



The essence of real-world study in cancer research

Chen Li^{1#}, Jian Zang^{2#}, Fan Li¹, Haixia Hu¹, Ling Wang¹, Jielai Xia¹

¹Department of Health Statistics, School of Preventive Medicine, Fourth Military Medical University, Xi'an, China; ²Department of Radiation Oncology, Xijing Hospital, Fourth Military Medical University, Xi'an, China

Contributions: (I) Conception and design: C Li, J Xia; (II) Administrative support: L Wang; (III) Provision of study materials or patients: F Li, H Hu; (IV) Collection and assembly of data: L Wang, J Zang; (V) Data analysis and interpretation: J Zang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Ling Wang; Jielai Xia. 169 Changlexilu Rd., Xi'an 710032, China. Email: wl.medstat@qq.com; jielaixia@yahoo.com.

Abstract: The widespread availability real-world data (RWD) offers valuable insights into cancer treatment, disease management, disease burden, and socio-economic status in routine practice, and thus has the potential to close gaps between the existed evidence and real-world populations. It can enrich and justify the health regulatory and intervention assessment to consider evidence from RWD in a considered and balanced manner. However, even with advanced methodologies to control various biases, many challenges and misunderstandings require resolutions before RWD can be integrated into decision-making. This paper provides an outline of the enormous value of real-world study (RWS) in cancer research, the framework for the use of RWD and the potential pitfalls inherent to using RWD for the comparative effectiveness. RWSs and randomized controlled trials (RCTs) remain complementary forms for evidence in medical researches each other. RWS can be informative, but should be interpreted cautiously. It is not appropriate to adopt new therapies on the basis of RWS in isolation, particularly if there is no evidence of treatment benefit from RCTs. In this situation, claims of effectiveness may probably be misleading and artificial, and thus researchers should be critical of the plausibility of such outcomes. However, systematically incorporating real-world evidence (RWE) into the benefit-risk assessment will ultimately improve patient care and better inform health policy decisions.

Keywords: Real-world data (RWD); real-world study (RWS); real-world evidence (RWE); bias

Received: 28 May 2020; Accepted: 10 July 2020; Published: 30 July 2020.

doi: 10.21037/tbcr-20-27

View this article at: <http://dx.doi.org/10.21037/tbcr-20-27>

Background

Despite the rapid development of gene engineering and other novel therapies, advanced cancer is still a largely intractable disease, and its high mortality necessitates the development of better treatment (1). In general, the regulatory approval of new cancer treatment is built on the objective criteria of efficacy and safety derived from randomized controlled trials (RCTs). However, with advances in genetic and molecular treatment, the number of patients available for a specific clinical trial may be too small for proper assessment of a benefit-risk paradigm (2). As a result, drug developers and regulators have become

increasingly interested in data external from clinical trials, and there is growing need to rethink the traditional approaches to treatment development and approval (3). One potential solution to this may be to further extend and validate the results from RCTs with the wider experience of real-world data (RWD) or real-world evidence (RWE).

Compared to the rigorous admission criteria and strict research conditions of RCTs, RWD is collected from the real world to provide evidence for clinical and policy-relevant questions that cannot be addressed with data from RCTs, such as patient health status, cancer incidence and mortality, quality of and access to care delivered in routine practice, rare cancers, and rare adverse events and toxicities

in the general population (4). RWD has drawn more attention from government departments, medical health practitioners, and other stakeholders, which have shown an increasingly important role in the benefit-risk assessment of cancer treatment in regulatory decisions and medical management. A search for the terms “real-world evidence” and “real-world data” in PubMed has seen an approximate 1,000% increase in citations between 2002 and 2019 (from 2,435 citations to 24,239 citations).

There are at least three key incentives to applying RWD in the development of treatment in cancer research (5). First, it is critical to have a good standard of comparison when translating a trial breakthrough into meaningful therapeutic regimens. RWD may inform the researchers in choosing the appropriate comparison in general medical practice. With a rapid influx of multiple treatments options, RWD can deliver information related to providing the best clinical benefit for a well-defined patient population. Second, with the increasing number of single-arm studies in the field of oncology, a large treatment effect can be expected but not feasibly confirmed by RCTs. In this scenario, RWD can act as a historical control for single-arm studies and provide a natural course of disease data (3). Third, the main focus of many oncology RCTs is survival benefit with the endpoints of overall survival (OS) or progression-free survival (PFS), which do not provide accurate long-term safety or quality of life data; RWD can be tailored to satisfy this informational shortfall (6).

China is confronted with many challenges and misunderstandings in the development of RWD, which include the lack of guidelines, a system of real-world study (RWS) design, and good accessibility and linkage of multi-source RWD, and other deficiencies. These issues have greatly affected the generation of accurate RWD, the development of high-quality RWSs, and the rational interpretation and use of RWE. The purpose of this paper is then to highlight the considerable value and potential drawbacks of RWE through on the basis of analysis of RWD in cancer research.

Types of RWSs in cancer research

Although RWD originates from general medical practice, this does not mean that RWS is the equivalent to observational study. In fact, RWS, as a new research concept, still constitutes goal-driven research, in which design and data collection remain based in conventional epidemiological methods.

Retrospective database study

The data for retrospective database research is mainly obtained from electronic medical records (EMRs), electronic health records (EHRs), the claims-based system, the death registry system, the health monitoring system, and other similar sources. These sources of information can not only be used to evaluate survival benefit but can also provide insight into long-term safety, rare adverse events, adjuvant therapy and care, quality of life, and socio-economic status (7).

The difficulty of retrospective database research lies in the database (8). It is critical to select the appropriate database, and extract and manage data so as to construct the research database. Thus, some key steps are essential for the quality control of the database, including the accessibility of the original database, the accuracy and integrality of the research population identity, the algorithm of data linkage, the rules for text structuring transformation, and other factors (9).

In this respect, a number of European Union (EU) member states have demonstrated the benefit of coordinating and utilizing EHRs. There exists an appreciable harmonization in the research community concerning the processing and quality control of each EHR across EU member states. For example, to facilitate the identification and assessment of the most eligible patients and the most relevant institutions for study sponsors, the EHR4CR program established a service system that reuses data from EHRs for clinical research and provides secure access to multiple EHR systems through the platform (10).

Registry study

When the available data are not useful to the purpose of the research, an observational study should be conducted to actively collect data. The authoritative definition of a registry study is located in “Registries for Evaluating Patient Outcomes: A User’s Guide” published by the Agency for Healthcare Research and Quality (AHRQ) in 2007. The updated 2014 version details the design, procedure, and evaluation in a registry study (11). In fact, the implementation of a registry study is quite similar to that of an RCT in terms of protocol formulation, the design of the case report form, data collection, and data management. Unlike the randomization in an RCT, a registry study is a prospective observational study. In some situations, a registry study can be combined with the retrospective database to

save costs and integrate a monitor system. Thus, the critical aspect of a quality registry study is the standardization and consistency of the data source. To integrate the different sources of data, one must consider the rules for combining the existing database with newly collected data, the external data linkage, and the updated rules for follow-up data.

The patient-based cancer registry is appreciated most particularly for its safety monitoring and evaluation, and the new indication development, especially for rare diseases. The well-known examples are the PARENT Joint Action and the European Network of Cancer Registries (ENCR) for EU-level initiatives. Other more recent examples include Flatiron's non-small-cell lung cancer cohort registries from 198 U.S. clinics (www.flatiron.com) and UNICANCER (www.unicancer.fr) based on 14,000 French breast cancer patients (www.unicancer.fr) (5). Population-based cancer registry studies mainly focus on the incidence and prognosis of cancers, cancer control activities, and cost-effectiveness analyses. In 1966, there were 32 cancer registries reporting on cancer incidence; in the 21st century, the number of cancer registries has expanded to 449, covering 21% of the world population as of 2006 (12).

Pragmatic clinical trials (PCTs)

Compared to the early efficacy detection of RCTs, PCTs evaluate the effectiveness of an intervention in a real-world circumstance to inform clinical decision making (13). This approach also ensures that the sorts of interventions tested can be plausibly rolled out in clinical practice and that the outcomes used to assess effectiveness are valid and easily understood by a range of users, including clinicians, patients, policy makers, and health commissioners (14). Furthermore, when novel cancer drugs meet specific criteria (i.e., conditional approval), PCTs may provide sufficient evidence for verification of clinical benefit and lead to full approval (15).

Generally, PCTs are designed as open-label, non-placebo-controlled or single-arm trials. The control group is often given conventional/standard treatment. A PCT is inclined to enroll all participants with the condition of interest who would be eligible to receive the intervention as part of usual care. Thus, patients are usually not blinded, and their compliance is intentionally neither maintained nor measured (16). However, the highly flexible design and implementation may result in more complex intercurrent events, such as a switch treatment, initiation of remedial treatment, unexpected terminal events, etc. which may result

in more challenges for the statistical design and analysis.

Values of RWS in cancer research

The widespread application of RWS offers enormous cornerstones and valuable insights into cancer research, especially on the population-level intervention studies, which have the potential to close gaps between the existed evidence and routine practice.

Population

Patients enrolled in clinical trials are usually not representative of those in routine practice, as they tend to be younger and have better health status than those in a real-world setting (17,18). RWD has the potential to provide insight into the complete distribution of the patient population, especially in the fields of demographics, comorbidity, and symptoms. For example, the Edmonton Symptom Assessment System (ESAS) is a registry assessment system that collects the symptom burden of cancer patients at the end of life. This form of RWD can provide important information for understanding symptoms of terminal stage cancer and inform the system-level planning of palliative care services (19).

RWD can establish data repositories to provide an overview of the social burden of population-based incidence and mortality. It not only forms the cornerstone of any initiative for guidance in cancer prevention and treatment but can also observe the disease burden on a population-level, enabling the health-care system to respond to future demands. In addition, population-based data can also be used to detect epidemiological transitions as disease burden shifts. One example observed is a parallel shift from cancers in a higher incidence in less-affluent populations (e.g., as cervical cancer and liver cancer) to cancers with increased incidence in affluent populations (e.g., breast, colorectal, or prostate cancer) (20). With the appropriate methodology and epidemiology, RWS can offer insight into shifting biological causes (e.g., cancers caused by viral infections), which elude the framing in RCTs. Meanwhile, incidence data can also shed light on the effectiveness of screening and prevention programs (e.g., the widespread hepatitis B vaccination and HPV vaccination) (21).

Intervention

In general, the standard intervention option guidelines for

recommendations are based on RCT-derived evidence, and thus do not take into account the conditions of patients who do not meet the rigorous admission criteria or who choose not to receive the treatment. RWD from health-care systems might reveal the adherence of real-world clinical practice with guidelines, along with the degree of underutilization and overutilization of cancer diagnosis and treatment. Evidence-based approaches can help the identification of the optimal utilization of treatment modalities in routine practice, while physician referral data can allow investigators to track patient management histories by identifying consultations with specialists (22). According to the past few years' reports, low utilization might be driven by upstream decision-making (not professional oncologists; e.g., surgeons) or the level of doctor-patient communication. Meanwhile, the overutilization of diagnosis and anticancer therapies based on clinical investigations usually further increase costs and delay treatment. Examples from the American and Canadian Choosing Wisely campaigns have listed ten commonly performed interventions that provide low-value options to patients so as to reduce unnecessary medical procedures and treatments (23,24). When several potential treatment options are feasible, RWD enables investigators to track the more commonly selected options from routine practice to make more informed policy choices (25). It can also identify the patient subgroups with impaired access to care or disproportionate disease burden, so that targeted programs or more economical measures can be designed.

RWD can also provide insights into the implementation and quality of decision-making and identify variations in practice. If practice variation is observed in a setting in which a better treatment option is feasible, RWD can inform initiatives to improve quality. If two treatment options are associated with similar survival outcomes, RWD can help policy makers decide which regimen is associated with lower morbidity and costs. In addition, the detailed EMR makes it possible to monitor the integrity of medical care delivery in routine practice, and, in this way, RWD can show the extent of incomplete treatment and modifiable factors.

Outcomes

Traditional treatment efficacy assessment in RCTs almost exclusively focuses on survival benefit, such as OS and PFS. However, emerging sources of RWD can document more comprehensive prognostic information in routine practice,

including improvement of patients' symptoms, quality of life, treatment-derived toxicities and/or complications, organ preservation rates, economic status, etc. Compared to the narrower perspective of survival outcomes, these data not only amply complement RCTs, but also act as a knowledge base for medical institutions to improve the quality of care in routine practice. With some short-term outcomes, such as 90-day postoperative mortality, readmission rate, or chemotherapy-derived toxicities, the metrics from RWD also enable investigators to compare their observations with those reported in the same/history period (26). With long-term outcomes, such as medical survival and 5-year survival, records from RWD provide a fundamental evaluation for a population of representative patients to verify the results from RCTs. A large gap between the trial and realistic outcomes may indicate the necessity for care improvement initiatives to be implemented (27).

Further to this, RWD offers clinicians insights into rare diseases and rare populations that can not be tested in RCTs of larger cohorts with longer follow-up durations (such as with elderly populations). The case of palbociclib approval by the FDA for the treatment of rare male breast cancer attests to the value of RWD in new drug development (28). Even for common diseases, RWD can offer more authentic evidence of the more rare or delayed toxicities [such as the cardiac event after breast radiotherapy (RT)] than RCTs can (29).

Generation of RWD

The variables selected in RWD depend on the study aims being explored. Booth *et al.* (17) summarized the fundamental variables often collected from RWD. For example, studies focusing on patients require variables of age, sex, occupation, diagnosis, disease history, and comorbidity; studies focusing on treatment need variables of extent of disease, intent of treatment, chemotherapy/RT dose, procedures, and treatment start and end date; and studies focusing on treatment outcomes always collect variables of adverse events, readmission rate, and survival benefit.

Original data

The completeness and accuracy of the original data are determinants for the quality of cancer research. The validity and reliability of the data should be considered carefully before the study is conducted. The first critical issue to

resolve is the data integrality. Random missing data reduce the data precision, whereas non-random missing data can lead to bias. Data feasibility is another significant factor and is especially relevant to the comprehensiveness of RWD. A prominent limitation for RWD sources is the lack of variables concerning patient prognosis, care, outcomes, life status, and disease burden, which becomes important for a comprehensive assessment of the effectiveness of treatment.

To capture RWD, a scheme for the RWD needs to be constructed. For example, a database study needs to follow a certain set of steps. First, the appropriate database according to the research purpose should be selected, and the accessibility and quality of the database should be evaluated; then, the original database structure, variable meaning, and source should be verified; after that, the data extraction method can be designed according to the database's advantages and limitations. Factors that can influence the original data quality include data collection methods, the skills and training of data managers, and some external factors (such as physician remuneration or hospital funding). After that, the primary data should be further checked. First, the accuracy of data extraction should be assessed. It can be checked and compared from a random sample of electronic data from regional administrative sources; then, the data check proceeds to an assessment of missing data, logical inconsistencies (such as impossible date of birth), and outliers or abnormal values (such as multiple organ excision); at last, data cleaning should be conducted to correct the wrong data.

To adjust to a lack of information in a single database, RWS can incorporate multiple linked databases. Theoretically, data linkage should be implemented with a unique identifier, such as ID number or health insurance number. Otherwise, a probabilistic link can be used, but the linkage quality may be reduced accordingly. Zhu *et al.* used simulation studies to obtain a matching score for each partial identifier (such as age, name, and address) according to the matching degree between databases, which was summed across fields with a specific threshold to distinguish matches (30).

Derived data

A complex and crucial task in RWD is to distinguish the similar but non-identical variables. This procedure often involves the “lumping and splitting” for all the exposure related to population, intervention, and outcomes (18). The study population must be typically defined into

clinically meaningful groups, which should be defined by the clinicians, whereas investigators need to complete the accurate identification and generation of variables under the existing situations. The definition of derived variables should be decided before conducting the analysis, as the changing of any of the critical definitions for the desired results may be tempting. Even without a deliberate attempt to obtain particular outcomes, attempts made with several different tentative groupings after obtaining the data may accidentally increase the probability of a statistically significant result.

Booth *et al.* have elaborated upon this topic using the treatment patterns of bladder cancer (31) and lung cancer (32). In these studies, groups were assigned by presumed treatment intent with the identification of adjuvant treatment time periods. Based on comprehensive consideration, the researcher defined any chemotherapy or RT within 16 weeks after surgery as adjuvant therapy. Otherwise, any treatment beyond this time point was defined as palliative therapy due to disease progression. From clinical insight and experience, if this threshold was set prematurely after surgery (such as 8 weeks), there was a risk of excluding some patients who received adjuvant therapy, whereas a later cut-off point (such as 32 weeks) might have included some patients receiving chemotherapy for early cancer metastasis. However, there is no standard, perfect approach for these types of problems, and indirect inferences about the intent of treatment should be prudently considered.

Implementation of RWS

Although RWS has the advantage of good extrapolation, it will still inevitably conceal any potential bias. The more variations the search environment and data sources present, the more complex the data processing and analysis techniques required.

Protocol

In clinical trials, researchers select patient subgroups in advance, whereas investigators of RWS tend to perform such stratification *post hoc* since RWD always exists before the study is implemented, in some situations. But it is critical to properly formulate and follow RWS protocol. The formulation of study design should be collaboratively implemented by researchers, clinicians, and statistical analysts, the key points of which are summarized in *Table 1*.

Table 1 The primary elements of RWS protocol

Items	Contents
Research purposes	
Background	Background and research issues
Aims	Assumptions and aims
Research design	
Type	Aims and data sources determine the type of research (such as database study, registry study, PCT, etc.)
Control	Standard routine treatment
Admission criteria	Variable definition rules
Exposure/intervention	Variable selection rules or detection methods
Follow-up time	The follow-up scheme, time-window, censor, etc.
Sample size	Generally, there is no strict rule on sample size determination, and all samples meeting the admission criteria are included as much as possible
Bias control	Bias is controlled in the design stage and measure in analysis
Statistical analysis	
Data check	Check and cleaning rules
Derived variables	Definition and transformation of derived variables
Screening	Selection of subjects
Missing data	Definition of missing data and the remedy scheme
Sensitivity analysis	Sensitivity analysis plan
Subgroup analysis	Subgroup analysis plan
Analytical method	Statistical analysis plan

PWS, real-world study; PCT, pragmatic clinical trial.

Bias

Although the RWD sample is generated from routine practice, it does not mean the respondents in RWS are truly representative of real-world patients. In RWS, bias is inevitable due to a lack of internal validity or incorrect assessment of the association between exposures and effects in the target population (33). Sanderson *et al.* have delineated three fundamental domains to evaluate bias (34): appropriate selection of participants, accurate measurement of exposure, and outcome and appropriate controls of

Table 2 Common biases in RWS

Type	Name
Selection bias	Neyman bias/prevalence-incidence bias
	Berkson's bias/admission rate bias
	Length bias
Information bias	Competing risks
	Immortal time bias/survivor treatment selection bias
	Healthcare access bias
	Misclassification bias
	Inducement bias
	Recall bias
	Lead time bias
	Protopathic bias

RWS, real-world study.

confounding. There are many measurement tools for bias evaluation, such as Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) (35). However, some of the assessment tools used weighted scores for various component elements, which are still inappropriate and inconsistent. Therefore, investigators suggest that the bias evaluation should always concentrate on a few, principal sources (33). *Table 2* lists the common biases in RWS, some of which are imperceptible but nonetheless significant for effect evaluation.

Immortal time bias (36), is also known as survivor treatment selection bias (37), guarantee time bias, or time-dependent bias. Immortal time is a time span in the observation or follow-up duration for a cohort during which the outcome under study could not have occurred (38) (otherwise, the follow-up should be terminated due to an outcome appearing before the intervention began). When a patient is not truly immortal during this span of time, the patient has to remain event-free until the exposure (treatment) begins. If this unexposed period is defined incorrectly (when, at the time of cohort definition, exposure starting and follow-up starting are not matched), immortal time bias usually occurs. In pharmacoepidemiology, this bias often appears in the treatment evaluation and safety monitoring for the database study. When a patient begins a given treatment before the cohort identification or after the doctor's actual order time, different approaches to form the cohort may lead to variations in resulting immortal time bias.

Carl *et al.* (39) reported that this type of bias was present in 7.6% of 682 articles from 9 reputable medical journals, 67.3% of which might have had mistaken conclusions as a consequence. If, for example, a study focusing on survival benefit assigns two groups according to whether the subjects receive a given treatment within 2 months after diagnosis. With this definition, patients dying within 2 months of diagnosis are grouped in the non-treatment group, which will reduce the average survival of this group if the survival time is measured from the diagnosis date. In RCTs, the intention-to treat analysis can be used to deal with this situation, whereas this method is not appropriate for database study since the treatment intention cannot be determined before it is actually received by patient. In view of this situation, Booth *et al.* recommend starting measurement of outcomes after 2 months to ignore the earlier deaths (17). Another method to prevent immortal time bias in RWS is to define the subjects as patients undergoing the given treatment for the first time. However, studies applying EMRs from a single situation may lack information related to the integrate exposure period due to data feasibility, which creates a methodology challenge for the left censoring data.

Misclassification bias occurs when the target subjects are classified as non-target subjects and vice versa, or if sensitivity/specificity for the exposure/outcome identification is not explicit (40). In routine practice, it is very uncommon to have a perfect tool to control misclassification. Since so many factors may affect it, like random error in data capture, missing data, and end-digit preference, it is unavoidable to a certain degree. Meanwhile, the codes for disease identification, drug application, and other computational procedures may also produce it. The most common misclassifications in RWS are drug exposure misclassification and outcome misclassification.

For cancer research, drug exposure misclassification often appears due to the incorrect identification of exposure dose and period. A counter example may be observed from a bladder cancer study which explored the survival benefit of cystectomy in patients with advanced bladder cancer. Though, it might have failed to show effectiveness correctly, the authors adopted propensity score (PS) methodologies to control confounding (41). Because the derived variable of disease stage leads to the misclassification of patients with an earlier disease stage (treated with curative intent) as having cystectomy with palliative intent (42-44).

Outcome misclassification is usually generated from the definition of survival time. For the studies with an OS

endpoint, the degree of misclassification is expected to be small. Due to the fact that the cause of death registration cannot be updated in real time, there is always a potential lag time. It is worth noting that the death registry system does not fully update on time and tends to lag. In this situation, the risk of misclassification may increase for the “softer” endpoint, such as PFS (5). To control this type of bias, restricted and combined definition of survival time can be helpful, and can include identification mode, combined discrimination, and sensitivity analysis (45,46).

Analysis

Treatment selection in routine practice is associated with a patient’s baseline characteristics; in turn, inherent imbalance between different treatment groups can lead to very different outcomes. Even if proper design for the control of biases has been performed, residual confounding may still occur since some of them may not be found or measured. Statistical methods, such as multivariate analysis, stratification, or matching, are helpful in mitigating this potential confounding, but entail a downside where in one can adjust only for the known variables.

Another possible solution to adjust confounding is PS. The idea of PS is to replace the confounding variables with a function which is then used to adjust the imbalance in patient characteristics between groups. This can be interpreted as a method of “post-randomization” for anon-randomized study. The main advantage of PS is that many covariates can be included simultaneously without the risk of overfitting the model. However, it can only be adjusted for observed confounding variables, and not for unmeasured variables (47).

A strategy to overcome the undiscovered or unmeasured confounding is instrumental variable analysis (IVA), in which subjects are grouped by a marker for different practice policies. However, it may be hard to find valid instruments in practice and challenging to apply (48).

Consideration of RWE

Although RWD offers essential experiences and insights into treatment effectiveness in a real-world population of patients, it is critical that RWD is carefully designed and considers the context of existing evidence. For the methodological limitations of RWS, we highlight some specific considerations to illustrate the benefits and pitfalls of RWE (4).

Validation of RCTs

RCT is the gold standard for efficacy evaluation. The use of RWE can validate/augment this efficacy into effectiveness in routine practice. For example, the results of several RCTs of adjuvant chemotherapy (ACT) for non-small-cell lung cancer have indicated improved outcomes, and RWE from the real-world population showed an improvement in OS in the era of ACT, which mitigates some common forms of bias via instrumental variables (32). On the other hand, RWS may also find non-effectiveness in spite of the efficacy demonstrated in RCTs. In the case of Sanoff's investigation, sorafenib for hepatocellular carcinoma in older and sicker patients showed unexpected outcomes in routine practice (49). The negative effectiveness of RWE is also important as those conformed efficacy for a novel treatment in the real world.

Lack of efficacy in RCTs

When previous RCTs have shown a lack of efficacy, the effectiveness demonstrated in RWE may be problematic. Using Bayesian logic for analysis, it is reasonable to find effectiveness in RWE where preexisting RCT evidence has established efficacy. However, it is difficult to infer the benefit of cancer treatment to patients in routine practice without the efficacy being demonstrated under the ideal situations of an RCT. In this case, we consider the prior risk of a negative result to be so high that the RWE for effectiveness in this circumstance is considered more likely to be an artifact. One relevant example is a series of comparative analyses from Karim and Booth (4) on the effectiveness of adjuvant therapy for stage II colorectal cancer. Previous RCTs (50,51) and a meta-analysis (52) had proven ACT did not offer a significant OS benefit in stage II colon cancer, especially the IMPACT B2 meta-analysis comprising 1,016 patients (53). However, a retrospective database study of 153,110 patients reported that the ACT resulted in an 18% absolute improvement in 5-year OS (hazard ratio =0.71, $P<0.001$), in which all subgroups showed a degree of benefit. With a rigorous review, Karim and Booth (4) claimed the investigator did not consider the effect of comorbidity on the primary analysis. Despite multivariate regression analysis and the PS methods used, this research was vulnerable to immortal time bias. Because ACT was defined as starting ACT within 30 days after surgery, those starting ACT after 30 days were assigned to the non-ACT group. Thus, early deaths between 30 days

postsurgery and the start of ACT in routine practice would distort the survival comparisons between groups.

In another example, RCTs confirmed that neoadjuvant RT for localized rectal cancer improved local control rates and reduced colostomy (54,55), but did not improve survival benefit in total mesorectal excision (TME) (56). Although an earlier meta-analysis reported conflicting results, studies included in the meta-analysis predated adoption of TME (57). In addition, RCT conducted by The Dutch Colorectal Cancer Group verified that neoadjuvant RT was associated with decreased local recurrence but not with OS (58). However, a retrospective study reported that complete contrast preoperative RT led to a 10% improvement in 5-year OS compared to incomplete RT but with no benefit in local control (59). In fact, this research is still problematic, since the large OS benefit is more likely due to "residual confounding from patient characteristics that allowed complete delivery of RT rather than the RT itself" (4).

Conclusions

The development of EHRs has greatly enhanced the feasibility of collecting RWD, and thus has the potential to fill the gap between clinical trial evidence and real-world populations. The widespread availability of RWD offers valuable insights into cancer treatment, disease management, disease burdens, and socio-economic status in routine practice that will improve patient care and health policy making.

There do exist inherent limitations in data quality and study design, while the internal validity remains an unknown in many RWSs. The research community should thus carefully consider the validity and reliability of such studies. Since clinicians are less likely to be familiar with the potential pitfalls of RWD as they are with those of RCTs, it is advisable to perform RWS with a multidisciplinary team involving clinicians, epidemiologists, and biostatisticians. RWD is suited for the validation of existing evidence of a given treatment (to ensure efficacy translates into effectiveness), whereas it is not appropriate to adopt new therapies on the basis of RWD in isolation, particularly if there is no evidence of treatment benefit from RCTs. In this situation, claims of effectiveness may probably be misleading and artificial, and thus researchers should be critical of the plausibility of such outcomes (4).

Although emerging sources of RWD provide important cornerstones for many clinical and health policy strategies

in cancer control, RCTs should remain the gold standard for validating the efficacies of new treatments, particularly new drugs (4,5,17). Even with advanced methodologies available for bias control, many challenges still need to be overcome before RWD can become integrated into clinical decision-making. Meanwhile, there is no standard strategy for the transformation of RWD into RWE. Nevertheless, RWD will offer increasingly relevant information and experience for benefit–risk evaluation in the field of oncology.

Acknowledgments

Funding: The authors gratefully acknowledge the financial support for this work provided by research grants from the National Natural Science Foundation of China (No. 81803328, 81773553, and 81973141) and from the Social Development Major Project of Shaanxi Province in China (No. 2020SF-282).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form. (available at <http://dx.doi.org/10.21037/tbcr-20-27>). JX serves as an unpaid editorial board member of *Translational Breast Cancer Research* from Mar 2020 to Feb 2022. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Veenstra CM, Wallner LP, Jagsi R, et al. Long-Term Economic and Employment Outcomes Among Partners of Women With Early-Stage Breast Cancer. *J Oncol Pract* 2017;13:e916-e926.
- Pignatti F, Luria X, Abadie E, et al. Regulators, payers, and prescribers: can we fill the gaps? *Lancet Oncol* 2011;12:930-1.
- Trotta F, Leufkens HG, Schellens JH, et al. Evaluation of oncology drugs at the European Medicines Agency and US Food and Drug Administration: when differences have an impact on clinical practice. *J Clin Oncol* 2011;29:2266-72.
- Karim S, Booth CM. Effectiveness in the Absence of Efficacy: Cautionary Tales From Real-World Evidence. *J Clin Oncol* 2019;37:1047-50.
- Skovlund E, Leufkens HGM, Smyth JF. The use of real-world data in cancer drug development. *Eur J Cancer* 2018;101:69-76.
- Kleijnen S, Leonardo Alves T, Meijboom K, et al. The impact of quality-of-life data in relative effectiveness assessments of new anti-cancer drugs in European countries. *Qual Life Res* 2017;26:2479-88.
- Sorensen HT, Sabroe S, Olsen J. A framework for evaluation of secondary data sources for epidemiological research. *Int J Epidemiol* 1996;25:435-42.
- Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *Z Evid Fortbild Qual Gesundheitswes* 2016;115-116:33-48.
- 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 Report--Part I. *Value Health* 2009;12:1044-52.
- Makady A, Stegenga H, Ciaglia A, et al. Practical Implications of Using Real-World Evidence (RWE) in Comparative Effectiveness Research: Learnings from IMI-GetReal. *J Comp Eff Res* 2017;6:485-90.
- AHRQ Methods for Effective Health Care. In: Gliklich RE, Dreyer NA. editors. *Registries for Evaluating Patient Outcomes: A User's Guide*. Rockville (MD): Agency for Healthcare Research and Quality (US); 2010.
- Parkin, Donald M. The evolution of the population-based cancer registry. *Nat Rev Cancer* 2006;6:603-12.
- Roland M, Torgerson DJ. What Are Pragmatic Trials? *BMJ* 1998;316:285.
- Ford I, Norrie J. Pragmatic Trials. *N Engl J Med* 2016;375:454-63.
- Koehler M, Donnelly ET, Kalanovic D, et al. Pragmatic randomized clinical trials: a proposal to enhance evaluation

- of new cancer therapies with early signs of exceptional activity. *Ann Oncol* 2016;27:1342-8.
16. 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.
 17. Booth CM, Karim S, Mackillop WJ. Real-world data: towards achieving the achievable in cancer care. *Nat Rev Clin Oncol* 2019;16:312-25.
 18. Mitchell AP, Harrison MR, Walker MS, et al. Clinical Trial Participants With Metastatic Renal Cell Carcinoma Differ From Patients Treated in Real-World Practice. *J Oncol Pract* 2015;11:491-7.
 19. Seow H, Barbera L, Sutradhar R, et al. Trajectory of performance status and symptom scores for patients with cancer during the last six months of life. *J Clin Oncol* 2011;29:1151-8.
 20. Bray F, Jemal A, Grey N, et al. Global Cancer Transitions According to the Human Development Index (2008-2030): A Population-Based Study. *Lancet Oncol* 2012;13:790-801.
 21. Rabkin CS, Biggar RJ, Horm JW. Increasing incidence of cancers associated with the human immunodeficiency virus epidemic. *Int J Cancer* 1991;47:692-6.
 22. Kankesan J, Shepherd FA, Peng Y, et al. Factors associated with referral to medical oncology and subsequent use of adjuvant chemotherapy for non-small-cell lung cancer: a population-based study. *Curr Oncol* 2013;20:30-7.
 23. Pramesh CS, Chaturvedi H, Reddy VA, et al. Choosing Wisely India: ten low-value or harmful practices that should be avoided in cancer care. *Lancet Oncol* 2019;20:e218-e223.
 24. Schnipper LE, Smith TJ, Raghavan D, et al. American Society of Clinical Oncology identifies five key opportunities to improve care and reduce costs: the top five list for oncology. *J Clin Oncol* 2012;30:1715-24.
 25. Quirt JS, Siemens DR, Zaza K, et al. Patterns of Referral to Radiation Oncology among Patients with Bladder Cancer: a Population-based Study. *Clin Oncol (R Coll Radiol)* 2017;29:171-9.
 26. Booth CM, Nanji S, Wei X, et al. Management and Outcome of Colorectal Cancer Liver Metastases in Elderly Patients: A Population-Based Study. *JAMA Oncol* 2015;1:1111-9.
 27. Patel MI, Bang A, Gillett D, et al. Poor survival of females with bladder cancer is limited to those aged 70 years or over: a population-wide linkage study, New South Wales, Australia. *Cancer Med* 2015;4:1145-52.
 28. Wedam S, Fashoyin-Aje L, Bloomquist E, et al. FDA Approval Summary: Palbociclib for Male Patients with Metastatic Breast Cancer. *Clin Cancer Res* 2020;26:1208-12.
 29. Seruga B, Sterling L, Wang L, et al. Reporting of serious adverse drug reactions of targeted anticancer agents in pivotal phase III clinical trials. *J Clin Oncol* 2011;29:174-85.
 30. Zhu Y, Matsuyama Y, Ohashi Y, et al. When to conduct probabilistic linkage vs. deterministic linkage? A simulation study. *J Biomed Inform* 2015;56:80-6.
 31. Booth CM, Siemens DR, Li G, et al. Perioperative chemotherapy for muscle-invasive bladder cancer: A population-based outcomes study. *Cancer* 2014;120:1630-8.
 32. Booth CM, Shepherd FA, Peng Y, et al. Adoption of adjuvant chemotherapy for non-small-cell lung cancer: a population-based outcomes study. *J Clin Oncol* 2010;28:3472-8.
 33. Delgado-Rodriguez M, Llorca J. Bias. *J Epidemiol Community Health* 2004;58:635-41.
 34. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007;36:666-76.
 35. Vandembroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Int J Surg* 2014;12:1500-24.
 36. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf* 2010;16:241-9.
 37. Glesby MJ, Hoover DR. Survivor treatment selection bias in observational studies: examples from the AIDS literature. *Ann Intern Med* 1996;124:999.
 38. Suissa S. Immortal Time Bias in Pharmaco-Epidemiology. *Am J Epidemiol* 2008;167:492-9.
 39. Carl VW, Darryl D, Forster AJ, et al. Time-dependent bias was common in survival analyses published in leading clinical journals. *J Clin Epidemiol* 2004;57:672-82.
 40. Copeland KT, Checkoway H, McMichael AJ, et al. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol* 1977;105:488-95.
 41. Seisen T, Sun M, Leow JJ, et al. Efficacy of High-Intensity Local Treatment for Metastatic Urothelial Carcinoma of the Bladder: A Propensity Score-Weighted Analysis From the National Cancer Data Base. *J Clin Oncol* 2016;34:3529-36.
 42. Mathieu R, Peyronnet B, Bensalah K. Re: Efficacy of High-Intensity Local Treatment for Metastatic Urothelial Carcinoma of the Bladder: A Propensity Score-Weighted Analysis from the National Cancer Data Base. *Eur Urol* 2016;70:894.

43. Woldu SL, Lotan Y. Re: Efficacy of High-intensity Local Treatment for Metastatic Urothelial Carcinoma of the Bladder: A Propensity Score-Weighted Analysis from the National Cancer Data Base. *Eur Urol* 2016;70:893.
44. Chang SS. Re: Efficacy of High-Intensity Local Treatment for Metastatic Urothelial Carcinoma of the Bladder: A Propensity Score-Weighted Analysis from the National Cancer Data Base. *J Urol* 2018;199:1396.
45. Keogh RH, Shaw PA, Gustafson P, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1-Basic theory and simple methods of adjustment. *Stat Med* 2020;39:2197-231.
46. Shaw PA, Gustafson P, Carroll RJ, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 2-More complex methods of adjustment and advanced topics. *Stat Med* 2020;39:2232-63.
47. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med* 1997;127:757-63.
48. Burgess S, Thompson SG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011;40:755-64.
49. Sanoff HK, Chang Y, Lund JL, et al. Sorafenib Effectiveness in Advanced Hepatocellular Carcinoma. *Oncologist* 2016;21:1113-20.
50. Matsuda C, Ishiguro M, Teramukai S, et al. A randomised-controlled trial of 1-year adjuvant chemotherapy with oral tegafur-uracil versus surgery alone in stage II colon cancer: SACURA trial. *Eur J Cancer* 2018;96:54-63.
51. Gray R, Barnwell J, McConkey C, et al. Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. *Lancet* 2007;370:2020-9.
52. Gill S, Loprinzi CL, Sargent DJ, et al. Pooled analysis of fluorouracil-based adjuvant therapy for stage II and III colon cancer: who benefits and by how much? *J Clin Oncol* 2004;22:1797-806.
53. Efficacy of adjuvant fluorouracil and folinic acid in B2 colon cancer. International Multicentre Pooled Analysis of B2 Colon Cancer Trials (IMPACT B2) Investigators. *J Clin Oncol* 1999;17:1356-63.
54. Glimelius B, Grönberg H, Järhult J, et al. A systematic overview of radiation therapy effects in rectal cancer. *Acta Oncol* 2003;42:476-92.
55. Colorectal Cancer Collaborative Group. Adjuvant radiotherapy for rectal cancer: a systematic overview of 8,507 patients from 22 randomised trials. *Lancet* 2001;358:1291-304.
56. Wong RK, Berry S, Spithoff K, et al. Preoperative or postoperative therapy for stage II or III rectal cancer: an updated practice guideline. *Clin Oncol (R Coll Radiol)* 2010;22:265-71.
57. Cammà C, Giunta M, Fiorica F, et al. Preoperative radiotherapy for resectable rectal cancer: A meta-analysis. *JAMA* 2000;284:1008-15.
58. Kapiteijn E, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *N Engl J Med* 2001;345:638-46.
59. Freischlag K, Sun Z, Adam MA, et al. Association Between Incomplete Neoadjuvant Radiotherapy and Survival for Patients With Locally Advanced Rectal Cancer. *JAMA Surg* 2017;152:558-64.

doi: 10.21037/tbcr-20-27

Cite this article as: Li C, Zang J, Li F, Hu H, Wang L, Xia J. The essence of real-world study in cancer research. *Transl Breast Cancer Res* 2020;1:17.