criteria.[120] The PrCT setting should also be representative of a routine clinical care setting rather than research site, and one strategy to improve long-term engagement is to involve the study sites in protocol development.[121] An appropriate comparison arm should be standard of care instead of a placebo arm. Finally, outcomes captured in a PrCT should reflect the information needed to make an informed decision by patients and physicians during routine care and have high relevance to the patient. One option is to collect relevant outcomes through direct extraction from the EMR. Carrying out blinding can be difficult for ethical reasons (e.g. administering sham surgery for a drug vs surgery comparison), and may in other cases be infeasible (e.g. blinding clinicians managing care of the patient). The following strategies are recommended when the patient or investigator cannot be blinded: (a) selecting only hard (not subjective) endpoints, (b) where event- or outcome-based endpoints are used, to have adjudication by blinded medical experts, and (c) the statistician or analyst can be blinded to the trial conduct and beyond for as long as possible.[122]

Box 3.11. (continued)

For further reading on good practices in PrCTs, see:

- Gamerman V, Cai T, Elsäßer A: Pragmatic Randomized Clinical Trials: Best Practices and Statistical Guidance[122]
- GetReal Work Package 3: Pragmatic Trials and Real World Evidence, Papers 1-8 published in the *Journal of Clinical Epidemiology*[123]
- Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M et al.: The PRagmatic-Explanatory Continuum Indicator Summary 2 (PRECIS-2) Tool – Designing Trials that are Fit for Purpose[124]
- ChinaREAL Consortium: Technical Guidance for Pragmatic Randomized Controlled Trials[125]
- Ford I, Norrie J: The Changing Face of Clinical Trials - Pragmatic Trials[126]

3.4.3. Single arm trials

Maintaining a control arm can be sometimes difficult due to the early phase of drug development or if there is a limited patient population due to the rarity of the disease under evaluation. One solution for single arm trials is generating an external control using RWD, known as a synthetic control arm.[127] A synthetic control arm uses patient level data from historical clinical trials in the same indication, where patients that meet the eligibility criteria and who have baseline characteristics that statistically match the current day experimental arm are carefully selected. This requires well-defined natural history and a covariate rich external control RWD dataset.

Single arm trials are used more for regulatory purposes than for HTA. South Korea does not recommend single arm trials for any use except safety. In the United States, precision oncology drugs have received accelerated regulatory approval based on evidence from early phase single arm trials.[128] A proof-of-concept study evaluated how well these external controls replicate the standard care arms in RCTs and found that results from an RCT for advanced non-small cell lung cancer (aNSCLC) could be approximated by substituting the EMR generated external control as a comparator.[128] Therefore, curated EMR data can serve as meaningful comparators in single arm trials in certain contexts.

Another use of single arm trials is in rare diseases. As an example, many epileptic conditions, especially if they affect children, are considered rare disorders with high unmet medical need.[129] In the history of development of anti-epileptic drugs as monotherapy, the US FDA required a placebo control. This was seen as ethically problematic because of the risk of injury from seizures. External controls were eventually permitted in 2006 to replace the placebo control group by FDA and since then, several anti-epileptic drugs have been evaluated using the single arm design.[130] [131]

3.5. Who to collect? Governance and accountability considerations for country adaptation

Box 3.12. Case study: Taiwan's National Health Insurance Research Database

The Taiwan National Health Insurance (NHI) Research Database covers more than 23 million residents (99.9% of the population) and is one of the largest nationwide population databases in the world. In 1995, a single-payer National Health Insurance plan was established by the government to provide Universal Health Coverage for its population.[132 133] This system's claims data are released as the National Health Insurance Research Database (NHIRD), and all data from primary outpatient departments and inpatient hospital care settings after 2000 are included. Some civil groups have filed lawsuits against the use of the NHIRD by the Ministry of Health and Welfare (MOHW) due to data privacy concerns, which led to the setup of the Health and Welfare Data Center (HWDC) to further strengthen the protection of health data. HWDC is a data repository that centralizes the NHIRD and other health-related databases. Cross-linkage of registries is comprehensive and linked to national surveys. Researchers and government alike are aligned in recognizing the utility of real-world health data as practical tools in medical research. Access can be granted to anyone conditional upon prior approval of the research or industry-sponsored project by an Institutional Review Board (IRB).[133] Safeguards to data privacy include deidentified datasets only accessible on-site, having statistical analysis syntax reviewed before access, and analyzed results examined before data export. There are no known NHIRD data breaches or leaks to date. [132]

The challenges of RWD collection in many countries extend beyond study design and data quality. Also relevant are issues of access and linkage involving data security, permissions from populations and database owners, and multiple stakeholders. Although the extensiveness and quality of Taiwan's database can answer a wide variety of research questions, not all in Taiwan are supportive of use of their personal information (Box 3.12). Civil groups have filed lawsuits against the use of the NHIRD by the Ministry of Health and Welfare (MOHW) due to data privacy concerns, which led to the setup of the Health and Welfare Data Center (HWDC) to further strengthen the protection of health data.[132 133] Taiwanese patients cannot opt out of inclusion in the database and this requirement is currently under review. This makes a strong case for the need to recognize that the standards and enforcement of privacy protection laws are country dependent. South Korea has some of the strongest data privacy laws in Asia, covering a person's image or voice, and linkage of health data is prohibited or limited to government operational purposes. Other countries may prefer an opt-in health records system. As these are contextual the working group encourages consideration of the following questions in the local adaptation of recommendations for data collection, including:[134]

- Which stakeholders are responsible for RWD collection, for which RWD source?
- Who bears the cost of RWD collection?
- Who manages and controls access to RWD?
- Who approves the ethics for research conduct, and who protects the privacy of RWD?
- Who can have access to RWD?
- What is the public's opinion on the use of their medical records for reimbursement and reassessment purposes?

3.6. Conclusion: General recommendations to improve RWD collection

3.6.1. Standardization of RWD variables between sources

Improvements to the collection and potential standardization of RWD variables will encourage further applications of RWE in reimbursement decision making. RWD format, completeness, and quality across registries, EMR vendors, and healthcare providers can vary significantly, and appropriate curation and validation are needed. In Malaysia, the Ministry of Health has initiated the Telemedicine Blueprint in 1996 and established National Health Data Dictionary to promote health information management and standardize health information in the country. Its terminology provides a common language that enables a consistent way of indexing, storing, retrieving, and aggregating clinical data across specialties and care settings. Thereafter, the Malaysian Health Data Warehouse Project in 2010 acts as a platform for the standardization and integration of health data from a variety of sources to better manage the health system, provide surveillance information and in addition provide a valuable source of data for research. The project is integrated into the ICT Strategic Plan in 2019. Taiwan has also set up the EMR Exchange Center (EEC) as an EMR gateway to facilitate the exchange of EMRs between different hospitals to avoid duplicating medications or examinations. Patients' informed consent are required to exchange the EMR. By the end of 2015, more than 80% of hospitals or clinics provide EMR exchange service through EEC.

Common data models that determine data fields of relevant capture can be implemented within specific disease areas to establish a broader consistency in data capture across providers and databases. An example of this is ASCO's CancerLinQ, which uses the National Cancer Institute (NCI) Metathesaurus and other vocabularies.[135] CancerLinQ aggregates data from EMRs through direct feeds and processes it through a series of transformations to standardize data elements across EMR systems.[136] Taiwan has also built common data models for collecting the clinical health information from clinical visits, discharge notes, surgery, pathology, examinations, blood tests, medical images, etc. The EEC and common data models are used as platforms for exchanging EMR and do not directly serve as sources of RWD. In Asia, national standardization of RWD can and should be an important consideration among countries that are already making plans for country-wide EMR systems (e.g. Bhutan, China).

3.6.2. Assess the costs and benefits of data collection

Evidence costs money. There are concerns in every resource constrained setting about whether the resources dedicated to an effort will be worthwhile. There is a need to prioritize RWD such that the benefits of collecting additional information can be expected to outweigh the costs. In Asian settings where data can be fragmented and incomplete, balance should be struck between the relevance of a new registry (or any other data) in relation to the burden of collection. The 'value-of-information (VOI) analysis' framework offers one approach to decision making, for when and what types of data to collect (see Box 3.13, for an example of VOI analysis in Thailand for coverage decision-making).[74] Formal use of decision analysis and VOI analysis can help determine whether an intervention should be adopted, whether additional evidence to further inform that decision is worth gathering, and what kind of information is of greatest value. The analysis evaluates the extent to which new evidence might improve expected benefits by reducing the chance for error and compares that improvement to the cost of the information.[8]

Box 3.13. Example of VOI analysis: Expected value of perfect information in palliative management vs peritoneal dialysis and hemodialysis for End Stage Renal Disease coverage decisions in Thailand

A study by Teerawattananon et al. examined the value for money of including peritoneal dialysis (PD) or hemodialysis (HD) for coverage in the universal health insurance scheme.[137] The results indicated that the government should not include dialysis services unless the social willingness to pay increases three times higher (700-750,000 Baht per QALY) than recommended by the commission on Microeconomics and Health. With uncertainty around input parameters of alternative treatment modalities, Figure 3.1 shows the expected opportunity loss of making a wrong decision for patients aged 50. The overall expected value of perfect information (EVPI) of treating 10,000 new ESRD cases per year and for a 10-year time period was highest (260,000 million Baht) at a ceiling ratio of 650,000 Baht per QALY.

The study also explored the effects of uncertainty around input parameters by looking at the value of obtaining further information on chosen parameters (partial EVPI) (Figure 3.2). Among the parameters, cost of PD and HD had the highest partial EVPIs.

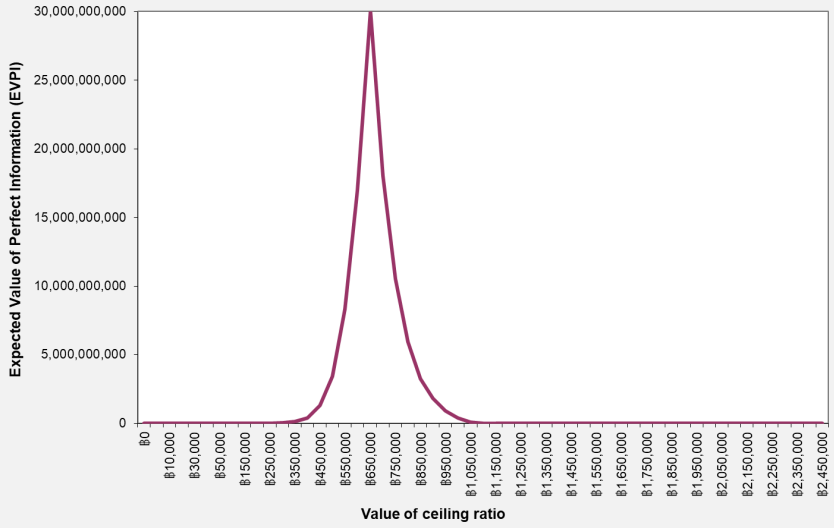

Figure 3.1. Population EVPI for a model using the 50 years age group

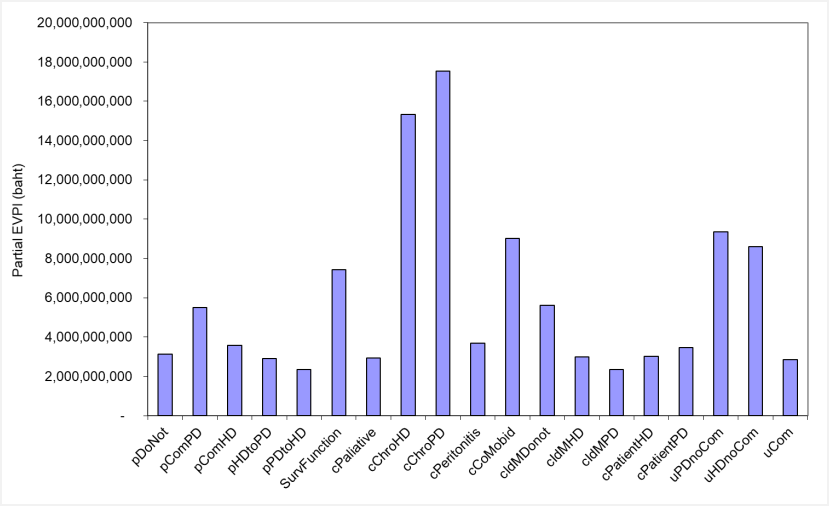

Figure 3.2. EVPI for input parameters. cChroHD: health care cost of HD, cChroPD: cost of PD

3.6.3. Develop incentives for quality capture of RWD

Incentives need to be developed for physicians and other providers, health care systems, payers, and patients to become invested stakeholders in the development and use of RWD and RWE.[138] These include financial incentives from public and private insurance payers for quicker processing of reimbursement claims if accurate data are captured by the EMRs;[139] for hospitals if they submit laboratory and test results efficiently; for reporting on patient outcomes; providing RWD for studies; and for further adoption of payments based on outcomes. In Korea, public funding for clinical research can be used to collect RWE such as pragmatic clinical trials / prospective cohort / retrospective cohort and others. NECA (Korea) manages US$20 million of research funding each year for extramural clinical research helpful from the public perspective.

Non-financial incentives include ensuring that RWD meets research needs and have clear value for those who collect them. The burden of RWD collection should be reduced by choosing meaningful RWD elements required for collection and that can be easily integrated into the workflow, thereby reducing errors due to time and resource constraints at the point of care. Clinicians must be convinced of the importance of the collected RWD in benefiting their patients and the delivery of clinical care.

3.6.4. Increase credibility of RWE relevant study designs (observational studies, PrCTs)

While many may perceive observational approaches as lacking credibility, researchers can exercise greater caution to overcome these concerns and improve the overall rigor of such studies. The use of checklists for good reporting practices is strongly encouraged (e.g. STROBE, RECORD for observational studies in 3.4.1), and submission of completed checklists is now required by some journals to validate manuscripts. Another way to increase credibility is by publishing detailed protocols of real-world studies in a public and online repository (as has been done for clinical trials, on ClinicalTrials.gov). This will enable researchers to see the study population, exposure and outcome variables, other key covariates, and the analysis plan that will be utilized before the study begins and can increase the validity of study results by ensuring that decisions made during the study process are not arbitrary, and that no data was mined to produce consequent study findings.[3] A public record encourages careful deliberation and accountability. Some observational studies have been registered on ClinicalTrials.gov, but not without some difficulty as the site is designed for RCTs.[3] One suggestion is to establish an online repository specific to RWE relevant study designs that an international audience can access. Comprehensive published protocols also allow for replication, where similar conclusions are derived from different data or with different analytical methods.

3.6.5. Balance patient data privacy protections and RWD as public good

Any consideration of patient medical records as a public good calls into question the safety and security of individual data. Data privacy laws are enforced to varying degrees in Asian countries, from requiring individual patient consent for every real-world study in South Korea, to permitted usage of de-identified and anonymized patient data in Taiwan and Singapore. Given that Japan is extremely conservative about data sharing, the process of anonymization before data can be shared with researchers or policymakers (e.g. the degree of disaggregation of registry data that is required before sharing) is an area of significant concern. We recommend that individual countries comply with their own countries' guidelines but at the same time, to promote active discussion of the tensions between access to RWD while ensuring adequate data protection, in order to arrive at a compromise between two needs. Block chain technology can be considered to link up patient consent such that it is required only once. In

South Korea, a privacy law was passed in 2012 that prohibits the linkage of data but permits government agencies to link limited data for their operational purposes.[140] This challenge of data privacy is mitigated in Korea by allowing researchers or industry to purchase anonymized sample claims datasets. It is also possible to apply masking techniques to create a synthetic dataset that replicates the key information needed for the specific research or policy question.[141] The Centers for Medicare and Medicaid Services (CMS) Data Entrepreneurs' Synthetic Public Use File (DE-SynPUF) is an example of a synthetic claims database which reflects real patient data but in a format that protects patients' identities so that the data can be used to train individuals in the appropriate use of claims data.[141]

A summary of the recommendations for Theme 2 on RWD collection are as follows:

Box 3.14. Recommendations for RWD collection

Recommendations on what RWD to collect for reimbursement decision making:

1. Collect patient variables that describe medical history/condition and practice variation across Asian countries to inform if patient groups and findings are comparable across different settings
2. Collect RWD on the intervention's optimal dosing, duration of treatment, waning of effect, and rate of discontinuation
3. Revisit trial efficacy data using person-level RWD once it is available to assess if the outcomes in the trial actually translate to clinically meaningful improvements for patients
4. Collect PROs where possible; the potential logistical barriers associated with data collection may be overcome by fostering collaborations between HTA agencies and academic units or clinicians
5. Identify important safety signals proactively with routine use of RWD such as administrative claims and electronic medical records (versus passive reporting)
6. Consider the relevance of cost items collected in relation to the burden of data collection and the perspective of the study (e.g. societal, public healthcare payer). Not all costs have to be included but sufficient justification should be given for including or excluding specific study cost items
7. Collect adherence data using a multi-measure approach to provide a better estimate of adherence, as existing methods of assessing adherence have their inherent limitations

Recommendations on where to collect:

1. Deliberate over the benefits and limitations of various sources in the local context, such as product or disease registries, claims databases, electronic medical records (EMRs), health surveys, or use of daily wearables and personal tracking devices.
2. Understand the potential reimbursement questions that may or may not be answerable because of the availability, access, and quality of the RWD sources. It may require integrating sources and leveraging the strengths of each

Box 3.14. (continued)

Recommendations on how to collect:

1. Determine the choice of study design (observational studies, pragmatic clinical trials) with considerations of access and completeness of local RWD sources, and the time available to collect the data to inform the research question
   Use a long enough timeframe of RWD collection beyond the pivotal RCT to allow longer-term safety and efficacy outcomes to be captured

Recommendations on who to collect:

1. Consider data governance and accountability in your local context, e.g. which stakeholders are responsible for RWD collection, who bears the cost of RWD collection, who manages and controls access to RWD, and who can access the data

Recommendations on improving the process of RWD collection:

1. Standardize RWD variables between sources
2. Assess the costs and benefits of data collection
3. Develop incentives for quality capture of RWD
4. Increase credibility of RWE relevant study designs
5. Balance patient data privacy protections and RWD as public good

4 Theme three: From RWD to RWE

Statistical approaches for using RWD in economic evaluations in HTA

4.1. Introduction and overview of the section

The previous theme of this guidance document discusses the different types of RWD, possible data sources and methods of collection. This section builds on the contents of Theme 2 and focuses on identifying the limitations of RWD and highlighting methods to analyze RWD for use in economic evaluations in the context of HTA, in order to improve the quality of RWE that is used for broader decision making in healthcare. The distinction between economic evaluation and HTA is made here because HTA has a broader scope beyond evaluation of the clinical and economic value of a new health technology. It also includes the social, distributional organizational and ethical issues of a health technology, which do not involve the use of statistical approaches. Nonetheless, it must be emphasized that RWD have many other uses in broader HTA aside from being inputs to an economic evaluation, most of which have been discussed in the previous section. While most examples given in the following parts of this section focus on the use of RWD in economic evaluations, many of the methods can be used to apply RWD in other parts of the HTA process such as specifying populations of interest, estimating incidence and prevalence for topic prioritization, and using information on adherence for measuring potential utilization, among others.

Several HTA guidelines recognize the potential biases that may arise from the use of RWD. The methods discussed in this section can be used to address limitations of RWD such as the lack of randomization of patients, lack of comparability, and missing data. Hence, when using data obtained from real-world sources to inform economic evaluations or any type of analysis, a good understanding of the potential biases is essential.

The aim of this theme is two-fold as described above: 1) to provide guidance on identifying limitations of RWD; and 2) to discuss statistical methods to address these limitations. While the target audience for this theme is analysts and researchers who prepare evidence submissions for HTA agencies or decision-makers, methods are summarized in a non-technical language supplemented by relevant examples to make them easily understandable by other knowledge users such as policy-makers and healthcare professionals who have an interest in the techniques employed in generating evidence for the submissions.

There is an abundance of literature on various statistical methods, tools and study designs to improve the use and reporting for observational studies. Many of these are cited throughout the chapter, and relevant links to these references are provided for further reading. Examples are also given to provide readers with practical examples of how observational studies are used in the context of economic evaluations and HTA. This section does not aim to provide a comprehensive review of all existing methods used to elicit and analyze observational data; rather, we seek to provide a high-level overview of how this type of evidence may be applied in the context of economic evaluations in HTA. For the interested readers, they should look up the cited references for details of the different methods.

4.2. How to use RWD to generate RWE?

4.2.1. Integrating RWD in economic evaluations

Economic evaluation is the comparative analysis of alternative courses of action in terms of both their costs and consequences. When applied to a broader context of healthcare decision-making, it involves a systematic process of identifying, measuring and valuing costs and outcomes of health technologies.[54] Majority of economic evaluations draw from various data sources, with RCTs as the gold standard, due to their high internal validity. In order to increase generalizability for use in broader policy making, modelling techniques are usually employed to utilize parameters coming from other sources such as observational studies.[142] While most guidance[142 143] asserts that RCTs should be the primary source for treatment effect parameters, studies have considered integrating real-world evidence to enhance the external validity of RCT-based evaluations or provide new and/or context-specific evidence when RCTs are not feasible.

There are several ways to integrate RWD into economic evaluations. Many published cost-effectiveness analyses use real-world data to extend trial results to estimate the treatment effect for the duration of the modelled time horizon, and use modelling techniques to incorporate different data sources into the analyses. Across the region, it is common for RWD to be used to derive local epidemiological estimates and cost data to inform an economic evaluation. These data are combined with published RCTs which are used to obtain efficacy estimates, and published studies from other jurisdictions to obtain utility data in the absence of local estimates. The study below provides a good example of how RWD was integrated into a model-based economic evaluation.

Box 4.1. Incorporating various data sources in a CEA: A study on denosumab for treating postmenopausal women with osteoporosis at high risk of fracture in Thailand[144]

A study was conducted in the Thai context, using a lifetime Markov model to compare denosumab vs. alendronate and no pharmacologic treatment for treating osteoporosis. Age-specific estimates of fracture incidence were derived from a local study, while treatment efficacy parameters were obtained from a foreign published meta-analysis. Costs were all reported in Thai Baht and were derived from a variety of sources: published local economic evaluations, a standard local cost list for HTA developed through a national survey, and some parameters calculated based on expert opinion. Health state utility estimates were also taken from foreign publications in the absence of local Thai data.

4.2.2. Using individual patient-level real-world data in an EE

The scope of the methods explained in this guidance is best applied to individual patient-level data (IPD). When compared to summary data or aggregate data (AD), IPD makes it possible to explore and adjust for individual patient characteristics that may influence outcomes and generate more accurate estimates of cost-effectiveness.[143] The National Institute for Health and Care Excellence (NICE) already sees the use of IPD as advantageous and has recommended stakeholders to use IPD whenever feasible. However, they also recognize that IPD is less accessible compared to AD since databases and other RWD sources are usually not designed with HTA in mind, hence, not research ready, or there is some unwillingness from the data custodian to share the data. Furthermore, most IPD provided to NICE is deemed commercial in confidence and has to be redacted from public documents, limiting the transparency of the data analysis. As an alternative, AD from registries or observational studies can be

used, but would heavily rely on the data managers to run such analyses, which may already be routinely done. While AD can be more easily accessed compared to IPD, it has limited usefulness in economic evaluations and analyses are limited to descriptive statistics instead of a more detailed analysis at the individual level. The information on patients' baseline characteristics is key to undertaking the statistical analyses described in the following sections and are needed to minimize the inevitable biases in RWD.

While IPD is seen as advantageous for use in economic models, it also has its drawbacks. Most IPD from the real-world are subject to stricter data privacy rules and obtaining relevant ethical approvals may be more challenging. Epidemiology-based literature claims that findings from observational studies or real-world sources are often less robust compared to RCTs, but economic evaluations usually cite this as a strength of the research.

Box 4.2. Using trial data, published literature and expert opinion to estimate clinical efficacy[145]

This study estimated the cost-effectiveness of sunitinib compared to interferon-alfa for treatment-naive patients with advanced and/or metastatic renal cell carcinoma in Singapore. A partitioned survival model was developed which applied the area under the curve to determine the mean time patients remained in the progression-free and progressive disease states. Overall survival and progression-free survival were extrapolated from the sunitinib pivotal trial. The duration of second-line treatment following disease progression was assumed to be consistent with the median progression-free survival reported in the literature from overseas jurisdictions and was supported by local expert opinion who confirmed that published estimates were applicable to Singapore's context.

4.3. What are the limitations of real-world data?

The use of RWD in cost-effectiveness analyses relies on a good understanding of the threats to its validity, limitations and possible solutions to address them. These threats affect either the internal or external validity of a study.

Internal validity refers to the ability of a study to measure what it is intended to do.[146] It can also be defined as the extent to which the observed difference in outcomes between two groups can be attributed to a certain intervention, instead of other factors.[147] The main determinant of this is how the treatment or intervention is assigned or given. The random allocation increases the internal validity as it minimizes selection bias and confounding.[147] For example, a well-designed RCT of a drug being studied for a particular condition will have high internal validity, since bias is minimized due to randomization of treatment assignment across two comparable groups, so measures of causal effect can be inferred from the results. On the other hand, observational studies and RWD have built-in biases, since many external factors or confounders influence the effect that can be estimated, given the absence of randomization and blinding.

External validity, on the other hand, refers to the ability to generalize results from a study to the population of interest. Selection bias, which is the distortion of estimates due to how patients are selected into the cohort or study sample, largely affects external validity. These usually arise due to poor choice in the selection of groups to be compared, or patients lost to follow-up and are no longer captured in a registry or through data collection.

4.3.1. Confounding

Confounding represents a mixing of effects between the treatment group and external factors that may also influence the outcome, potentially obscuring or distorting the relationship that can be inferred.[41] These factors that influence the association between a treatment and the effect may either be known or unknown.[42] The most common concern in observational studies and real-world sources like patient registries is confounding by indication. In RCTs, treatment assignment is random and not dependent on the patient's characteristics such as age, medical history or current health status. Most patients have a relatively good baseline due to the strict selection criteria and would usually have no severe comorbidities. However, in the real world, patients are more heterogeneous, and the prescribed treatment would depend on the choice of the physician. Other possible factors such as clinical guidelines informing local practice, availability of treatments (e.g. in hospital formularies), insurance coverage or affordability may influence the choice of treatment. This is usually termed as 'healthcare access bias' in the literature, alongside other types (e.g. confounding by contraindication, functional status, cognitive impairment, healthy user, and healthy adherence bias).[73 148]

Confounding can be minimized either through careful study design or through data analysis. The study cohort or the choice of the studied population can be restricted to current users / treatment group, and a comparison / control group. RWD studies can also restrict the selection of patients who are new users of a particular treatment, or only to adherent patients. Hard matching can be done based on observed confounders. Other approaches for adjusting for confounding are discussed in Section 4.4.

4.3.2. Selection bias

Another limitation to using RWD is selection bias, which arises when the observed subgroup of patients is not representative of the broader population of interest.[43] This is a key issue when using IPD from real-world sources and is a threat to both the internal and external validity of the study and its generalizability to a larger population. It is important to note the difference between confounding and selection bias and that methods to control for the former may not address the latter. In some jurisdictions, patients may be involved in the HTA process. It is important in such circumstances to avoid an upward selection bias, called "Methuselah bias",[149] where surveyed patients are the only people who are left alive after using the technology, and they actually lie on the end of the survival curve rather than represent the average patients who used the technology.

Since the majority of HTA decisions in Asia are made at the national level, there is a need to ensure that economic evaluations conducted using RWD are generalizable to the whole population likely to receive treatment. Most RWD come from single institutions or service delivery networks and they cover only a limited number of patients, usually confined to a specific practice or limited to a smaller geographical area. Integrating these smaller databases or registries to create a national repository would increase external validity when used to inform HTA. This is usually the case for claims databases of public payers as all patients are captured in one central database. On the other hand, if integration of databases is not yet possible and the RWD comes from a single facility, researchers must justify the choice, explain why it is appropriate and how representative it is of the broader population (Box 4.3).

Box 4.3. Justifying using data from a single center to make a decision at the national level[150]

An evaluation was conducted in the Philippines to compare treatment protocols for using hemodialysis first, peritoneal dialysis first, or undertaking a pre-emptive kidney transplant in patients with end-stage renal disease from the health system perspective. Data came from a variety of sources: epidemiological data (baseline transition probabilities, baseline distribution, mortality data) was derived from the Philippines renal disease registry, costs were from the national health insurance claims database, while outcomes were obtained from a single facility, the National Kidney and Transplant Institute (NKTI). NKTI is a tertiary hospital located in Metro Manila and is the national referral center for renal care. While patients in NKTI may not be exactly similar to patients in the other regions and island groups of the Philippines, clinicians agreed during a stakeholder consultation meeting to use data from NKTI since they have the largest number of patients with end-stage renal disease in the country, and the hospital is seen to hold the highest standards of renal care. Health utilities were obtained from a patient survey in NKTI, as well as other clinical parameters (survival, complication rates, transplant mortality) that were used as inputs to develop the economic model for the study.

Table 4.1. Types of confounders and potential solutions on how to control for them

Type of confounder	Example	Strategy
Measured	Age and sex	Regression Matching Stratification Propensity Scores
Unmeasured but measurable	Smoking status Body mass index Disease severity	External adjustment Proxy measures Imputation
Unmeasurable	Frailty	Self-controlled design Instrumental variable Mendelian randomization Regression discontinuity design Sensitivity analysis

Adapted from: Norgaard et al.[151]

4.3.3. Missing data

Having missing data is not uncommon in most real-world studies. Often, administrative errors occur, or it may not be feasible to collect all data points at one time. In this case, limiting the analysis to only complete entries or cases may not be a good approach, as this might decrease the sample size significantly. It will also affect the generalizability of the findings as the data may not be missing at random, and there could be systematic differences between those with and without missing data. Approaches for dealing with missing data are discussed in the next sub-section.

It is essential that decision makers understand these potential threats to the study validity when assessing and appraising data from RWD sources or observational studies. Using RWD as supplementary evidence alongside RCTs, can improve the internal and external validity of the available evidence and address uncertainties that may arise when conducting the HTA or during the decision-making process.

4.4. What are the statistical methods to analyze real-world data?

There are different statistical methods available to analyze RWD. Which method to consider depends on several factors such as data availability, data quality, sample size, and follow-up time. Each method has strengths and limitations. The following section highlights commonly used/mentioned approaches when analyzing RWD.

Box 4.4. Recommended readings

For readers who would like a refresher on basic statistics, kindly refer to Morshed et al.'s **'Analysis of Observational Studies: A Guide to Understanding Statistical Methods'**.[41] The text provides a simple and concise overview of key concepts in statistics such as populations and distributions, estimation and hypothesis testing and different analytical techniques for observational studies.

The **STROBE Statement** (Strengthening the reporting of observational studies in epidemiology)[152] [153] and its extended version **RECORD** (Reporting of studies conducted using observational routinely collected data),[154] provide a checklist of 22 items that are considered essential for good reporting of observational studies.

There are existing guidance documents to support analysis of RWD. For example, NICE Decision Support Unit (DSU) Technical Support Document 17[155] on the use of observational data to inform estimates of treatment effectiveness proposes an algorithm to help inform the selection of appropriate methods for the analysis of comparative IPD. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) also has a Special Task Force on Real-World Evidence in Health Care Decision Making that has made recommendations on good practices for using RWD for estimating treatment and/or comparative effectiveness.[3] For survival outcomes, the CanREValue collaboration methods working group in Canada[156] also has a forthcoming paper on using statistical methods for survival outcomes. Our document seeks to briefly summarize the information from these published sources and to provide examples using Asian studies whenever possible.

4.4.1. Propensity score analyses

Propensity score matching

To understand propensity score matching (PSM), the concept of matching and a propensity score need to be understood separately first. **Matching** involves finding untreated or unexposed patients and similar patients who received the intervention. The matching can be done based on observed characteristics of the patients such as age, sex, and duration on treatment.[157] It is like simulating your own treatment and control group in an RCT. On the other hand, a **propensity score** is an estimate of the likelihood of being exposed based on a given set of characteristics or covariates. It is usually estimated by logistic regression to calculate the chance of receiving or being assigned a treatment. So, in PSM, instead of matching with observed characteristics, groups are matched based on their propensity score. Most studies use propensity scores when comparing two treatment groups, hence it is less attractive when multiple treatment strategies are being considered.[73] However, advances in methodologies have shown that PSM can be extended to three or more treatments.[158] Lastly, propensity score methods must assume that there is no unobserved or unknown confounding. However, this assumption is likely to be violated because many patients are excluded. Hence, the matched cohort only represents a smaller subset of the overall eligible population.

Box 4.5. Application of Propensity Score Matching in an economic evaluation[159]

A study in Singapore used real-world data to estimate the cost-effectiveness of an adherence-enhancing program in the rheumatology division of a public tertiary hospital. Data was collected from the hospital's Patient Affordability Simulation System (PASS) database, a 10-year standing database of historical patient electronic medical records (EMR). To create a comparable control group from the database, propensity score matching was used based on measured confounders. Matching variables chosen were demographics (age, sex, ethnicity), comorbidities (Charlson comorbidity index), serum urate at baseline, gout hospitalization history at baseline and gout medication use (NSAID, colchicine, glucocorticoids). PSM was conducted using the MatchIt package in R. The authors adopted a nearest neighbor, 1-to-1 method of matching and adopted a 'caliper' of 0.25 standard deviations, known as the maximum permitted distance between matched subjects.

Inverse probability weights

Inverse probability weights (IPW), is a method that uses the propensity score function, and applies the same assumption of no unmeasured confounding. IPW using propensity scores researchers to obtain unbiased estimates of average treatment effects,[160] however, the usefulness of this method depends on how well the model for propensity score predicts the probability of treatment.[161] When comparing two groups, the average treatment effect would correspond to the difference in weighted means. When using IPW, the weighted means are calculated using the inverse of the propensity scores as weights. This new estimator of treatment effect corrects for missing data and compensates for this by giving more weight to the small number of observations which appear in one group but have a small probability of being found in the other group. The strength of this method is its ability to directly see confounding through the distribution of the propensity score. However, there is a chance for the effect to still have bias due to unknown confounding.[161] It appears that the use of IPW in economic evaluations is not common in Asia. However, we are seeing emerging use of IPW in studies of treatment effectiveness.[46 162]

Box 4.6. Application of Inverse Probability Weighting in economic evaluation[163]

This study, conducted from the perspective of a provincial healthcare system (Ontario, Canada), used linked administrative databases to evaluate value for money of adding rituximab (RCHOP) to the standard chemotherapy regimen (CHOP). The Ontario Cancer Registry (OCR) was the main source of RWD used in this study. The authors applied the Inverse Probability Weighting (IPW) nonparametric method to adjust for censoring in the cost data. To do this, the study period was partitioned and the total observed cost in each time interval was divided by the probability of not being censored at the beginning. Mean costs were estimated by summing the totals across all intervals and then dividing the sum by the sample size. Weights were constructed separately for each treatment group, and 3-year and 5-year costs were estimated. Overall survival was estimated using the Kaplan-Meier method for each cohort. Survival was defined as the time from diagnosis to date of death from any cause or the end of the study timeframe. To estimate mean survival time, the survival data were partitioned the same way as the cost data and determined using the same IPW methodology. The study also used propensity scores to match the RCHOP vs CHOP groups using the following baseline characteristics: sex, income quintile, adjusted clinical group score, and histological diagnosis code (primary).

4.4.2. Covariate adjustment

Adjusting using regression aims to reduce bias related to confounding by introducing known confounders in a regression model to estimate the final outcomes. Similar to the first two methods described above, conducting regression adjustments using multiple variables or covariates assumes that there are no unknown confounders. In addition, this approach similarly requires patient-level information on baseline demographics and disease characteristics. The outcome of interest is the dependent variable, and the baseline characteristics and the intervention are independent variables. Multivariable regression allows the association between the dependent and independent variables to be estimated which is called the adjusted effect, while controlling for the influence of other independent variables.[164] The most commonly used methods for this are linear and logistic regression. However, when modelling time-to-event outcomes, survival analysis techniques such as Cox proportional hazards model must be applied instead. A strength of this approach is its efficiency in being able to simultaneously adjust for multiple confounders, and its ability to easily assess the effects of each confounder. However, the quality of the estimates depends on the model fit and its corresponding assumptions.[41] Below is a list of multivariable adjustment models for common types of outcomes:

Table 4.2. Appropriate multivariable adjustment models for common types of outcomes[41]

Type of outcome	Example	Model	Estimate of effect
Binary	Prevalence of postoperative infection	Logistic regression	Odds ratio
Continuous	Range of motion or functional outcome score (i.e. SF-36)	Linear regression	Mean difference
Time-to-event	Time to reoperation following total hip arthroplasty	Cox proportional hazards	Hazard ratio
Rate	National rates of total joint replacement	Poisson regression	Rate ratio

A more specific approach, called regression adjustment, as recommended by the NICE DSU Technical Support Document 17 requires a two-step process to adjust for confounding. First, two regression equations are estimated, one for the treatment group and another for the control group. Afterwards, individual differences in the predictions for the two outcomes are averaged across individuals. This method requires the correction of standard errors in the second step considering that the potential outcomes in the first step are estimated, which can be addressed by any of the standard methods for two step estimators.

4.4.3. Instrumental variables

Unlike the previous statistical controls mentioned above that can only control observed characteristics, the use of instrumental variables can control for both observed and unobserved confounding. An instrumental variable (IV) is a factor associated with the treatment but only correlated with the outcome through its effect on the treatment.[165] It is also known as the exclusion restriction, since the IV is exclusive to the treatment choice equation. A diagram is provided below to help describe the relationship of the IV with the treatment, confounding variables and the outcome. There must be no association between the IV and the confounders. Similarly, the IV's effect on the outcome must not be direct and should only be through the treatment.

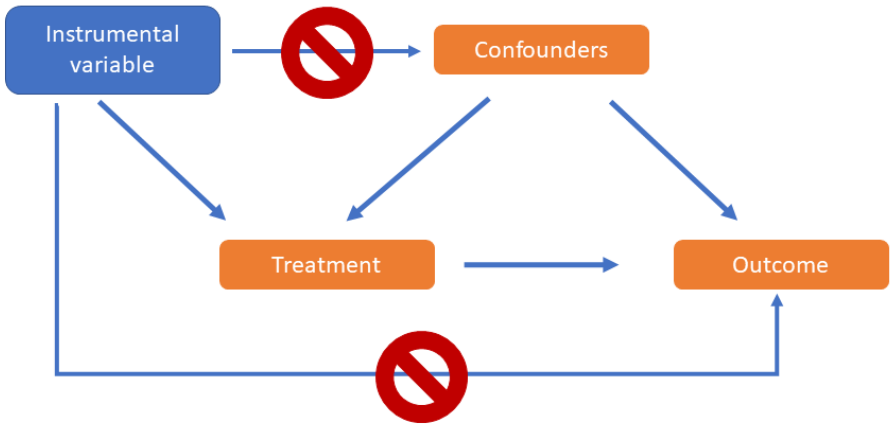

Figure 4.1. Instrumental variable[41]

In practice, finding an instrument that satisfies the criteria given above is challenging. However, once an instrument is identified, it offers the possibility to control for both known and unknown confounders in the estimate of the outcome. The downside, however, is that it is usually difficult to test all IV assumptions, and the inference is restricted only to the patients whose treatment is influenced by the IV ("compliers"), thus the estimate is unlikely to be generalizable to the population of interest. A tutorial on the use of instrumental variables in pharmacoepidemiology was developed by Ertefaie et al.[166] and details a step-by-step guide on how to use IV to adjust for unobserved confounding. A list of examples of IV is also provided below.

Table 4.3. Examples of instrumental variables in published studies[151]

Association examined	Instrumental variable	Reference
NSAID treatment for persistent ductus arteriosus and mortality and moderate/severe bronchopulmonary dysplasia	Institution variation in NSAID treatment frequency	Slaughter et al.
Use of non-invasive ventilation in patients with pneumonia and 30-day mortality	Differential distance, i.e. The difference between the distance from a patient's residence to the nearest hospital of any type	Valley et al.
Second-generation versus third-generation oral contraceptives and risk of venous thromboembolism	Proportion of prescriptions for third-generation oral contraceptives by the general practitioner in the year preceding the current prescription	Boef et al.
Readmission destination and risk of mortality after major surgery	Regional index hospital readmission rates	Brooke et al.

Box 4.7. Using an instrumental variable to estimate real-world effectiveness of hematopoietic transplant among elderly individuals with multiple myeloma[167]

The study compared the effectiveness of hematopoietic stem cell transplant (HSCT) using information from the Surveillance, Epidemiology, and End Results (SEER) database in the United States. A mix of approaches were used, such as PSM-based analysis and IV analysis. The authors used geographic variation as their instrument and created the IV in two steps. First, they used logistic regression to predict the probability of having an HSCT as a function of age, race, sex, marital status, urban status, area level poverty, and comorbidity score. Second, for each year and health service area, they calculated the difference between the observed and predicted number of transplants. They found that geographic variation was a valid instrument because there was statistically significant variation across geographic areas and a strong association between receiving treatment and location. Location was also deemed independent of the patient's characteristics. The IV analysis was conducted using the two-step residual inclusion method with hazard rates estimated in the second stage equation. This analysis produced a hazard ratio that was consistent with the results of other analyses conducted (multivariable analysis, PSM), and showed that HSCT improved overall survival for older individuals with multiple myeloma.

4.4.4. Imputation for addressing missing data

Various imputation methods can be used to address missing data. **Simply imputing mean values where variables are missing is not recommended** since patients from the real-world are likely to be more heterogeneous compared to controlled settings. IPD usually has a higher rate of missing data, hence multiple imputation methods are usually applied as they not only impute the missing values, but also account for the uncertainty associated with the imputed values. Manca et al. described in detail the methods to handle missing data in individual patient-level CEAs. While the methods described are along the context of cost-effectiveness analyses alongside RCTs, the same approaches can be applied to IPD from RWD.[168] Any study that collects IPD, whether in an RCT or real-world setting suffers from issues associated with incomplete observations. There are many types of missing data, and the methods to address them depends on the context why missing data has occurred. The most common approach is multiple imputation (MI) for data missing at random, since this method not only imputes missing values but also accounts for the uncertainty associated with the imputed values through the creation of multiple datasets.[73] MI requires a three-stage process:[169]

1. **Imputation:** A dataset is created by simultaneously filling in each of the incomplete observations with values generated from their predictive distributions. This variation among the imputations reflects the uncertainty with which the missing values can be predicted from the observed values.
2. **Analysis:** The datasets are subsequently analyzed using standard methods.
3. **Pooling:** The results from step 2 are combined to produce estimates and confidence intervals that incorporate the missing data uncertainty.

Box 4.8. Multiple imputation to address missing data when using various databases in an economic evaluation[170]

A combination of curative and combination treatment strategies were evaluated from the perspective of Ontario, Canada's provincial health system. The Ontario Cancer Registry (OCR) was the main source of IPD RWD and was linked to the Discharge Abstract Database maintained by the Canadian Institute for Health Information, the Ontario Health Insurance Plan (OHIP), the Ontario Drug Benefit Program and the Canadian census data. The IPD contained information on sociodemographic, screening, treatment and clinical factors. Multiple imputation (MI) was used to impute the values for variables with a high degree of missing data such as cancer stage at diagnosis, birth country, and Charlson-Deyo comorbidity index. Five independent draws from an imputation model were used to create five completed datasets and results were combined to obtain one imputation inference.

Another approach to handle missing data is through the use of stratification. When dealing with continuous data, an indicator variable for missing value can be used; alternatively, a separate category or stratum for "missing" can be added for discrete variables.[171] This approach is relatively easy to implement and does not make an assumption on the type or mechanism of missingness.

There are many other ways to deal with missing data, and below is a list of other methods to consider:[172] [173]

- (Simple) Substitution
- Mean imputation
- Hot deck imputations
- Cold deck imputation
- Regression imputation
- Stochastic regression imputation
- Interpolation and extrapolation

4.4.5. Net benefit regression

The net-benefit regression (NBR) framework was introduced by Stinnett AA. and Mullahy K. as a new framework for the analysis of uncertainty in cost-effectiveness analyses. It can be an alternative or additional analysis to the estimation of the incremental cost-effectiveness ratio (ICER) and can be applied to individual patient level data from an RCT, or data sets collected from the real-world.[174] The net-benefit framework reformulates the cost-effectiveness problem and avoids the reliance on the CER and the statistical issues associated with it such as negative ICERs. With the use of IPD, individual level covariates' impact on the marginal cost-effectiveness of an intervention can be estimated. The equation for estimating the NMB is found below:

$$NMB_i = E_i \times \lambda - C_i$$

Where NMB_i is the net monetary benefit for each patient *i*, and E_i is the individual effectiveness parameter, C_i is individual cost, and λ is the willingness to pay (WTP) value. Various estimates of WTP can be used based on a country threshold or individual level estimates of willingness-to-pay. One advantage of using this framework is that the generation of the outcome, the net benefit statistic, can be generated through basic regression methods such as the Ordinary Least Squares (OLS). This approach recognized the limitation of the ICER to give an unequivocal treatment recommendation.[175]

Moreover, this framework provides methods to characterize the uncertainty of cost-effectiveness findings (such as cost-effectiveness acceptability curves and 95% confidence intervals).[176] In addition, the use of NBR also allows researchers to conduct subgroup analyses and adjust for known confounders.

Box 4.9. Example of an NMB analysis of laparoscopy versus open colectomy for colon cancer in Taiwan[177]

A study in Taiwan compared the effectiveness and costs associated with laparoscopic and open colectomy from the perspective of the National Health Insurance (NHI) system. A cohort was observed by linking the Taiwan Cancer Registry, claims data from the NHI and the National Death Registry. Cohorts were matched using propensity scores to estimate overall survival, recurrence-free survival and disease-free survival. Health outcomes and net monetary benefits (NMB) were verified using multivariate mixed-effect models. The NMB estimate was calculated based on a range of threshold values from 1 – 3 GDP per capita. On all three outcomes, the model demonstrated that laparoscopic colectomy (LC) was more cost-effective than open colectomy, regardless of the WTP level. Even when WTP was 0, there was a 92% chance of LC being cost-effective.

4.5. What other topics to consider when analyzing RWD?

As this guidance document does not seek to provide an exhaustive list of methods for dealing with the limitations of RWD, there are other topics that we have not considered at length.

For limitations of RWD, in addition to confounding and selection bias, there are other limitations to consider:

- Immortal bias[178]
- Competing risk[179]
- Time-varying confounders[180]
- Left truncation[181]

For statistical methods to analyze RWD, there are other methods to consider:

- Genetic matching[182-185]
- G methods[186]
- High dimensional propensity score[187]
- Multistate models[188 189]

Additionally, below are additional topics to consider when planning to analyze patient-level real-world data

- Censoring when each individual has different follow-up times[190 191]
- Sample size calculation for economic evaluations using person-level data (as both costs and effects need to be considered in the calculation)[192 193]

4.6. Conclusion

Many more statistical approaches were not mentioned in this chapter, and the available methods increase everyday as researchers continue to find new ways to address potential biases and limitations.

While most of the methods discussed in this chapter may go beyond the toolkit of a typical HTA researcher, it is helpful to know and understand the methods that are available in the event they are required at any stage in your future research. It is important to review the scenarios when to use RWD in the first theme of this guidance, to ensure that RWD will be useful in your planned analysis, and the methods described here will be worthwhile undertaking to improve the rigor of the analysis. Finally, it is important to note that there is a limit to what statistical analyses can do. Their usefulness may be limited if the RWD collected is of poor quality and derived from unreliable sources. Users of this guidance are encouraged to have a good understanding of the best practices for collecting RWD as they are complementary to the methods reported in this section and will likely become more useful and relevant in the field of economic evaluation and HTA.

Below are some of the procedural recommendations for data analysis:

Box 4.10. Good practices for data analysis

1. Clearly specify the primary and secondary outcome measures that are used in analysing RWD
2. Discuss the rationale for adjusting for confounding and biases and justify the choice of method used
3. Where possible, use more than one approach to undertake the data analysis, as a tool for sensitivity analysis (e.g. using both covariate adjustment and PSM and comparing the results thereafter)
4. Include statistical codes and packages used in the analysis as supplementary material / annex to the published study to improve transparency of the work undertaken

5 Moving forward: What we will address in future versions

We conclude by affirming the need for better quality RWD and RWE to inform HTA in Asia. There are several opportunities to use RWD/RWE that can provide clear advantages for understanding outcomes of drug therapies in the real-world setting, especially in diseases areas involving diverse patients whose treatment regimens and clinical needs are not driven by trial protocols. Many RWD sources can contribute to RWE efforts but steps for mitigating erroneous data, standardizing its collection, and ensuring quality management/assurance will need to be developed and applied across efforts in the Asian setting. Study designs involving RWD can combine benefits of collecting data from real-world settings while incorporating best practice methods (e.g. randomization methods from traditional RCTs). We emphasize important limitations of RWD that users should be aware of, including potential for biases, lack of comparability, and missing data; and introduce techniques for analysis that are appropriate for the evolving data environment to ensure that best use can be made of data that is available. A summary of our guidance recommendations can be found below:

Table 5.1. Summary of recommendations

Theme one

Recommendations on when to use RWD/RWE:

1. Consider RWD/RWE when RCTs are of poor quality, or impossible to conduct for ethical reasons or due to small numbers (e.g. for rare diseases)
2. Consider use of RWD/RWE to contextualize and localize economic models, extrapolate RCT data beyond trials, and for price negotiations based on real world outcomes
3. Note that RWD and RWE is generalizable to routine clinical practice only if the data represents the target clinical population reflective of the local context

Theme two

Recommendations on what RWD to collect for reimbursement decision making:

1. Collect patient variables that describe medical history/condition and practice variation across Asian countries to inform if patient groups and findings are comparable across different settings
2. Collect RWD on the intervention's optimal dosing, duration of treatment, waning of effect, and rate of discontinuation
3. Revisit trial efficacy data using person-level RWD once it is available to assess if the outcomes in the trial actually translate to clinically meaningful improvements for patients
4. Collect PROs where possible; the potential logistical barriers associated with data collection may be overcome by fostering collaborations between HTA agencies and academic units or clinicians
5. Identify important safety signals proactively with routine use of RWD such as administrative claims and electronic medical records (versus passive reporting)
6. Consider the relevance of cost items collected in relation to the burden of data collection and the perspective of the study (e.g. societal, public healthcare payer). Not all costs have to be included but sufficient justification should be given for including or excluding specific study cost items
7. Collect adherence data using a multi-measure approach to provide a better estimate of adherence, as existing methods of assessing adherence have their inherent limitations

Table 5.1. Summary of recommendations (continued)

Recommendations on where to collect:

1. Deliberate over the benefits and limitations of various sources in the local context, such as product or disease registries, claims databases, electronic medical records (EMRs), health surveys, or use of daily wearables and personal tracking devices.
2. Understand the potential reimbursement questions that may or may not be answerable because of the availability, access, and quality of the RWD sources. It may require integrating sources and leveraging the strengths of each

Recommendations on how to collect:

1. Determine the choice of study design (observational studies, pragmatic clinical trials) with considerations of access and completeness of local RWD sources, and the time available to collect the data to inform the research question
2. Use a long enough timeframe of RWD collection beyond the pivotal RCT to allow longer-term safety and efficacy outcomes to be captured

Recommendations on who to collect:

1. Consider data governance and accountability in your local context, e.g. which stakeholders are responsible for RWD collection, who bears the cost of RWD collection, who manages and controls access to RWD, and who can access the data

Recommendations on improving the process of RWD collection:

1. Standardize RWD variables between sources
2. Assess the costs and benefits of data collection
3. Develop incentives for quality capture of RWD
4. Increase credibility of RWE relevant study designs
5. Balance patient data privacy protections and RWD as public good

Theme three

Procedural recommendations for RWD analysis:

1. Clearly specify the primary and secondary outcome measures that are used in analyzing RWD
2. Discuss the rationale for adjusting for confounding and biases and justify the choice of method used
3. Where possible, use more than one approach to undertake the data analysis, as a tool for sensitivity analysis (e.g. using both covariate adjustment and PSM and comparing the results thereafter)
4. Include statistical codes and packages used in the analysis as supplementary material / annex to the published study to improve transparency of the work undertaken

The aim of this guidance is to provide a framework for anyone involved in HTA in Asia to generate and use RWD/RWE and improve the quality of such evidence when used to inform reimbursement activities. Recommendations from this guidance may be useful for some countries to include in their local HTA methods and process guidelines to clearly explain the role of RWD/RWE in informing HTAs. It is also important to promote dialogue among stakeholders regarding the use of RWD/RWE, and to have commitment from research funders as well as providers and professional associations at all levels (national and regional) to support the development of infrastructure for collecting and analyzing RWD in the region.

Therefore, while we have prepared the full version of the report with HTA agencies and 'doers' of HTA in mind, the working group acknowledges that different audiences may have different expectations and needs. Two abridged reports will be produced for (a) policy and decision-makers, industry, and patient groups who use HTA; and (b) clinicians and research staff who generate RWD at study sites.

The version for policy makers, funders, and other stakeholders, who may be unfamiliar with the concepts of RWD and RWE, will focus on the scenarios they can consider its use; governance, accountability and ethical considerations; and general recommendations for RWD collection. In the report intended for clinicians and research staff collecting RWD, emphasis will be on various aspects of the collection itself (what, where, and how to collect); include ethical considerations for patient and data privacy in handling data; and conclude with a short overview of data analysis and statistical methods the data they collect is used for.

While this guidance attempts to be as comprehensive as possible, there are other aspects of RWD generation and analyses that we have not covered in detail. For example, reimbursement for medical devices and companion diagnostics, and statistical methods on extrapolation of survival. These topics may be considered in future work of the REALISE working group. Experiences of the REALISE working group members in using this guidance will also be documented so that future updates of this guidance document will reflect the feasibility and outcomes of our recommendations.

Figure 5.1. REALISE guidance document infographic

6 Glossary of terms and list of abbreviations

6.1. Glossary

Accuracy: In the context of a study, the quality of a measurement (e.g. the mean estimate of a treatment effect) that is correct or that reflects the actual effectiveness of the treatment.[2]

Clinical effectiveness: The benefit of using a technology, program or intervention to address a specific problem under general or routine conditions, rather than under controlled conditions, for example, by a physician in a hospital or by a patient at home.[2]

Confounding: A mixing of effects between the treatment group and external factors that may also influence the outcome, potentially obscuring or distorting the relationship that can be inferred. These factors that influence the association between a treatment and the effect may either be known or unknown.[41]

Economic evaluation: The comparative analysis of the costs and consequences of two or more possible options. Depending on whether the consequences are expressed as monetary, physical or qualitative variables, the analysis may be a cost-benefit, cost-effectiveness or cost-utility analysis.[2]

Efficacy: The benefit of using a technology, program or intervention to treat a particular problem under ideal conditions—for example, in the context of research in a laboratory or a rigorous protocol for a randomized clinical trial.[2]

Efficacy-effectiveness gap: The difference in benefit–risk between effectiveness and efficacy.[2]

External validity: The ability of a research design to provide findings that can be generalized to other populations, contexts and periods.[2]

Health-related quality of life: The measures of the impact of an intervention on patients' health status, extending beyond the traditional measures of mortality and morbidity to include dimensions such as physiology, function, social life, cognition, emotions, sleep and rest, energy and vitality, health perception and general life satisfaction.[2]

Health technology assessment: The systematic evaluation of the properties and effects of a health technology, addressing the direct and intended effects of this technology, as well as its indirect and unintended consequences, and aimed mainly at informing decision making regarding health technologies. HTA is conducted by interdisciplinary groups that use explicit analytical frameworks drawing on a variety of methods.[2]

Medication possession ratio: The sum of the days' supply for all fills of a given drug for a specified time period divided by the number of days in the period.[77]

National Health Insurance Research Database, Taiwan: One of the largest nationwide population databases in the world, covering approximately 23 million residents in Taiwan. The NHI program was established in 1995 to deliver universal coverage provided by a government-run, single-payer compulsory insurance plan, covering more than 99.9% of the population. This system's claims data are released as the National Health Insurance Research Database (NHIRD), and all data from primary outpatient departments and inpatient hospital care settings after 2000 are included. The Health and Welfare Data Center (HWDC) was set up by the Ministry of Health and Welfare (MOHW) to further strengthen the protection of health data. HWDC is a data repository that centralizes the NHIRD and

other health-related databases. Cross-linkage of registries is comprehensive and linked to national surveys.[133]

Observational studies: A study in which the investigators do not intervene, but only observe subjects who are (and sometimes who are not, for comparison purposes) exposed to a given factor, and interpret the outcomes. This type of study is more subject to bias than is an experimental study such as a randomized controlled trial.[2]

Orphan drug: A drug used to treat, prevent, or diagnose an orphan disease.[129]

Pharmacovigilance: The pharmacological science relating to the collection, detection, assessment, monitoring, and prevention of adverse effects with pharmaceutical products.[65]

Pragmatic trials: A trial that measures the effects of an intervention in routine clinical practice, to evaluate the intervention's actual effectiveness.[2]

REALISE: A working group comprising the (a) International Advisory Panel (IAP), (b) HTAsiaLink working group, and (c) Core Team. The IAP are prominent experts from leading HTA organizations in Australia, Canada, the UK and the US. The HTAsiaLink working group includes representatives from 11 Asian health systems (Bhutan, China, India, Indonesia, Japan, Malaysia, Philippines, Singapore, South Korea, Taiwan and Thailand). The core team comprises staff from Saw Swee Hock School of Public Health (SSHSPH), National University of Singapore (NUS), and Health Intervention and Technology Assessment Program (HITAP), Ministry of Health, Thailand.

Real-world data: Data collected during the routine delivery of health care. Sources may include observational data, administrative data, research data, patient-generated data or professional-generated data. These data may be collected in administrative datasets, case notes, surveys, product and disease registries, social media, electronic health records, claims and billing datasets, or mobile health applications.[2]

Real-world evidence: Evidence derived from the analysis of real-world data. Real world data are primarily analyzed through observational study designs. Real world evidence is characterized by the actual use of the technology in practice and by findings that are generalizable to the target population for the technology.[2]

Single arm trials: An analysis or evaluation of a study with only one branch, i.e. a trial in which there was no parallel comparison group and all the subjects received the same intervention.[128]

Surrogate endpoint: An indicator that, while not being of direct interest for the patient, may reflect important outcomes. For example, blood pressure is not of direct clinical interest to the patient, but is often used as an evaluation criterion in clinical trials because it is a risk factor for stroke and heart attacks.[2]

US Food and Drug Administration and EU European Medicines Agency: Regulators that are responsible for protecting the public health by ensuring the safety, efficacy, and security of human and veterinary drugs, biological products, and medical devices[107]

Validity: The ability of a measurement or a study to estimate the true value free of systematic errors (bias).[2]

6.2. Abbreviations

AD	Aggregate data
ADE	Adverse drug event
CEA	Cost effectiveness analysis
CUA	Cost utility analysis
CTTI	Clinical Trials Transformation Initiative (US)
EE	Economic evaluation
EMR	Electronic medical records
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
FDA	Food and Drug Administration (US)
HIRA	Health Insurance Review and Assessment Service (Korea)
HITAP	Health Intervention and Technology Assessment Program (Thailand)
HPV	Human papillomavirus
HRQoL	Health-related quality of life
HMIS	Health Management and Information System (Bhutan)
HTA	Health technology assessment
HWDC	Health and Welfare Data Center (Taiwan)
INAHTA	International Network of Agencies for Health Technology Assessment
IPD	Individual patient-level data
IPW	Inverse probability weights
IRB	Institutional Review Board
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
IV	Instrumental variable
KCDC	Korea Centers for Disease Control and Prevention
MDS	Myelodysplastic syndrome
MES	Managed entry scheme
MI	Multiple imputation
MPR	Medication possession ratio
MSAC	Medical Services Advisory Committee (Australia)
NBR	Net-benefit regression
NCSP	National Cervical Cancer Screening Program (Australia)

NHIRD	National Health Insurance Research Database (Taiwan)
NICE	National Institute for Health and Care Excellence (UK)
NKTI	National Kidney and Transplant Institute (Philippines)
NSCLC	Non-small cell lung cancer
NSSO	National Sample Survey Office (India)
OS	Overall survival
OSCCD	Ontario Steering Committee for Cancer Drugs
PBAC	Pharmaceutical Benefits Advisory Committee (Australia)
PCV	Pneumococcal conjugate vaccine
PFS	Progression free survival
PGHD	Person-generated health data
PICO	Patient, Intervention, Comparators and Outcomes (framework)
PRO	Patient reported outcomes
PrCT	Pragmatic clinical trial
PSA	Prostate Specific Antigen
PSM	Propensity score matching
QALY	Quality-adjusted life year
REALISE	REAL World Data In ASia for HEalth Technology Assessment in Reimbursement
RCT	Randomized clinical trial
RWD	Real-world data
RWE	Real-world evidence
SDG	Sustainable Development Goal
SEER	Surveillance, Epidemiology, and End Results database (US)
VOI	Value of information (analysis)
WHO INTDIS	World Health Organization International Drug Information System

7 References

1. Berger M, Monz B, Crown W. Improving transparency in non-interventional research for hypothesis testing — WHY, WHAT, and HOW: considerations from the Real-World Evidence Transparency Initiative 2019. Available from: https://www.ispor.org/docs/default-source/strategic-initiatives/improving-transparency-in-non-interventional-research-for-hypothesis-testing_final.pdf [Accessed 6 Feb 2019].
2. INAHTA, HTAi. HTA glossary. Available from: http://htaglossary.net/HomePage [Accessed 17 Mar 2020].
3. Berger ML, Sox H, Willke RJ, et al. 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. *Value Health* 2017;20(8):1003-08. doi: 10.1016/j.jval.2017.08.3019
4. O'Rourke B, Oortwijn W, Schuller T. The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care* 2020:1-4. doi: 10.1017/S0266462320000215
5. NASEM. National Academies of Sciences, Engineering, Medicine. Examining the impact of real-world evidence on medical product development: proceedings of a workshop series. Washington, DC: The National Academies Press 2019.
6. Chow IHI, Tang CH, You SL, et al. Cost-effectiveness analysis of human papillomavirus DNA testing and Pap smear for cervical cancer screening in a publicly financed health-care system. *Br J Cancer* 2010;103(12):1773-82. doi: 10.1038/sj.bjc.6605974
7. Kim HS, Lee SH, Kim H, et al. Statin-related aminotransferase elevation according to baseline aminotransferases level in real practice in Korea. *J Clin Pharm Ther* 2016;41(3):266-72. doi: 10.1111/jcpt.12377
8. Garrison LP, Jr., Neumann PJ, Erickson P, et al. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report. *Value Health* 2007;10(5):326-35. doi: 10.1111/j.1524-4733.2007.00186.x
9. Motheral BR, Fairman KA. The use of claims databases for outcomes research: rationale, challenges, and strategies. *Clin Ther* 1997;19:346-66.
10. Makady A, de Boer A, Hillege H, et al. What Is real-world data? A review of definitions based on literature and stakeholder interviews. *Value Health* 2017;20(7):858-65. doi: 10.1016/j.jval.2017.03.008
11. Li L, Jick S, Gopalakrishnan C, et al. Metformin use and risk of lactic acidosis in people with diabetes with and without renal impairment: a cohort study in Denmark and the UK. *Diabet Med* 2017;34(4):485-89. doi: 10.1111/dme.13203
12. FDA. FDA's Sentinel Initiative: US Food and Drug Administration. Available from: http://www.fda.gov/safety/fdas-sentinel-initiative [Accessed 3 Mar 2020].
13. FDA. Real-world evidence: US Food and Drug Administration. Available from: https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence [Accessed 22 Jun 2019].
14. Hampson G, Towse A, Dreitlein B, et al. Real world evidence for coverage decisions: opportunities and challenges. A report from the 2017 ICER membership policy summit: Boston Institute for Clinical and Economic Review; 2018. Available from: https://icer-review.org/wp-content/uploads/2018/03/ICER-Real-World-Evidence-White-Paper-03282018.pdf [Accessed 22 Dec 2019].
15. Murphy G. Use of real-world evidence in single drug technology assessment processes by health technology assessment and regulatory organizations Ottawa, ON: CADTH; 2018. Available from: https://www.cadth.ca/use-real-world-evidence-single-drug-technology-assessment-processes-health-technology-assessment-and [Accessed 19 Dec 2019].

16. Fashoyin-Aje LA, Fernandes LL, Lemery S, et al. Asian representation in clinical trials of new drugs for the treatment of cancer. *Journal of Clinical Oncology* 2017;35(15_suppl):6564-64. doi: 10.1200/JCO.2017.35.15_suppl.6564
17. Alemayehu C, Mitchell G, Nikles J. Barriers for conducting clinical trials in developing countries- a systematic review. *International Journal for Equity in Health* 2018;17(1) doi: 10.1186/s12939-018-0748-6
18. Downing NS, Shah ND, Neiman JH, et al. Participation of the elderly, women, and minorities in pivotal trials supporting 2011–2013 U.S. Food and Drug Administration approvals. *Trials* 2016;17(1) doi: 10.1186/s13063-016-1322-4
19. Lee E, Ryan S, Birmingham B, et al. Rosuvastatin pharmacokinetics and pharmacogenetics in white and Asian subjects residing in the same environment. *Clin Pharmacol Ther* 2005;78(4):330-41. doi: 10.1016/j.clpt.2005.06.013
20. NICE. Rheumatoid arthritis in adults: management. Available from: https://www.nice.org.uk/guidance/ng100 [Accessed 20 Jul 2019].
21. Louthrenoo W, Kasitanon N, Katchamart W, et al. 2016 updated Thai Rheumatism Association recommendations for the use of biologic and targeted synthetic disease-modifying anti-rheumatic drugs in patients with rheumatoid arthritis. *Int J Rheu Dis* 2017;20(9):1166-84. doi: 10.1111/1756-185x.13130
22. NICE. Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral compression fractures. Technology appraisal guidance [TA279]: UK National Institute for Health and Care Excellence; 2013. Available from: https://www.nice.org.uk/guidance/ta279 [Accessed 3 Feb 2020].
23. Cheng HM, Chiou LJ, Chen TC, et al. Real-world cost-effectiveness of drug-eluting stents vs. bare-metal stents for coronary heart disease: a five-year follow-up study. *Health Policy* 2019;123(2):229-34. doi: 10.1016/j.healthpol.2018.11.010
24. NHISS. The statistics of national health insurance in Korea: population with health insurance coverage: National Health Insurance Sharing Service. Available from: https://nhiss.nhis.or.kr/bd/ad/bdada013cv.do [Accessed 9 Jun 2019].
25. ACE. Drug evaluation methods and process guide (Version 2): Agency for Care Effectiveness; 2019. Available from: https://www.ace-hta.gov.sg/public-data/our-process-and-methods/ACE%20methods%20and%20process%20guide%20for%20drug%20evaluation%20(20%20Dec%202019).pdf [Accessed 15 Sep 2020].
26. Youngkong S, Baltussen R, Tantivess S, et al. Multicriteria decision analysis for including health interventions in the Universal Health Coverage benefit package in Thailand. *Value Health* 2012;15(6):961-70. doi: 10.1016/j.jval.2012.06.006
27. Rubinstein YR, Groft SC, Bartek R, et al. Creating a global rare disease patient registry linked to a rare diseases biorepository database: Rare Disease-HUB (RD-HUB). *Contemp Clin Trials* 2010;31(5):394-404. doi: 10.1016/j.cct.2010.06.007
28. Ozer-Stillman I, Strand L, Chang J, et al. Meta-analysis for the association between overall survival and progression-free survival in gastrointestinal stromal tumor. *Clin Cancer Res* 2015;21(2):295-302. doi: 10.1158/1078-0432.ccr-14-1779
29. Pazdur R. Endpoints for assessing drug activity in clinical trials. *Oncologist* 2008;13(S2):19-21. doi: 10.1634/theoncologist.13-s2-19
30. Sullivan E. Clinical trial endpoints: US Department of Health and Human Services, Food and Drug Administration (US FDA). Available from: https://www.fda.gov/media/84987/download [Accessed 22 Jun 2019].
31. NICE. Sorafenib for treating advanced hepatocellular carcinoma. Technology appraisal guidance [TA474]: UK National Institute for Health and Care Excellence; 2017. Available from: https://www.nice.org.uk/guidance/ta474 [Accessed 22 Jun 2019].

32. LLSC. Facts and statistics: blood cancers: Leukemia & Lymphoma Society of Canada. Available from: https://www.llscanada.org/disease-information/facts-and-statistics [Accessed 22 Jun 2019].
33. Ramagopalan SV, Simpson A, Sammon C. Can real-world data really replace randomised clinical trials? *BMC Med* 2020;18(1) doi: 10.1186/s12916-019-1481-8
34. Eichler H-G, Abadie E, Breckenridge A, et al. Bridging the efficacy–effectiveness gap: a regulator's perspective on addressing variability of drug response. *Nat Rev Drug Discov* 2011;10(7):495-506. doi: 10.1038/nrd3501
35. Sabirin J, Md Fuzi SA, Shafie AA. Single use dialyser versus reuse dialyser. Putrajaya: Malaysian Health Technology Assessment Section, Ministry of Health, 2014.
36. MSAC. 1276 - National Cervical Screening Program renewal: Australian Government, Department of Health, Medical Services Advisory Committee; 2016. Available from: http://www.msac.gov.au/internet/msac/publishing.nsf/Content/1276-public [Accessed 2 Mar 2020].
37. Ronco G, Dillner J, Elfström KM, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet* 2014;383(9916):524-32. doi: 10.1016/s0140-6736(13)62218-7
38. Kitchener HC, Gilham C, Sargent A, et al. A comparison of HPV DNA testing and liquid based cytology over three rounds of primary cervical screening: extended follow up in the ARTISTIC trial. *Eur J Cancer* 2011;47(6):864-71. doi: 10.1016/j.ejca.2011.01.008
39. Gertig DM, Brotherton JM, Budd AC, et al. Impact of a population-based HPV vaccination program on cervical abnormalities: a data linkage study. *BMC Med* 2013;11(1):227. doi: 10.1186/1741-7015-11-227
40. Tuffaha HW, Scuffham PA. The Australian managed entry scheme: are we getting it right? *Pharmacoeconomics* 2018;36(5):555-65. doi: 10.1007/s40273-018-0633-6
41. Morshed S, Tornetta P, 3rd, Bhandari M. Analysis of observational studies: a guide to understanding statistical methods. *J Bone Joint Surg Am* 2009;91 Suppl 3:50-60. doi: 10.2106/JBJS.H.01577
42. Bowrin K, Briere JB, Levy P, et al. Cost-effectiveness analyses using real-world data: an overview of the literature. *J Med Econ* 2019;22(6):545-53. doi: 10.1080/13696998.2019.1588737
43. Haneuse S. Distinguishing Selection Bias and Confounding Bias in Comparative Effectiveness Research. *Med Care* 2016;54(4):e23-9. doi: 10.1097/MLR.0000000000000011
44. NICE. Lesinurad for treating chronic hyperuricaemia in people with gout. Technology appraisal guidance [TA506]: UK National Institute for Health and Care Excellence; 2018. Available from: https://www.nice.org.uk/guidance/ta506 [Accessed 22 Jun 2019].
45. Lee C, Luo Z, Ngiam KY, et al. Big healthcare data analytics: challenges and applications. In: Khan SU, Zomaya AY, Abbas A, eds. Handbook of large-scale distributed computing in smart healthcare: Springer 2017:11-41.
46. Shafie AA, Tan YP, Liau SY, et al. Registry based analysis of cost-of-illness study among stage C heart failure patients at Hospital Queen Elizabeth II, Sabah, Malaysia. *Health Policy Techn* 2019;8:51-60. doi: 10.1016/j.hlpt.2019.01.002
47. Morton SC, Costlow MR, Graff JS, et al. Standards and guidelines for observational studies: quality is in the eye of the beholder. *J Clin Epidemiol* 2016;71:3-10. doi: 10.1016/j.jclinepi.2015.10.014
48. Schardt C, Adams MB, Owens T, et al. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak* 2007;7(1):16. doi: 10.1186/1472-6947-7-16
49. Pruis S-L, Aziz MIA, Pearce F, et al. Cost-effectiveness analysis of Sunitinib versus Interferon-Alfa for first-line treatment of advanced and/or metastatic renal cell carcinoma in Singapore. *Int J Technol Assess Health Care* 2019;35(2):126-33. doi: 10.1017/s0266462319000059

50. Sander JW. New antiepileptic drugs in practice - how do they perform in the real world? *Acta Neurol Scand* 2005;112(s181):26-29. doi: 10.1111/j.1600-0404.2005.00505.x
51. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 1996;313(7052):275-83. doi: 10.1136/bmj.313.7052.275
52. Makady A, Ham RT, de Boer A, et al. Policies for use of real-world data in Health Technology Assessment (HTA): a comparative study of six HTA agencies. *Value Health* 2017;20(4):520-32. doi: 10.1016/j.jval.2016.12.003
53. Elston J, Taylor RS. Use of surrogate outcomes in cost-effectiveness models: a review of United Kingdom health technology assessment reports. *Int J Technol Assess Health Care* 2009;25(1):6-13. doi: 10.1017/S0266462309090023
54. Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the economic evaluation of health care programme. Third edition: Oxford: Oxford University Press 2005.
55. Wang P, Liu GG, Jo M-W, et al. Valuation of EQ-5D-5L health states: a comparison of seven Asian populations. *Expert Rev Pharmacoecon Outcomes Res* 2019;19(4):445-51. doi: 10.1080/14737167.2019.1557048
56. Lam H. Estimating the EQ-5D-5L value set for the Philippines (final report): Foundation for the Advancement of Clinical Epidemiology Incorporated, 2018.
57. Lin H-W, Li C-I, Lin F-J, et al. Valuation of the EQ-5D-5L in Taiwan. *PLoS One* 2018;13(12):e0209344. doi: 10.1371/journal.pone.0209344
58. Shiroiwa T, Ikeda S, Noto S, et al. Comparison of value set based on DCE and/or TTO data: scoring for EQ-5D-5L health states in Japan. *Value Health* 2016;19(5):648-54. doi: 10.1016/j.jval.2016.03.1834
59. Kim S-H, Ahn J, Ock M, et al. The EQ-5D-5L valuation study in Korea. *Qual Life Res* 2016;25(7):1845-52. doi: 10.1007/s11136-015-1205-2
60. Jensen RE, Potosky AL, Moinpour CM, et al. United States population-based estimates of Patient-Reported Outcomes measurement information system symptom and functional status reference values for individuals with cancer. *J Clin Oncol* 2017;35(17):1913-20. doi: 10.1200/jco.2016.71.4410
61. Teerawattananon Y, Rattanavipapong W, Lin LW, et al. Landscape analysis of health technology assessment (HTA): systems and practices in Asia. *Int J Technol Assess Health Care* 2019;35(6):416-21. doi: 10.1017/s0266462319000667
62. Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res* 2012;21(8):1305-14. doi: 10.1007/s11136-011-0054-x
63. Ueno N, Banno S, Endo Y, et al. Treatment status and safety of crizotinib in 2028 Japanese patients with ALK-positive NSCLC in clinical settings. *Jpn J Clin Oncol* 2019;49(7):676-86. doi: 10.1093/jjco/hyz049
64. Blackhall F, Kim D-W, Besse B, et al. Patient-reported outcomes and quality of life in PROFILE 1007: a randomized trial of crizotinib compared with chemotherapy in previously treated patients with ALK-positive advanced non–small-cell lung cancer. *J Thorac Oncol* 2014;9(11):1625-33. doi: 10.1097/jto.0000000000000318
65. Biswas P. Pharmacovigilance in Asia. *J Pharmacol Pharmacother* 2013;4(Suppl 1):S7-S19. doi: 10.4103/0976-500X.120941
66. Yang Y, Wu F-P, Wang W-J, et al. Real life efficacy and safety of direct-acting antiviral therapy for treatment of patients infected with hepatitis C virus genotypes 1, 2 and 3 in northwest China. *World J Gastroenterol* 2019;25(44):6551-60. doi: 10.3748/wjg.v25.i44.6551
67. Lindquist M. VigiBase, the WHO Global ICSR Database System: basic facts. *Drug Inf J* 2008;42(5):409-19. doi: 10.1177/009286150804200501
68. Huang YL, Moon J, Segal JB. A comparison of active adverse event surveillance systems worldwide. *Drug Saf* 2014;37(8):581-96. doi: 10.1007/s40264-014-0194-3

69. CIOMS. Final Report of CIOMS Working Group VIII: practical aspects of signal detection in pharmacovigilance. Geneva: CIOMS; 2010. Available from: https://cioms.ch/working_groups/working-group-viii [Accessed 9 Jan 2020].
70. Sculpher M, Pang F, Manca A, et al. Generalisability in economic evaluation studies in healthcare: a review and case studies. *Health Technol Assess* 2004;8(49) doi: 10.3310/hta8490
71. Loong HH, Wong CKH, Leung LKS, et al. Cost-effectiveness analysis of pembrolizumab compared to standard of care as first line treatment for patients with advanced melanoma in Hong Kong. *Cost Eff Resour Alloc* 2020;18(1) doi: 10.1186/s12962-020-0200-9
72. Schachter J, Ribas A, Long GV, et al. Pembrolizumab versus ipilimumab for advanced melanoma: final overall survival results of a multicentre, randomised, open-label phase 3 study (KEYNOTE-006). *Lancet* 2017;390(10105):1853-62. doi: 10.1016/s0140-6736(17)31601-x
73. Blommestein HM, Franken MG, Uyl-de Groot CA. A practical guide for using registry data to inform decisions about the cost effectiveness of new cancer drugs: lessons learned from the PHAROS registry. *Pharmacoeconomics* 2015;33(6):551-60. doi: 10.1007/s40273-015-0260-4
74. Minelli C, Baio G. Value of information: a tool to improve research prioritization and reduce waste. *PLoS Med* 2015;12(9):e1001882. doi: 10.1371/journal.pmed.1001882
75. Shafie AA, Ahmad N, Naidoo J, et al. Estimating the population health and economic impacts of introducing a pneumococcal conjugate vaccine in Malaysia- an economic evaluation. *Hum Vaccin Immunother* 2020:1-9. doi: 10.1080/21645515.2019.1701911
76. Arnet I, Kooij MJ, Messerli M, et al. Proposal of standardization to assess adherence with medication records: methodology matters. *Ann Pharmacother* 2016;50(5):360-8. doi: 10.1177/1060028016634106
77. Lam WY, Fresco P. Medication adherence measures: an overview. *Biomed Res Int* 2015;2015:217047. doi: 10.1155/2015/217047
78. Tu JV, Willison DJ, Silver FL, et al. Impracticability of informed consent in the registry of the Canadian Stroke Network. *N Engl J Med* 2004;350(14):1414-21. doi: 10.1056/nejmsa031697
79. Voon HY, Shafie AA, Bujang MA, et al. Cost effectiveness analysis of carbetocin during cesarean section in a high volume maternity unit. *Journal of Obstetrics and Gynaecology Research* 2018;44(1):109-16. doi: 10.1111/jog.13486
80. Gliklich RE, Dreyer NA, Leavy MB. Registries for evaluating patient outcomes: a user's guide (3rd edition) Rockville (MD): Agency for Healthcare Research and Quality; 2014. Available from: https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/registries-guide-3rd-edition_research.pdf [Accessed 23 Dec 2019].
81. Tan J, Peng X, Shu X, et al. Technical guidance for developing patient registry databases. *Chin J Evid Based Med* 2019;19(7):771-8. doi: 10.7507/1672-2531.201904161
82. SSO. Singapore Statutes Online. National Registry of Diseases Act (Chapter 201B) Available from: https://sso.agc.gov.sg/Act/NRDA2007 [Accessed 23 Dec 2019].
83. Ferver K, Burton B, Jesilow P. The use of claims data in healthcare research. *Open Public Health J* 2009;2(1):11-24. doi: 10.2174/1874944500902010011
84. Suwantika AA, Zakiyah N, Kusuma ASW, et al. Impact of switch options on the economics of Pneumococcal Conjugate Vaccine (PCV) introduction in Indonesia. *Vaccines* 2020;8(2):233. doi: 10.3390/vaccines8020233
85. Shin S, Kim YH, Hwang JS, et al. Economic evaluation of prostate cancer screening test as a national cancer screening program in South Korea. *Asian Pac J Cancer Prev* 2014;15(8):3383-9. doi: 10.7314/apjcp.2014.15.8.3383
86. Birnbaum H, Cremieux P, Greenberg P, et al. Using healthcare claims data for outcomes research and pharmacoeconomic analyses. *Pharmacoeconomics* 1999;16(1):1-8.
87. Pempa, Luz ACG, Luangasanatip N, et al. Economic evaluation of rotavirus vaccination in children of Bhutan. *Vaccine* 2020;38(32):5049-59. doi: 10.1016/j.vaccine.2020.05.035
88. FDA. US FDA Guidance document: use of electronic health record data in clinical investigations guidance for industry 2018. Available from: https://www.fda.gov/regulatory-

information/search-fda-guidance-documents/use-electronic-health-record-data-clinical-investigations-guidance-industry [Accessed 26 Dec 2019].
89. Mahendraratnam N. Determining real world data's fitness for use and the role of reliability 2019. Available from: https://healthpolicy.duke.edu/publications/determining-real-world-data%E2%80%99s-fitness-use-and-role-reliability [Accessed 20 Dec 2019].
90. Wang W, Gao P, Wu J, et al. Technical guidance for developing research databases using existing health and medical data. *Chin J Evid Based Med* 2019;19(7):763-70. doi: 10.7507/1672-2531.201904160
91. NSSO. Key indicators of social consumption in India: Health (NSSO 71st round, January-June 2014). New Delhi: National Sample Survey Office (National Statistical Office, Government of India); 2015. Available from: https://ruralindiaonline.org/library/resource/health-in-india-nss-71st-round-january-june-2014/ [Accessed 22 Feb 2020].
92. Tantivess S, Yothasamut J, Saengsri W. Utilisation of evidence from Thailand's National Health Examination Survey in policy development: finding the weakest link. *Health Res Policy Syst* 2019;17(1):104. doi: 10.1186/s12961-019-0512-4
93. Hunger M, Schwarzkopf L, Heier M, et al. Official statistics and claims data records indicate non-response and recall bias within survey-based estimates of health care utilization in the older population. *BMC Health Serv Res* 2013;13(1):1. doi: 10.1186/1472-6963-13-1
94. Chugh Y, Dhiman RK, Premkumar M, et al. Real-world cost-effectiveness of pan-genotypic Sofosbuvir-Velpatasvir combination versus genotype dependent directly acting anti-viral drugs for treatment of hepatitis C patients in the universal coverage scheme of Punjab state in India. *PLoS One* 2019;14(8):e0221769. doi: 10.1371/journal.pone.0221769
95. Eurostat. Guidelines for the development and criteria for the adoption of health survey instruments Luxembourg: Office for Official Publications of the European Communities; 2005. Available from: https://ec.europa.eu/eurostat/documents/3859598/5890585/KS-CC-05-003-EN.PDF/240dbb94-25be-4d07-aa73-601552886f84 [Accessed 22 Feb 2020].
96. Kelley K, Clark B, Brown V, et al. Good practice in the conduct and reporting of survey research. *Int J Qual Health Care* 2003;15:261-6.
97. Radin JM, Wineinger NE, Topol EJ, et al. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. *Lancet Digit Health* 2020;2(2):e85-e93. doi: 10.1016/s2589-7500(19)30222-5
98. Oliver N, Lepri B, Sterly H, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Sci Adv* 2020;6(23):eabc0764. doi: 10.1126/sciadv.abc0764
99. Han H, Ma X, Oyama K. Flexible detection of fall events using bidirectional EMG sensor. *Stud Health Technol Inform* 2017;245:1225.
100. CTTI. Clinical Trials Transformation Initiative (CTTI) recommendations: advancing the use of mobile technologies for data capture and improved clinical trials 2017. Available from: www.ctti-clinicaltrials.org/files/mobile-devices-recommendations.pdf [Accessed 4 Feb 2020].
101. Byrom B, Watson C, Doll H, et al. Selection of and evidentiary considerations for wearable devices and their measurements for use in regulatory decision making: Recommendations from the ePRO Consortium. *Value Health* 2018;21(6):631-39. doi: 10.1016/j.jval.2017.09.012
102. Sauerbrei W, Abrahamowicz M, Altman DG, et al. STRengthening analytical thinking for observational studies: the STRATOS initiative. *Stat Med* 2014;33(30):5413-32. doi: 10.1002/sim.6265
103. AHRQ. Developing a protocol for observational comparative effectiveness research: a user's guide Rockville, MD: Agency for Healthcare Research and Quality; 2013. Available from: http://www.ncbi.nlm.nih.gov/books/NBK126190/ [Accessed 6 Feb 2020].
104. Berger ML, Martin BC, Husereau D, et al. A questionnaire to assess the relevance and credibility of observational studies to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health* 2014;17(2):143-56. doi: 10.1016/j.jval.2013.12.011

105. ENCePP. Checklist for study protocols (revision 2): The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. Available from: http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml [Accessed 6 Feb 2020].
106. ENCePP. Guide on methodological standards in pharmacoepidemiology (revision 4). EMA/95098/2010. Available from: http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml.
107. FDA. US FDA Guidance for industry: good pharmacovigilance practices and pharmacoepidemiologic assessment (pharmacovigilance guidance): Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER); 2005. Available from: https://www.fda.gov/media/71546/download [Accessed 6 Feb 2020].
108. Dreyer NA, Schneeweiss S, McNeil BJ, et al. GRACE principles: recognizing high-quality observational studies of comparative effectiveness. *Am J Manag Care* 2010;16:467e71.
109. Dreyer NA, Velentgas P, Westrich K, et al. The GRACE checklist for rating the quality of observational studies of comparative effectiveness: a tale of hope and caution. *J Manage Care Pharm* 2014;20(3):301-08. doi: 10.18553/jmcp.2014.20.3.301
110. Berger ML, Mamdani M, Atkins D, et al. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force. *Value Health* 2009;12(8):1044-52. doi: 10.1111/j.1524-4733.2009.00600.x
111. Cox E, Martin BC, Van Staa T, et al. Good research practices for comparative effectiveness research: Approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: The International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force report—part II. *Value Health* 2009;12(8):1053-61. doi: 10.1111/j.1524-4733.2009.00601.x
112. Johnson ML, Crown W, Martin BC, et al. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force report -- part III. *Value in Health* 2009;12(8):1062-73. doi: 10.1111/j.1524-4733.2009.00602.x
113. Motheral B, Brooks J, Clark MA, et al. A checklist for retrospective database studies—report of the ISPOR task force on retrospective databases. *Value Health* 2003;6:90-7.
114. Berger ML, Dreyer N, Anderson F, et al. Prospective observational studies to assess comparative effectiveness: the ISPOR Good Research Practices Task Force report. *Value Health* 2012;15(2):217-30. doi: 10.1016/j.jval.2011.12.010
115. PCORI. The PCORI methodology report 2013: PCORI (Patient-Centered Outcomes Research Institute) Methodology Committee; 2013. Available from: http://www.pcori.org/assets/2013/11/PCORI-Methodology-Report.pdf [Accessed 7 Feb 2020].
116. Peng X, Shu X, Tan J, et al. Technical guidance for designing observational studies to assess therapeutic outcomes using real-world data. *Chin J Evid Based Med* 2019;19(7):779-86. doi: 10.7507/1672-2531.201904164
117. Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61(4):344-49. doi: 10.1016/j.jclinepi.2007.11.008
118. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med* 2015;12(10):e1001885. doi: 10.1371/journal.pmed.1001885

119. Qiu Y, Xiao Y, Zhao K, et al. Health technology assessment on the Da Vinci Surgical System using real world data in China. *Int J Technol Assess Health Care*;33:41.
120. Oude Rengerink K, Kalkman S, Collier S, et al. Series: pragmatic trials and real world evidence: paper 3. patient selection challenges and consequences. *J Clin Epidemiol* 2017;89:173-80. doi: 10.1016/j.jclinepi.2016.12.021
121. Worsley SD, Oude Rengerink K, Irving E, et al. Series: pragmatic trials and real world evidence: paper 2. setting, sites, and investigator selection. *J Clin Epidemiol* 2017;88:14-20. doi: 10.1016/j.jclinepi.2017.05.003
122. Gamerman V, Cai T, Elsäßer A. Pragmatic randomized clinical trials: best practices and statistical guidance. *Health Serv Outcomes Res Methodol* 2018;19(1):23-35. doi: 10.1007/s10742-018-0192-5
123. GetReal. Publications: pragmatic trial paper series. Available from: https://www.pragmagic.eu/JCEPragmaticTrialsSeries [Accessed 7 Feb 2020].
124. Loudon K, Treweek S, Sullivan F, et al. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147. doi: 10.1136/bmj.h2147
125. Wen Z, Li L, Liu Y, et al. Technical guidance for pragmatic randomized controlled trials. *Chin J Evid Based Med* 2019;19(7):794-802. doi: 10.7507/1672-2531.201904163
126. Ford I, Norrie J. Pragmatic trials. *N Engl J Med* 2016;375(5):454-63. doi: 10.1056/NEJMra1510059
127. FDA. FDA-AACR Real world evidence workshop (presentation slides). MD: US FDA, American Association for Cancer Research, 2019.
128. Carrigan G, Whipple S, Capra WB, et al. Using electronic health records to derive control arms for early phase single-arm lung cancer trials: Proof-of-concept in randomized controlled trials. *Clin Pharmacol Ther* 2020;107(2):369-77. doi: 10.1002/cpt.1586
129. Doring JH, Lampert A, Hoffmann GF, et al. Thirty years of orphan drug legislation and the development of drugs to treat rare seizure conditions: a cross sectional analysis. *PLoS One* 2016;11(8):e0161660. doi: 10.1371/journal.pone.0161660
130. Schmidli H, Haring DA, Thomas M, et al. Beyond randomized clinical trials: use of external controls. *Clin Pharmacol Ther* 2020;107(4):806-16. doi: 10.1002/cpt.1723
131. Katz R. FDA update. *Epilepsy Res* 2006;68(1):85-94. doi: 10.1016/j.eplepsyres.2005.09.034
132. Hsing AW, Ioannidis JP. Nationwide population science: lessons from the Taiwan National Health Insurance Research Database. *JAMA Intern Med* 2015;175(9):1527-9. doi: 10.1001/jamainternmed.2015.3540
133. Lin LY, Warren-Gash C, Smeeth L, et al. Data resource profile: the National Health Insurance Research Database (NHIRD). *Epidemiol Health* 2018;40:e2018062. doi: 10.4178/epih.e2018062
134. HTAi. Background paper for HTAi Asia Policy Forum 2018: improving access to high cost technologies in the Asia Pacific region Jakarta, Indonesia2018. Available from: https://htai.org/wp-content/uploads/2019/06/HTAi_asia-policy-forum_2018-background-paper_final.pdf [Accessed 9 Apr 2020].
135. NCI. National Cancer Institute (NCI) Metathesaurus. Available from: https://ncimeta.nci.nih.gov/ncimbrowser/ [Accessed 6 Feb 2020].
136. Miller RS, Wong JL. Using oncology real-world evidence for quality improvement and discovery: the case for ASCO's CancerLinQ. *Future Oncol* 2018;14(1):5-8. doi: 10.2217/fon-2017-0521
137. Teerawattananon Y, Mugford M, Tangcharoensathien V. Economic evaluation of palliative management versus peritoneal dialysis and hemodialysis for end-stage renal disease: evidence for coverage decisions in Thailand. 2007;10(1):61-72. doi: 10.1111/j.1524-4733.2006.00145.x
138. Berger ML, Daniel G, Frank K. A framework for regulatory use of real-world evidence 2017. Available from: https://healthpolicy.duke.edu/publications/real-world-evidence-white-paper [Accessed 20 Dec 2019].

139. Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology. *J Natl Cancer Inst* 2017;109(11) doi: 10.1093/jnci/djx187
140. KLRI. Personal information protection act (English version): Korea Legislative Research Institute. Available from: http://www.law.go.kr/lsInfoP.do?lsiSeq=142563&chrClsCd=010203&urlMode=engLsInfoR&viewCls=engLsInfoR#0000 [Accessed 20 Aug 2020].
141. CMS. Centers for Medicare & Medicaid Services (CMS) 2008-2010 Data Entrpreneurs' Synthetic Public Use File (DE-SynPUF). Available from: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/SynPUFs/DE_Syn_PUF [Accessed 3 Feb 2020].
142. Baltussen R, Leidl R, Ament A. Real world designs in economic evaluation: bridging the gap between clinical research and policy-making. *Pharmacoeconomics* 1999;16(5)
143. Bell H, Wailoo A, Hernandez M, et al. The use of real-world data for the estimation of treatment effects in NICE decision making. *Report by the Decision Support Unit* 2016
144. Pongchaiyakul C, Nanagara R, Songpatanasilp T, et al. Cost-effectiveness of denosumab for high-risk postmenopausal women with osteoporosis in Thailand. *J Med Econ* 2020:1. doi: 10.1080/13696998.2020.1730381
145. Pruis SL, Aziz MIA, Pearce F, et al. Cost-effectiveness analysis of sunitinib versus interferon-alfa for first-line treatment of advanced and/or metastatic renal cell carcinoma in Singapore. *Int J Technol Assess Health Care* 2019;35(2):126-33. doi: 10.1017/S0266462319000059
146. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002;359(9302):248-52. doi: 10.1016/s0140-6736(02)07451-2
147. Gurwitz J, Sykora K, Mamdani M, et al. Reader's guide to critical appraisal of cohort studies: 1. Role and design. *The BMJ* 2005;330
148. Corrao G. Building reliable evidence from realworld data: methods, cautiousness and recommendations. 2013;10(3) doi: 10.2427/8981
149. Black DA, Hsu Y-C, Sanders SG, et al. The Methuselah Effect: the pernicious impact of unreported deaths on old-age mortality estimates. *Demography* 2017;54(6):2001-24. doi: 10.1007/s13524-017-0623-x
150. Bayani D, Uy G, Almirol B, et al. Economic evaluation of renal replacement coverage policies in the Philippines. unpublished, 2019.
151. Norgaard M, Ehrenstein V, Vandenbroucke JP. Confounding in observational studies based on large health care databases: problems and potential solutions - a primer for the clinician. *Clin Epidemiol* 2017;9:185-93. doi: 10.2147/CLEP.S129879
152. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Epidemiology* 2007;18(6):805-35. doi: 10.1097/EDE.0b013e3181577511
153. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;370(9596):1453-57. doi: 10.1016/S0140-6736(07)61602-X
154. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLOS Medicine* 2015;12(10):e1001885. doi: 10.1371/journal.pmed.1001885
155. Faria R, Alava M, Manca A, et al. NICE Decision Support Unit Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: Methods for comparative individual patient data 2015. [Accessed January 1 2020].
156. Chan K, Nam S, Evans B, et al. Developing a framework to incorporate real-world evidence in cancer drug funding decisions: the Canadian Real-world Evidence for Value of Cancer Drugs (CanREValue) collaboration. *BMJ Open* 2020;10(1):e032884. doi: 10.1136/bmjopen-2019-032884
157. Council MR. Using natural experiments to evaluate population health interventions: guidance for producers and users of evidence. [Accessed February 10 2020].

158. McCaffrey DF, Griffin BA, Almirall D, et al. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Stat Med* 2013;32(19):3388-414. doi: 10.1002/sim.5753
159. Lin LW, Teng GG, Lim AYN, et al. Cost-effectiveness of an adherence-enhancing intervention for gout based on real-world data. *Int J Rheum Dis* 2019;22(4):545-54. doi: 10.1111/1756-185X.13446
160. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med* 2015;34(28):3661-79. doi: 10.1002/sim.6607
161. Faria R, Alava M, Manca A, et al. NICE Desicion Support Unit Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: Methods for comparative individual patient data. *Report by the Decision Support Unit* 2015
162. Kamo T, Momosaki R, Suzuki K, et al. Effectiveness of Intensive Rehabilitation Therapy on Functional Outcomes After Stroke: A Propensity Score Analysis Based on Japan Rehabilitation Database. *Journal of Stroke and Cerebrovascular Diseases* 2019;28(9):2537-42. doi: https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.06.007
163. Khor Sea. Real world costs and cost-effectiveness of rituximab for diffuse large B-cell lymphoma patients: a population-based analysis. *BMC Cancer* 2014;14(586) doi: http://www.biomedcentral.com/1471-2407/14/586
164. Normand S, Sykora K, Li P, et al. Readers guide to critical appraisal of cohort studies: 3. Analytical strategies to reduce confounding. *The BMJ* 2005;330
165. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc* 1996;91(434):444-55. doi: 10.2307/2291629
166. Ertefaie A, Small DS, Flory JH, et al. A tutorial on the use of instrumental variables in pharmacoepidemiology. *Pharmacoepidemiol Drug Saf* 2017;26(4):357-67. doi: 10.1002/pds.4158
167. Winn AN, Shah GL, Cohen JT, et al. The real world effectiveness of hematopoietic transplant among elderly individuals with multiple myeloma. *J Natl Cancer Inst* 2015;107(8) doi: 10.1093/jnci/djv139
168. Manca A, Palmer S. Handling missing data in patient-level cost-effectiveness analysis alongside randomised clinical trials. *Appl Health Econ Health Policy* 2005;4(2)
169. Manca A, Palmer S. Handling Missing Data in Patient-Level Cost-Effectiveness Analysis alongside Randomised Clinical Trials. *Applied Health Economincs and Health Policy* 2005;4(2)
170. Thein HH, Isaranuwatchai W, Qiao Y, et al. Cost-effectiveness analysis of potentially curative and combination treatments for hepatocellular carcinoma with person-level data in a Canadian setting. *Cancer Med* 2017;6(9):2017-33. doi: 10.1002/cam4.1119
171. Jones MP. Indicator and stratification methods for missing explanatory variables in multiple linear regression. *J Am Stat Assoc* 1996;91(433):222-30. doi: 10.2307/2291399
172. Mason AJ, Gomes M, Grieve R, et al. A Bayesian framework for health economic evaluation in studies with missing data. *Health Econ* 2018;27(11):1670-83. doi: 10.1002/hec.3793
173. Faria R, Gomes M, Epstein D, et al. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *PharmacoEconomics* 2014;32(12):1157-70. doi: 10.1007/s40273-014-0193-3
174. Stinnett A, Mullahy J. Net Health Benefits: A New Framework for the Analysis of Uncertainty in Cost-Effectiveness Analysis. *Med Decision Making* 1998
175. Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Econ* 2002;11(5):415-30. doi: 10.1002/hec.678
176. Hoch JS, Rockx MA, Krahn AD. Using the net benefit regression framework to construct cost-effectiveness acceptability curves: an example using data from a trial of external loop

recorders versus Holter monitoring for ambulatory monitoring of "community acquired" syncope. *BMC Health Serv Res* 2006;6:68. doi: 10.1186/1472-6963-6-68
177. Liao CH, Tan EC, Chen CC, et al. Real-world cost-effectiveness of laparoscopy versus open colectomy for colon cancer: a nationwide population-based study. *Surg Endosc* 2017;31(4):1796-805. doi: 10.1007/s00464-016-5176-3
178. Platt RW, Hutcheon JA, Suissa S. Immortal time bias in epidemiology. *Curr Epidemiol Rep* 2019;6(1):23-27. doi: 10.1007/s40471-019-0180-5
179. Andersen PK, Geskus RB, de Witte T, et al. Competing risks in epidemiology: possibilities and pitfalls. *Int J Epidemiol* 2012;41(3):861-70. doi: 10.1093/ije/dyr213
180. Mansournia MA, Etminan M, Danaei G, et al. Handling time varying confounding in observational research. *BMJ* 2017;359:j4587. doi: 10.1136/bmj.j4587
181. Cain KC, Harlow SD, Little RJ, et al. Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. *Am J Epidemiol* 2011;173(9):1078-84. doi: 10.1093/aje/kwq481
182. Kreif N, Grieve R, Radice R, et al. Regression-adjusted matching and double-robust methods for estimating average treatment effects in health economic evaluation. *Health Services and Outcomes Research Methodology* 2013;13(2):174-202. doi: 10.1007/s10742-013-0109-2
183. Abadie A, Imbens GW. Bias-Corrected Matching Estimators for Average Treatment Effects. *Journal of Business & Economic Statistics* 2011;29(1):1-11. doi: 10.1198/jbes.2009.07333
184. van der Laan MJ, Gruber S. Collaborative double robust targeted maximum likelihood estimation. *Int J Biostat* 2010;6(1):17-17. doi: 10.2202/1557-4679.1181
185. Kang JDY, Schafer JL. Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data. *Statist Sci* 2007;22(4):523-39. doi: 10.1214/07-STS227
186. Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. *Int J Epidemiol* 2017;46(2):756-62. doi: 10.1093/ije/dyw323
187. Austin PC, Wu CF, Lee DS, et al. Comparing the high-dimensional propensity score for use with administrative data with propensity scores derived from high-quality clinical data. *Stat Methods Med Res* 2020;29(2):568-88. doi: 10.1177/0962280219842362
188. Sutradhar R, Cook RJ. Analysis of interval-censored data from clustered multistate processes: application to joint damage in psoriatic arthritis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2008;57(5):553-66. doi: 10.1111/j.1467-9876.2008.00630.x
189. Sutradhar R, Barbera L, Seow H, et al. Multistate Analysis of Interval-Censored Longitudinal Data: Application to a Cohort Study on Performance Status Among Patients Diagnosed With Cancer. *American Journal of Epidemiology* 2010;173(4):468-75. doi: 10.1093/aje/kwq384
190. Lin DY. Linear regression analysis of censored medical costs. *Biostatistics* 2000;1(1):35-47. doi: 10.1093/biostatistics/1.1.35
191. Bang H, Tsiatis A. Estimating medical costs with censored data. *Biometrika* 2000;87(2):329-43. doi: 10.1093/biomet/87.2.329
192. Glick HA. Sample size and power for cost-effectiveness analysis (Part 1). *Pharmacoeconomics* 2011;29(3):189-98. doi: 10.2165/11585070-000000000-00000
193. Glick HA. Sample size and power for cost-effectiveness analysis (Part 2). *Pharmacoeconomics* 2011;29(4):287-96. doi: 10.2165/11585080-000000000-00000
194. Milea D, Azmi S, Reginald P, et al. A review of accessibility of administrative healthcare databases in the Asia-Pacific region. *J Mark Access Health Policy* 2015;3 doi: 10.3402/jmahp.v3.28076

8 Appendix

8.1. List of real-world data sources available in Asia (non-exhaustive)

Table 8.1. Adapted from: Milea et al.[194] RWD sources supplemented by the REALISE guidance authors.

Country / RWD source	RWD source
Bhutan	
• Bhutan cancer Registry	Registry
• DHIS-2 (District Health Information System)	EMR
• Mother and Child Health tracking system	EMR
• Lab Information system	EMR
• e-BMSIS (Electronic Bhutan Medical Supplies Inventory System)	EMR
China	
• Zhongnan Hospital	EMR
• West China Hospital	EMR
• Jinhua Municipal Central Hospital database	EMR
• The 306th Hospital of PLA, Beijing	EMR
• Chinese People's liberation army general hospital database	EMR
• Soochow University Affiliates Children's Hospital database	EMR
• Urban Employee Basic Medical Insurance database (UEBMI) (Tianjin)	Reimbursement
• UEBMI (Guangzhou)	Reimbursement
• UEBMI (Hebei)	Reimbursement
• Urban Resident Basic Medical Insurance database (URBMI) (Guangzhou)	Reimbursement
• China Health Insurance Research Association (CHIRA) database	Reimbursement
• Zhongshan Hospital	EMR
• New Rural Cooperative Medical Scheme (NRCMS)	Reimbursement
• Electronic Health Records (EHR) system - Minhang, Shanghai	EMR
• Guangzhao Psychiatric Hospital	EMR
India	
• National Health System Cost Database	Reimbursement
Japan	
• Convergence CT Global Research Network (CGRN)	EMR
• Medical Data Vision (MDV) EBM Provider	Hospital administration

• Hamamatsu Medical University Database	EMR
• Osaka University Database Computer Operating System Database Dokkyo Medical University	EMR
• EMR Retrieval System-Kyoto University Hospital	EMR
• Nihon University School of Medicine Clinical Data Warehouse	EMR
• Japan Medical Data Centre (JMDC) Claims database	Reimbursement
• JammNet	Reimbursement
• Diagnosis Procedure Code Database	EMR
• NHI Database	Reimbursement
• Medi-Trend	Prescription
• Nihon-Chouzai Pharmacy Claim Database	Prescription
• JMIRI Pharmacy Claims Database	Prescription
• Long term care insurance (Kaigo Hoken)	Reimbursement
Malaysia	
• United Nations University (UNU)-Casemix database	Hospital administration (discharge records)
• Electronic Health Management Information System	Hospital administration
• Acute Stroke Registry Malaysia	Registry
• Malaysian Thalassaemia Registry	Registry
• National Dermatology Registry	Registry
• Malaysian Gastrointestinal Registry	Registry
• Malaysian National Neonatal Registry	Registry
• National Cardiovascular Disease Database	Registry
• National Cancer Patient Registry	Registry
• National Cardiovascular and thoracic surgical database	Registry
• National Eye Database	Registry
• National Inflammatory Arthritis Registry	Registry
• National Neurology Registry	Registry
• National Obstetrics Registry (NOR) Malaysia	Registry
• National Renal Registry	Registry
• Malaysian Registry of Renal Biopsy	Registry
• National Transplant Registry	Registry
• National Trauma Database	Registry
• National Orthopaedic Registry Malaysia	Registry
• National Suicide Registry Malaysia	Registry
• Malaysian Registry of Intensive Care	Registry

Philippines	
• Philippine Health Insurance Corporation (PhilHealth) Claims Database	Reimbursement
• Philippine Health Insurance Corporation (PhilHealth) Cost Database for Z-Benefits	Reimbursement
• Philippine Integrated Disease Surveillance and Response (PIDSR)	Registry
• Philippine Renal Disease Registry (PRDR)	Registry
• Field Health Service Information System	Database
• National Injury Surveillance System	Database
Singapore	
• Casemix database	Hospital
• Medisave database	Reimbursement
• National Electronics Health Records Database	EMR
• National Immunization Registry	Registry
• Various condition-specific registries including Singapore Acute Myocardial Infarction Registry, Singapore Cancer Registry, Singapore Diabetes Registry, Singapore Stroke Registry, Singapore Renal Registry, etc.	Registry
South Korea	
• National Health Insurance Corporation (NHIC) database	Reimbursement
• Health Insurance Review and Assessment (HIRA) database	Reimbursement
• NCI Central Cancer Registry	Registry
• National Institute of Health Clinical Research Information System (CRIS)	Registry (only publicly funded research registries)
• All hospitals have EMR systems compliant to HL7 and connected with PACS and OCS for various research	EMR (with image data)
Taiwan	
• National Health Insurance (NHI) Database	Reimbursement
• NHI DAA-treated patients registry	Registry/Reimbursement
• NHI immune oncology drugs treated patient registry	Registry/Reimbursement
• Taiwan Cancer Registry	Registry
• Cancer screening registries	Registry
• Adult preventive health information file	Registry
• Rare disease data	Registry
• Data of genetic disease	Registry
• Notifiable disease dataset of confirmed cases	Registry
• Symptom Surveillance and Reporting System Database	Registry

• Infectious diseases database (tuberculosis, HIV/AIDS)	Registry
• Database of National Immunization Information System	Registry
• Birth certificate application	Registry
• Survey for three-hypers series	Health survey
• National Health Interview Survey	Health survey
• Chang Gung Research Database (CGRD)	EMR
• China Medical University Hospital Clinical Research Data Repository (CMUH-CRDR)	EMR
• National Taiwan University Hospital Integrated Health Care Information System (NTUH-IHIS)	EMR
• Taipei Medical University Healthcare System Clinical Data	EMR
• Taipei Veterans General Hospital Big Data Center (Taipei VGH BDC)	EMR
Thailand	
• Universal Coverage Scheme	Reimbursement
• Civil Servant Medical Benefits Scheme	Reimbursement
• Social Security Scheme	Reimbursement
• Ramathibodi Hospital Database	EMR
• Buddhachinaraj Hospital Database	EMR
• Sunpasitthiprasong Hospital	EMR
• Nakhon Thai Crown Prince Hospital	EMR