Real-world data: towards achieving the achievable in cancer care

Christopher M. Booth[1], Safiya Karim[2] and William J. Mackillop[1]*

Abstract | The use of data from the real world to address clinical and policy-relevant questions that cannot be answered using data from clinical trials is garnering increased interest. Indeed, data from cancer registries and linked treatment records can provide unique insights into patients, treatments and outcomes in routine oncology practice. In this Review, we explore the quality of real-world data (RWD), provide a framework for the use of RWD and draw attention to the methodological pitfalls inherent to using RWD in studies of comparative effectiveness. Randomized controlled trials and RWD remain complementary forms of medical evidence; studies using RWD should not be used as substitutes for clinical trials. The comparison of outcomes between nonrandomized groups of patients who have received different treatments in routine practice remains problematic. Accordingly, comparative effectiveness studies need to be designed and interpreted very carefully. With due diligence, RWD can be used to identify and close gaps in health care, offering the potential for short-term improvement in health-care systems by enabling them to achieve the achievable.

Within the oncology community, growing interest exists in using data from the real world to address clinical and policy-relevant questions that cannot be answered with data from clinical trials. In this article, the term real-world data (RWD) refers to population-level data obtained from cancer registries and not n-of-1 patient-based studies[1]. The FDA has defined RWD as data relating to patient health status and/or the delivery of health care that are collected from sources that include electronic health records (EHRs), billing claims and product and disease registries[2]. The term RWD is being used increasingly for research purposes; a search for the terms "real-world data", "real-world evidence" or "registry" in PubMed would have yielded 2,435 citations in 2002 but 14,956 citations in 2016. By way of control, a similar search for studies on colon cancer would result in 3,028 and 6,420 hits, respectively (FIG. 1).

Cancer registries in the first half of the 20th century enabled early descriptions of cancer incidence. In the 1990s, investigators began to expand this work to include information about cancer treatment and outcomes, spawning early use of RWD in oncology. At that time, a parallel process in the USA[3] and Ontario[4] (Canada) led to linkage of treatment records to original cancer registry data[3,4]. Population-based cancer registries comprise information from a variety of sources; for example, cancer incidence can be derived using hospital and/or pathology records. Population-level data sources also enable the evaluation of health-care system performance and care at the level of an individual patient with cancer. The contemporary sources of oncology RWD sources typically fall into one of two categories: an opportunistic collection of existing administrative data sources (such as Surveillance, Epidemiology and End Results (SEER)–Medicare in the USA and Ontario Cancer Registry-linked treatment files in Canada) or a purposefully designed comprehensive single database (for example, the National Cancer Database in the USA). Further details regarding the evolution of population-based cancer registries have been elegantly described elsewhere[5]. Most published studies with RWD include a combination of data from cancer registries, hospital treatment records and insurance claims. In the contemporary era, substantial interest exists in using emerging sources of RWD from multi-institutional EHRs, mobile applications and wearable technologies[6]. These new data sources offer the potential benefit of improved granularity (including details on characteristics such as smoking status, physical activity and/or biomarkers), although their internal validity remains largely unknown. Future efforts are needed to better describe the validity and applicability of these new data sources in the peer-reviewed scientific literature.

The three fundamental elements of RWD include descriptions of patients, treatments and outcomes (BOX 1). The relative importance of each variable will depend on the research question being asked; for example, studies of incidence require substantially fewer variables (such as age, sex, date of diagnosis or tumour histology) than do studies of treatment outcomes (which require, for example, data on the extent of disease, intent of treatment or chemotherapy agents and/or dosing).

[1]*Division of Cancer Care and Epidemiology, Queen's University Cancer Research Institute, Kingston, Ontario, Canada.*

[2]*Department of Oncology, University of Calgary, Calgary, Alberta, Canada.*

**e-mail: booth@queensu.ca*

https://doi.org/10.1038/s41571-019-0167-7

Key points

- In the past decade, interest in linking electronic health records (EHRs) of treatment and outcome to cancer registry data has increased; these sources of real-world data (RWD) can offer unique insights into patients, treatments and outcomes in routine oncology practice.
- The quality of RWD relates to the quality of the primary data (completeness, accuracy and comprehensiveness), data linkages and derived variables assessed.
- Emerging sources of RWD from EHRs, mobile applications and wearable technologies, can offer improved granularity over traditional sources of RWD, but their internal validity and applicability remain largely unknown.
- For health-care systems, the use of RWD enables measurement of performance and can be used to identify targets for future quality improvement interventions.
- RWD can offer important insights into outcomes achieved with new anticancer therapies in routine practice; however, the comparison of outcomes between nonrandomized groups of patients who have received different treatments in routine practice remains problematic and is not a substitute for randomized controlled trials.
- Investigators working with RWD need to move beyond simply describing gaps in care and towards designing intervention studies to improve patient care and outcomes.

The challenges related to the latter include identifying treatment toxicities that do not lead to hospitalization or death (such as neuropathy, severe nausea and/or fatigue) and ascribing cause of death. In this Review, we discuss these challenges and those related to the quality of RWD. We also provide a framework for the use of RWD, outlining contexts in which RWD can be informative (TABLE 1), and draw attention to the pitfalls inherent in using RWD for studies of comparative effectiveness.

Quality of real-world data

Quality of the primary data. In studies using RWD, the quality of the primary data depends on its completeness and accuracy[7]. A critical consideration is the extent to which data are missing at random: random missing data decrease the precision of observations, whereas non-random missing data can lead to biased results. Thus, investigators working with RWD need to carefully consider data validity (that is, do the data tell us what really happened?) and reliability (that is, are data elements captured consistently?). Distinct from data quality is the comprehensiveness of available data; a limitation of most sources of RWD is the lack of information about patient prognosis, care and outcomes, including data on performance status, disease stage, intent of treatment or disease burden, which becomes particularly relevant in studies of comparative effectiveness. Some emerging sources of RWD have mitigated these limitations by incorporating additional patient-related and disease-related variables. In a 2018 study, Khozin et al.[8] used EHRs from clinics within the Flatiron Health network to describe the outcomes of patients with non-small-cell lung cancer (NSCLC) treated with nivolumab and pembrolizumab in routine practice. This study included important prognostic variables, such as smoking status and relevant biomarkers (including programmed cell death 1 ligand 1 (PD-L1), EGFR or ALK) not commonly available in previous sources of RWD. Investigators working with RWD can improve the comprehensiveness of existing data sets by adding new sources over time. Important examples of this approach include the addition of patient-reported symptoms[9] or educational outcomes of paediatric cancer survivors[10]. The feasibility of linking additional data sources to cancer registry files will be guided in part by ethics and privacy regulations specific to each jurisdiction.

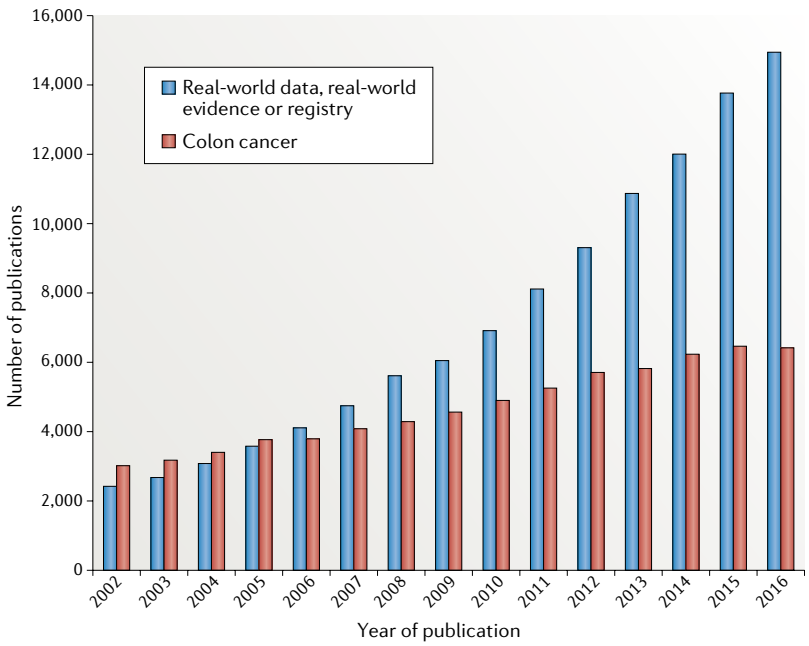

Fig. 1 | **Increased reporting of studies using real-world data.** Temporal trends in publications identified in the PubMed database for each year between 2002 and 2016 using the search terms "real-world data", "real-world evidence" or "registry" (sum of results) versus "colon cancer" are shown.

Data quality can be checked using several mechanisms: validation studies compare electronic data from administrative sources against a random sample of patient charts[11]; electronic data sources can be analysed to evaluate accuracy[12]; and data cleaning involves checking for logical inconsistencies (for example, impossible dates of birth or the presence of single patients with multiple records of radical prostatectomy). In our studies of the outcomes of patients with bladder cancer receiving cystectomy, we identified inconsistencies in the electronic data that prompted a detailed review of primary records[13]. Specifically, redundant sources of administrative treatment records (for example, physician billing records versus hospital procedure records) suggested that some patients with bladder cancer classified as having a cystectomy might have actually had cystoscopy (thus limiting the validity of the potentially duplicated data). Moreover, this coding error was more likely to occur in hospitals admitting lower numbers of patients (thus limiting the reliability of data from those centres). To address this differential bias, we obtained primary pathology reports for all patients with bladder cancer in Ontario and reviewed source documents to accurately identify which patients underwent cystectomy and at which disease stage.

Factors that influence the quality of primary data include how the data are collated (passive collection versus active collection), the skills, training and oversight of those individuals who collect and collate data, whether external influences might promote accurate data capture (for example, when data collection influences

Box 1 | **Variables for different domains in cancer-related data sets**

Patient
- Age
- Sex
- Date of diagnosis
- Tumour histology
- Tumour stage at diagnosis
- Tumour pathological and molecular details (margin status, extent of nodal harvest, lymphovascular invasion, grade and molecular markers (such as hormone receptor or mutation status))
- Socio-economic status
- Performance status
- Comorbidity
- Organ function (for example, renal)
- Height, weight and/or body surface area
- Second primary cancer

Treatment
- Extent of disease at time of treatment
- Intent of treatment
- Surgical procedure
- Surgical admission and/or discharge date
- Radiotherapy dose, fractionation, volume and technique
- Chemotherapy drugs, doses and schedule
- Treatment start and end dates
- Consultations with specialists

Outcome
- Date of death
- Cause of death
- Diagnosis and dates of admission and/or discharge from hospital
- Diagnosis and dates of emergency room visits
- Patient-reported symptom burden and quality of life
- Functional status
- System resource utilization and/or costs

physician remuneration or hospital funding, data completeness is favoured) and whether and how often the data quality is audited. In a passive cancer registry, data flow routinely from participating institutions to a central repository with minimal oversight (using, for example, routine automated data capture). An active cancer registry involves trained staff (such as cancer registrars), who actively seek cases and send linked records to the central repository.

Quality of data linkages. Most studies of RWD include information from different data sets, which therefore need to be linked. Ideally, data linkage should be performed with a unique identifier. For example, in Ontario (a province with a single payer for health care), government-issued health-insurance numbers are used to link data sets. This form of linkage is generally of higher quality than that of probabilistic linkage, in which each partial identifier (for example, age, name or address) is assigned a score on the basis of how well it matches between data sets. This score is summed across fields, and a specific threshold is assigned to distinguish between matches and non-matches[14,15].

Quality of derived variables. A labour-intensive and crucial step in working with RWD is deciding how to categorize similar but non-identical patients, treatments and outcomes. This approach is often referred to as the 'lumping and splitting' of patient groups or exposures. These decisions must be made at the outset of the study, before conducting the main analysis. Otherwise, changing these critical definitions as the results emerge in order to obtain the expected or desired outcome would be tempting. Even without a deliberate attempt to obtain a particular result, experimenting with several different tentative groupings increases the probability of obtaining a statistically significant result owing to chance alone. Furthermore, without considerable rigour, even keeping track of the number of different combinations of rules that have been used, in order to determine the significance of any apparently positive results, might be difficult. Clinical trialists typically select patient subgroups a priori, whereas investigators of RWD studies tend to perform such stratification post hoc. Study populations must be divided into clinically meaningful groups — this step requires input from clinicians. In our studies of bladder cancer and lung cancer patterns of care[13,16], groups were assigned by presumed treatment intent. Using clinical insight and expertise, we defined any chemotherapy or radiotherapy within 16 weeks after surgery as adjuvant therapy and assumed that any therapy started beyond that time point was palliative treatment initiated because of new evidence of disease progression. In a parallel approach, we also analysed the distribution of values to determine whether a logical cut-off point was apparent. If we made the cut-off time point too soon after surgery (such as 6 weeks), we would risk excluding some patients who received adjuvant treatment, whereas a later threshold (for example, 32 weeks) would result in the inclusion of some patients treated with chemotherapy for early metastatic disease. No approach to this type of problem is perfect, and indirect inferences about the intent of treatment should always be approached with caution. A report that explored cystectomy in patients with advanced-stage bladder cancer probably made false conclusions because the use of the derived variable disease stage at the time of treatment led to the misclassification of patients with an earlier disease stage (treated with curative intent) as having cystectomy with palliative intent[17,18]. Some data sets will be of lower quality than others, for example, the indirect measures of comorbidity generated through the use of hospital diagnostic codes. Investigators in the UK have proposed frameworks to evaluate the quality of RWD, which address many of these concepts[19,20]. Another interesting initiative is the publication in 2018 in *JAMA Surgery* of a series of articles aiming to elevate the science of surgical database research[21–24].

Using real-world data about patients

Burden of disease. Cancer registries are repositories for data that enable the societal burden of cancer to be described through population-level estimates of incidence and mortality. This information forms the cornerstone of any cancer plan to guide initiatives in prevention, screening or treatment; shifts in disease burden observed at the population level can enable health-care

Table 1 | **Illustrative examples of studies using real-world data about patients, treatments and outcomes**

Domain	Theme	Data	Key findings	Refs
Using RWD about patients				
Burden of disease	Global cancer transitions	Registry-based data of cancer incidence in 184 countries (worldwide)	In countries in economic transition, a shift from cancers associated with low-and-middle-income settings (such as cervical or gastric cancer) towards cancers associated with high-income settings (such as breast cancer, prostate cancer or CRC) was observed	25
Routine patients	Differences between trials and routine practice	Multi-site registry of 438 patients with metastatic renal cell cancer (USA)	• Patients in routine practice are less fit and older than those enrolled in clinical trials • 39% of patients in routine practice would not be eligible for relevant trials	30
Symptom burden	Symptoms at the end of life	Registry-based study of patient-reported symptoms among 10,752 patients who died of cancer (Canada)	• Pain, nausea, anxiety and/or depression remained stable in the last 6 months of life, but dyspnoea, fatigue and well-being worsened • One-third of patients had moderate-to-severe symptoms in the last month of life	31
Disease biology	HPV-related oral cancers	Registry-based study of temporal trends in HPV causality in 45,769 patients with oral cancer (USA)	• Marked increase in HPV-related oral cancer seen in younger men • HPV-related disease associated with better survival outcomes than non-HPV-related disease	27
Using RWD about treatments: access to care				
Underutilization	Use of curative-intent surgery	Registry-based study of treatment among 2,445 patients with bladder cancer (Netherlands)	• One-third of patients did not have definitive treatment with cystectomy or radiotherapy • Lack of treatment was associated with advanced age and comorbidity	34
Overutilization	Radiotherapy fractionation for bone metastases	Registry-based study of palliative radiotherapy delivery to bone metastases in 80,899 patients (Canada)	Despite evidence and guidelines supporting the use of single-fraction radiotherapy for bone metastases, more than one-half of patients continued to receive longer treatment courses	60
Referral patterns	Referral for adjuvant chemotherapy	Registry-based study of medical oncology referral patterns among 3,354 patients with resected NSCLC (Canada)	• Only half of all patients with resected NSCLC were referred to medical oncology • Lack of referral (especially among patients older than 70 years of age) is a major barrier to the use of adjuvant chemotherapy	39
Impaired access	Disparities in treatment	Registry-based study of 9,015 women with stage I–stage III breast cancer (New Zealand)	Female Māori and Pacific Islanders were less likely to receive adjuvant systemic therapy and experienced higher mortality than others	47
Country analysis	Global access to opioids	Country level reported per capita morphine consumption from International Narcotics Control Board (worldwide)	• Striking differences in access to and/or use of opioids • Worldwide, two-thirds of the population has virtually no access, and only 8% of the population has adequate access	50
Waiting times	Waiting times for radiotherapy	Registry-based study of 74,444 patients treated with radiotherapy (UK)	• Substantial differences in waiting time for radiotherapy across geographical districts (median wait time range is from 42 to 65 days) • Wait times increased over time	52
Benchmarking	Estimating the need for palliative radiotherapy	Registry-based study of geographical and system-level factors associated with the delivery of palliative radiotherapy (Canada)	One-third of patients who die from cancer need palliative radiotherapy, but a substantial proportion of them are never treated	148
Using RWD about treatments: quality of care				
Guideline concordance	Surgical treatment of CRC	Registry-based study of practice patterns among 3,095 patients with CRC (Australia)	• Median guideline concordance rate was 67% • Patient age was inversely associated with guideline concordance • Surgeon case volume was positively associated with improved measures of perioperative care	72
Integrity of treatment delivery	Radiotherapy delivery in low-resource settings	Hospital-registry description of radiotherapy delivery and factors associated with treatment completion among 1,975 consecutive patients (India)	• Only one-half of patients planned for treatment actually received surgery, radiotherapy and/or chemotherapy • Among patients who started treatment, only two-thirds completed the planned treatment course	74
Treatment abandonment	Paediatric cancer treatment abandonment	Registry-based study of treatment refusal or abandonment rate in 1,789 patients with paediatric cancer (Guatemala)	• Rates of treatment abandonment decreased over time (27% in 2001 and 7% in 2008) following implementation of an integrated health-care plan • Greater distance to centre and younger age were associated with higher rates of treatment abandonment • Treatment abandonment was associated with inferior survival	76

Table 1 (cont.) | **Illustrative examples of studies using real-world data about patients, treatments and outcomes**

Domain	Theme	Data	Key findings	Refs
Using RWD about outcomes				
Rare disease	Thymoma	Registry-based descriptive study of 668 patients with thymoma (Sweden)	• Thymoma is associated with autoimmune disease and other cancers • Survival of patients with thymoma is improving over time, but patients have inferior survival compared with population-matched controls	77
Rare patient populations	Cancer during pregnancy	Registry-based study of 1,798 patients diagnosed with cancer during pregnancy (Australia)	• Cancer incidence during pregnancy increased over time • Most common cancers in pregnant women are melanoma and breast cancer • Pregnant women with cancer had higher than average rates of early delivery and caesarean section	78
Rare outcomes	Stroke among patients with head and neck cancer	Insurance database study of 13,390 patients with head and neck cancer (Taiwan)	• Patients with head and neck cancer have a 1.44-fold increased risk of stroke compared with that of population-matched controls • The relative risk of stroke is highest in patients younger than 40 years of age and is positively associated with the receipt of combined chemoradiotherapy	79
Outcome reality check	Real-world outcomes of patients with bladder cancer	Registry-based study of 5,259 patients with bladder cancer treated with curative-intent surgery and/or radiotherapy (Canada)	• 5-year overall survival (~30%) in routine practice is lower than that observed in single-centre case series and clinical trials (~50–66%) • Quality improvement initiatives are needed	90
Systems of care	Complex surgery	Implementation of a regional model of hepatobiliary surgery care across the Veterans Health Administration health-care system (USA)	Establishment of a designated surgical programme led to increased referral, increased multidisciplinary assessment, increased treatment delivery rates and reduced levels of postoperative morbidity	92
Comparative patient outcomes	Outcomes across Europe	Registry-based study of cancer outcomes across 3 million patients in 23 European countries	Marked differences in survival across countries, with the highest survival generally noted in northern European countries and lowest survival in eastern European countries	97
Cancer economics	Costs of breast cancer care	Registry-based study of 12,580 patients with breast cancer (Italy)	• Quantification of costs of breast cancer diagnosis (€414), treatment (€8,780) and follow-up assessments (€2,351) in routine practice • Substantial variation in costs was observed between hospitals	84
Survivorship	Educational outcomes of survivors of paediatric cancer	Registry-based comparison of the educational outcomes of 782 paediatric cancer survivors with those of the general population (Canada)	Identification of domains of poor educational outcomes across various paediatric cancers and treatment modalities	10

CRC, colorectal cancer; HPV, human papilloma virus; NSCLC, non-small-cell lung cancer; RWD, real-world data.

systems to anticipate future demands. Cancer registries have also been used to describe epidemiological transitions in low-or-middle-income countries (LMICs) as disease burden shifts from communicable to non-communicable diseases; this change is accompanied by a parallel shift from cancers with a higher incidence in less-affluent populations (such as cervical and head and neck carcinomas) towards cancers with increased incidence in affluent populations (such as breast, colorectal or prostate cancer)[25]. RWD studies offer insights into shifting biological causes (for example, in cancers caused by viral infection) that cannot be learned from clinical trials[26,27]. Incidence data from cancer registries can also inform on the effectiveness of prevention and screening programmes (for example, widespread hepatitis B vaccination and cervical cancer screening). Finally, cancer registry data can enable the identification of communities with a disproportionate burden of cancer[28].

In the past few years, RWD studies have gained attention in the clinical literature; however, for decades, epidemiologists have been formulating hypotheses on cancer aetiology by carefully studying registry data. Epidemiologists have developed widely used methodological approaches to using such data, including case–control studies and cohort studies. In these studies, geographical variation, studies of migrant individuals or time trends are used to make causal inferences. Observations using RWD have contributed to a better understanding of the roles of smoking, diet and obesity in cancer causation. Thus, clinicians who analyse RWD should engage with epidemiologists to prevent methodological pitfalls related to data quality and biases that are inherent to these data sets. An optimal multidisciplinary research team includes clinicians (to ensure that the right questions are asked and that the lumping and splitting of data has clinical validity) and experts in epidemiology and biostatistics methods (to ensure that accurate answers are obtained for the clinical questions).

Patients in routine practice. Cancer registry-based studies have the potential to provide information about case distribution in an entire patient population — specifically about the demographics (age, sex and geography), comorbidity and symptom burden in patients.

Notably, the patient populations included in clinical trials are usually not representative of those seen in routine practice[29]. Knowing that patients routinely seen by doctors tend to be older and to have a poor health status compared with patients included in randomized controlled trials (RCTs)[30] is important for clinicians in treatment decision-making and for policymakers in interpreting data on utilization rates and outcomes. For example, RWD have enabled providers to understand the symptom burden of patients at the end of life using tools such as the Edmonton Symptom Assessment System (ESAS)[31]; such real-world data sets contain important information for system-level planning of palliative care services.

Using real-world data about treatments

Access to care. The extent to which patients access cancer services can be determined using RWD. Several studies have demonstrated underutilization of radiotherapy, surgery and chemotherapy for many cancers[13,32–34]. Physician referral data can enable investigators to track the management history of a patient by identifying consultations with specialists[35,36]. For example, the low utilization rates of (neo)adjuvant chemotherapy for various cancers has been described in many studies[37,38]. Moreover, in the past few years, several reports have documented the extent to which this problem is driven by upstream decision-making by the surgeon (owing to non-referral to a medical oncologist) or at the level of the patient–medical oncologist conversation[35,39]. When two potential treatment options are available (for example, definitive surgery or radiotherapy for cancer of the prostate or bladder), RWD can enable tracking of whether the patient met with both a surgeon and radiation oncologist to make an informed choice[40].

As already discussed, patient subpopulations with impaired access to care and a disproportionate disease burden can be identified using RWD. Subsequently, this information can be used to design targeted programmes to improve care among those patients who need it the most. An ecological measure of socio-economic status (for example, neighbourhood income) has frequently been used to demonstrate that disease burden, treatment and outcome are independently associated with socio-economic status[28,41,42]. A study from the early 2000s in Ontario demonstrated a gradient in access to palliative radiotherapy across socio-economic groups[43]. Other studies have illustrated substantial differences in care between rural and urban communities[44,45]. RWD have also facilitated studies that identify differences in the utilization of health-care services and in outcome across ethnic groups[46,47]. Striking differences in treatment rates across age groups were evidenced in studies of practice patterns[48]. Finally, the results of regional analyses can illustrate inequity in access to anticancer therapy[49,50]. This information can be used to improve access to care by locating new clinical services in regions identified as having impaired access[51].

The timeliness of cancer diagnosis and therapy is commonly described using RWD[52,53] and, thus, modifiable system-level barriers responsible for delays can be addressed with interventions[54]. RWD studies of timeliness of care have offered insights into the association between delays, stage of disease and inferior outcome[55,56]. In a cohort of 1,795 women with breast cancer across 5 countries in Africa, McKenzie et al. identified social factors associated with advanced stage of disease at the time of diagnosis[56]. This information can be used by public health authorities to design targeted educational and outreach activities to facilitate timely diagnosis.

The use of RWD can reveal underutilization and overutilization of anticancer therapies and diagnostic tests relative to practice guidelines. Results of a study published in 2015 showed marked overutilization of diagnostic and surveillance imaging among women with breast cancer[57,58]; the consequences of overutilization are usually further clinical investigations that can lead to increased costs and delayed treatment[59]. Another aspect highlighted by studies using RWD is that the delivery of radiotherapy in routine practice often includes more fractions than are recommended in practice guidelines[60,61]. These studies address topics on the American and Canadian Choosing Wisely lists[62,63]. In these lists, specialist organizations identify ten commonly performed interventions that offer low value to patients. The goal of this initiative is to reduce unnecessary medical tests, treatments and procedures. The use of RWD can also facilitate quality improvement to evidence-based care[64]. While the aforementioned examples compare the adherence of real-world clinical practice with guidelines, the identification of the optimal utilization of imaging or treatment modalities is challenging. Guidelines recommend standard treatment options on the basis of data from RCTs but do not take into account outcomes in patients who are not eligible for treatment and/or choose not to undergo a therapy. Evidence-based approaches for estimating the need for anticancer treatment have been developed[65] and are used widely in health-care system planning[66] but might lead to an overestimation of the percentage of patients who receive appropriate (guideline-compliant) treatment[67]. Criteria-based benchmarking is an empirical method that uses RWD to estimate the appropriate rate of use of a specific therapy in routine practice[68]; this approach is used by health-care system managers for planning and monitoring (for example, of radiotherapy delivery in Ontario)[69].

Quality of care. RWD can offer insights into the quality of decision-making and how decisions are implemented. Studies that describe patterns of care facilitate measurement of performance compared with guideline recommendations[70–72]; investigators can also identify factors associated with the underutilization of recommended therapies. RWD can enable the identification of variations in practice. Potential drivers of practice variability can relate to equity, efficiency and clinical uncertainty, and RWD can help to identify questions for future RCTs. If practice variation is observed in a setting in which a superior treatment option is available, RWD can inform initiatives to improve quality. If two treatment options are associated with similar survival outcomes, RWD can inform policymakers in planning initiatives to support the option associated with less morbidity or reduced costs.

The availability of detailed treatment records makes it possible to describe the integrity of treatment delivery in routine practice. Records of chemotherapy or radiation administration can describe dose intensity and completion rates that can be compared with those recommended in RCTs and practice guidelines[73,74]. Pathology records can offer insight into the quality of surgical care and provide data to drive quality improvement[75]. Porter et al.[75] used a cancer registry and linked treatment records to evaluate the extent to which an audit and feedback programme led to improved lymph node harvest among patients with colorectal cancer. In LMICs, failure to complete treatment is a major problem for a substantial proportion of patients; RWD can be used to describe the extent of incomplete treatment and offer insights into modifiable factors[76].

Using real-world data about outcomes

Augmenting knowledge from clinical trials. RWD can offer important insights into rare diseases for which large RCTs have not been conducted[77], rare populations of patients (such as those with cancer during pregnancy)[78] and rare events (such as certain toxicities) that are not identified in the usually short-term follow-up study of patients included in clinical trials[79]. One of the greatest strengths of RWD is the ability to improve the generalizability of conclusions from clinical research and to contribute knowledge by enabling the study of larger cohorts and/or by allowing longer follow-up durations than those typically used in RCTs. This information is helpful for clinicians and patients when considering treatment options and long-term prognosis among patients who are under-represented in clinical trials (the most common example being elderly patients)[80]. Even among patients with common diseases, the incidence of less common or delayed toxicities (for example, a cardiac event after breast radiotherapy) can be better described using RWD than data from RCTs[81,82].

Real-world outcomes: a reality check. RWD can be used to document the outcomes of patients with common cancers in routine practice, thereby providing a fundamental evaluation of quality of care. While cancer registries have traditionally contained limited information (mostly on survival outcomes), contemporary sources of RWD enable investigators to describe other outcomes, including patient symptoms[9], treatment-derived toxicities and/or complications[79], organ preservation rates[83] and costs of care[84]. Reporting of short-term outcomes (for example, 90-day postoperative mortality[85], readmission rates or mortality from chemotherapy-derived toxicities) can enable institutions to establish a benchmark from which quality can be improved[86]. These metrics also enable health-care providers to compare the outcomes of their patients with those reported in the published literature and/or patients treated by their peers. Longer-term outcomes (for example, median survival of patients with advanced-stage disease or 5-year survival of those with early stage disease) can be compared with results reported in clinical trials and/or case series from centres of excellence. A large gap between expected and observed outcomes would suggest the need for quality improvement initiatives to be implemented[87,88]. Such differences, however, also occur because survival estimated using the results of RCTs — which is often presented to patients in routine clinical care despite trials including patients with a better performance status — might not be applicable[89]. For example, 5-year overall survival among 3,879 patients with bladder cancer treated with cystectomy in Ontario was found to be 36%, a value substantially inferior to the 66% reported for 1,054 patients operated on at a leading centre in the USA[90,91]. In addition to survival, other traditional outcomes that can be evaluated with RWD include organ preservation and toxicity, whereas emerging outcomes include patient-reported outcomes and economic costs in routine practice.

System-level perspectives. RWD can inform system-level analyses of quality of care in the context of poor outcomes, which are distinct from studies of the comparative effectiveness of two different treatment strategies. Using RWD to identify variability in outcomes and to understand the causes thereof is fundamentally different from using RWD to explore differences in the outcomes of two different treatments. For example, the identification of high postoperative mortality and readmission rates following complex surgery has led to system-level changes resulting in improved outcomes[92,93]. After initiating a dedicated regional hepatobiliary surgical programme within the Veterans Affairs system, Lau et al.[92] identified a substantial increase in regional case volumes together with improved processes of care (for example, increased use of surgical resection and chemotherapy) and outcomes (such as reduced postoperative complications). RWD have also been used to describe the association between the number of patients treated by a surgeon and/or hospital volume and the outcomes of complex surgical procedures[85,94] and have been used in many jurisdictions (such as Ontario and the Netherlands) to regionalize complex surgery and thereby improve quality of care[95,96].

Comparisons of outcome can also be made across borders to provide a benchmark for achievable outcomes. The EUROCARE-4 study identified marked differences in cancer outcomes across 23 countries in Europe and showed a relationship between health-care expenditures and outcomes[97]. This pivotal study prompted major efforts to change cancer care systems and focused investments in several countries[98]. RWD can also be used to set a benchmark for emerging health-care systems in LMICs, who can compare their outcomes with those observed in high-income countries to measure the gap between current and potentially achievable outcomes[99–101].

The use of RWD can enable estimates of the costs of therapy and disease management at the societal level[84,102] and can inform societal decisions about how to optimally allocate resources for the screening, prevention and treatment of cancer[103–105]. The improvement of outcomes at the population level requires the identification of potentially modifiable factors associated with suboptimal access and/or quality of care. We propose a general framework to design any programme in health-care

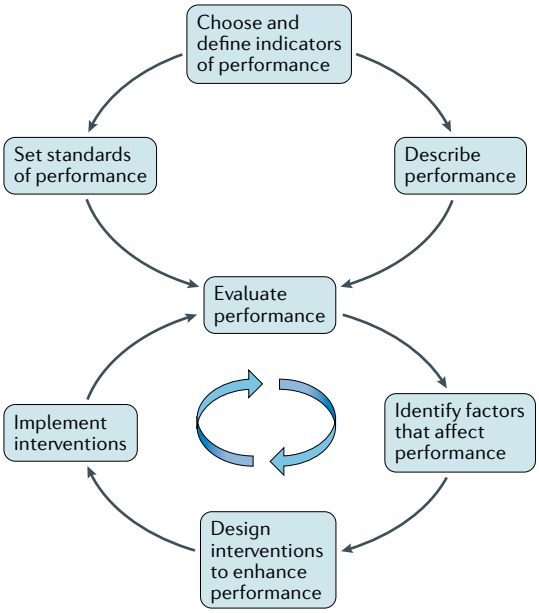

Fig. 2 | **Conceptual framework depicting a general scheme for research on health-care system performance.** The first step is to select, define and validate appropriate indicators of the aspect of performance that has been targeted for investigation. The next two steps are to develop methods to measure system performance in terms of the chosen indicators and to prescribe standards or targets for system performance in terms of the chosen indicators. Once such standards have been set, and methods for measuring performance have been established and validated, the performance of the health-care system against the standards can be evaluated. This process, in turn, enables conducting further explanatory studies aimed at identifying factors that are associated with better or worse performance. The information obtained can be used to design interventions aimed at improving performance. Figure adapted with permission from REF.[106], Elsevier.

services research aimed at improving a specific aspect related to performance[106] (FIG. 2).

Comparative effectiveness

Randomized trials and real-world data. RCTs and studies of RWD are very different forms of research that have different goals; while the outcomes of RCTs define what is achievable for patients with cancer, the outcomes observed analysing RWD define what is currently achieved. Thus, the interface of RCTs and RWD can help systems move towards achieving the achievable[106]. With RWD, many issues of cancer control that cannot be addressed with an RCT can be explored — RWD studies should not be viewed as alternatives to RCTs. This principle needs to be carefully considered in light of the interest observed in the past few years towards using RWD for the purposes of regulatory approval[107]. When RWD are used to compare the effectiveness of two anticancer therapies, the studies should be regarded as hypothesis-generating only; if the study is poorly designed, its results could potentially be harmful to patients[18,108–110].

The strength of RCTs rests with the internal validity of the data, but their weakness lies in their limited generalizability, because trial populations are often not representative of the general population at large. Owing in part to strict eligibility criteria, <5% of adult patients with cancer in the USA participate in a clinical trial[111]. However, no reason exists for clinical trials not to be more representative of the general population. The paediatric oncology community has a long history of undertaking trials in which ~60% of all children with cancer are treated on protocol[112]. Clinical trials for adults with cancer should adopt broader entry criteria, as proposed by Peto and colleagues[113]; several large, simple trials conducted in the UK (such as ICON4/AGO-OVAR-2.2 and STAMPEDE) allowed the inclusion of all patients considered fit enough to receive the treatment, with consequent wider application[114,115]. Therefore, RWD should not be expected to fix the generalizability problem of RCTs, but, instead, RCTs need to be made more applicable to patients in the real world. Finally, that RCTs (a scientific methodology) and studies of RWD (a data source) are not mutually exclusive is worth highlighting. While RWD are most often used to study practice and outcomes after (or in the absence of) an RCT, increasing interest is being focused on implementing RCTs in real-world settings and using data from EHRs and administrative records to evaluate the toxicity and outcomes of the compared treatments[116,117].

Real-world data studies of comparative effectiveness. The comparison of outcomes of nonrandomized groups of patients who have received different treatments in the real world is problematic, primarily because the groups of patients might differ with respect to prognostic factors. This aspect is as prevalent in current studies using RWD as it was 50 years ago when investigators used institutional data to evaluate treatment effectiveness. Population-based studies might have greater external validity than that of institution-based studies, but they have a similar internal validity, and, therefore, both are classified as providing level 3 evidence according to Sackett's hierarchy of evidence[118,119].

Investigators have three general approaches to mitigate selection bias in observational studies using RWD. In the first approach, the investigator can undertake multivariate analyses to adjust for potential confounders; the downside to this approach is that one can adjust only for variables that are known and can be measured.

In the second approach, the investigator can apply a propensity score, which is the estimated probability that a patient will receive a treatment given the values of all covariates that can affect the choice of treatment. Patients can be stratified into subgroups defined by the propensity score, and the outcomes of these subgroups are compared. The method can decrease confounding because known prognostic characteristics should be randomly distributed within each stratum between patients who did and did not receive the intervention. Despite the increasing use of propensity score analyses, the extent to which this approach provides a less biased output than that of traditional Cox models is modest[120]. Both approaches are limited because the factors associated strongly with both treatment selection and outcome include stage of disease (data on which are

either missing or unreliable in many cancer registries), performance status (data on which are almost universally missing) and comorbidities. Efforts to control for the latter rely on measures such as the Charlson and/or Elixhauser comorbidity scales, which have been adapted for use with administrative data sets; however, these instruments have limited predictive ability[121].

The third approach to reducing selection bias is instrumental variable analysis (IVA), in which patients are grouped by a marker for different practice policies (usually, a temporal period or geographical region)[122]. This quasi-experimental design assumes that the underlying population and other treatments are comparable across time periods or geographical regions, and, therefore, any difference in outcome between two periods or regions can be attributed to the differences in therapy. IVA avoids the risk of selection bias and is therefore superior to multivariate or propensity score methodologies[122]. Our group has exploited temporal and/or geographical differences in practice to evaluate the uptake of adjuvant chemotherapy for lung cancer[16], cisplatin concurrent with radiotherapy for cervical cancer[123] and definitive management of head and neck cancer[124]. We described time-related adoption of adjuvant chemotherapy for NSCLC (2001–2006) and of chemoradiotherapy for cervical cancer (1992–2001). During the years of study, the case distribution and other elements of care did not change. We used a before-and-after design to show that the survival of treated patients (at the population level) improved with adoption of the novel therapy, with a magnitude of benefit that was comparable to that observed in RCTs. In the study of head and neck cancer, we also exploited geographical differences in treatment between the USA (radical surgery) and Ontario (radiotherapy)[124]. In both regions, the survival across each region was remarkably similar, with a higher rate of larynx preservation in Ontario. This type of study relies on the assumption that case distribution, other elements of care and disease biology are comparable across time periods and/or regions. Our study of the adoption of chemoradiotherapy for head and neck cancer provides an example in which this assumption failed. Following the adoption of concurrent treatment with cisplatin, the observed improvement in survival in patients with oropharyngeal cancer was largely attributable to a simultaneous increase in the incidence of the more radiosensitive, human papilloma virus (HPV)-related disease subtype, rather than to the change in therapy[125].

The validity of observational studies of comparative effectiveness relates to the pretest probability that an observed association is valid. In this regard, we propose a hierarchy of scenarios for analyses of comparative effectiveness. At the top of this hierarchy are those studies seeking to measure whether the efficacy observed in a pivotal RCT translates into effectiveness in routine practice[16,123]. While these studies remain prone to bias, if a treatment effect emerges in routine practice, the confidence in this effect being real is high because level 1 RCT evidence demonstrates that the treatment works. Following publication of pivotal RCTs of adjuvant chemotherapy in patients with NSCLC, our group described the uptake and outcomes associated with such treatments, which we found to be associated with a survival benefit in the general population[126].

The second scenario is that in which data other than level 1 evidence support efficacy. In this case, the aim is typically to seek insight into an intervention for which a growing body of literature suggests efficacy. For example, our work related to adjuvant chemotherapy for bladder cancer added to knowledge from a series of small RCTs that suggested survival benefit with cisplatin-based therapy[13].

In the third scenario, a clinical trial has failed to show efficacy, and no evidence suggests benefit from treatment. If an observational study suggests benefit in this setting, the pretest probability of a type 1 error is so high that the conclusion should be regarded with scepticism. This phenomenon has been demonstrated by several studies using the US National Cancer Database (NCDB). RCTs have clearly demonstrated, for example, that the survival benefit of adjuvant chemotherapy in patients with stage II colon cancer is very modest and that radiotherapy improves local disease control but not overall survival in patients with rectal cancer[127,128]. In contrast with the results of RCTs, two studies using the NCDB and published in 2016 and 2017 found implausibly large survival benefits in patients with low-risk stage II colon cancer receiving adjuvant chemotherapy compared with those in patients who did not receive adjuvant chemotherapy and similarly large benefits in patients with locally advanced rectal cancer completing neoadjuvant radiotherapy compared with those in patients who did not complete treatment[129,130]. The observed benefits are more likely to be caused by residual confounding, immortal time bias and measurement errors than by these treatments being more effective in routine practice than in RCTs. In another study using the NCDB, a very large improvement in overall survival was observed among patients with metastatic bladder cancer who underwent radical cystectomy; despite a lack of RCT data to refute these findings, the study results are vulnerable to many forms of bias (such as selection bias, misclassification bias and immortal time bias) and are likely misleading and potentially dangerous because they support an aggressive therapeutic strategy[17,18]. These studies are striking examples of how RWD can misinform policy and practice in studies of comparative effectiveness.

Quality of treatment-effectiveness studies. Grading the quality of clinical evidence on the basis of study design only is a simplistic approach; the validity of observational studies varies widely, depending on the details of their implementation and analysis[131]. Epidemiologists, who can rarely perform RCTs to test hypotheses about disease causation, rely heavily on observational data and have developed methods for evaluating the quality of such data. In a systematic review, Sanderson et al.[132] identified 86 different tools for assessing the quality of observational studies. The content of the tools was evaluated for domains subject to bias using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines[133]. The authors were unable to identify any single tool that was optimal but concluded that three fundamental domains must be evaluated: appropriate selection

of participants, appropriate measurement of exposure and outcome variables and appropriate control of confounding. Other design-specific sources of bias should be evaluated. Many of the tools identified by Sanderson and collaborators[123] provide numerical summary scores; the authors concluded that these are inappropriate because the process of generating summary scores involves weighting of the component elements, which is a variable and inconsistent process. Thus, these investigators recommended evaluating observational studies using a simple checklist that "concentrates on the few, principal, potential sources of bias in a study's findings"[123].

The approach proposed by Sanderson et al. is also useful in evaluating the quality of population-based studies that compare outcomes between cohorts of patients who have received different treatments. We provide a checklist of questions that should be addressed before reaching a conclusion about the validity of a population-based outcome study in oncology (BOX 2). This framework is based partly on the Newcastle–Ottawa Scale for evaluating the quality of nonrandomized studies[134,135] but has been adapted for the qualitative evaluation of cancer-registry-based outcome studies and does not yield a summary score. In creating this checklist, we adhere to the broad recommendations of Sanderson et al., who proposed that tools should "1) include a small number of key domains; 2) be as specific as possible (with consideration of study design and topic); 3) be a simple checklist rather than a scale; and 4) show evidence of careful development, and of validity and reliability". The first three aspects are further discussed herein, and future work is necessary for the last item.

To avoid design-specific sources of bias in population-based studies, the analytical strategy used must, as far as possible, follow the logic of a clinical trial. Investigators of population-based outcome studies usually simultaneously receive all the data that they will use to identify the study population, assign patients to treatment categories

Box 2 | **Checklist for the evaluation and interpretation of population-based studies of treatment effectiveness**

Selection of the study population

- Was the study population clearly and unambiguously defined?
- Was the study population representative of the overall population of interest?
- Was the study population homogeneous enough to enable general conclusions about the effects of treatment to be made?
- How complete were the fields, and how accurate were the variables used to identify the study population? Was this information estimated or audited?
- Was it possible for ineligible patients to be included in the study population or for eligible patients to be excluded? How frequently did inclusion and/or exclusion occur? Was this information estimated or audited?

Identification of treatment subgroups (for example, according to exposure)

- Were the treatment subgroups clearly and unambiguously defined?
- How complete were the fields, and how accurate were the variables used to identify the treatment groups? Was this information estimated or audited?
- Is there any risk of misclassification of treatment? What is the risk of misclassification? Was this information estimated or audited?

Measurement of outcomes

- Were the outcome measures clearly and unambiguously defined?
- Were the outcome measures appropriate for evaluating the treatment of interest?
- How complete were the fields, and how accurate were the variables used to measure outcomes? Was this information estimated or audited?
- How complete was the follow-up information? Was the follow-up duration long enough?
- Was follow-up monitoring completed equally among the treatment groups?

Temporal relationship between creation of the study population, assignment to treatment groups and evaluation of outcomes

- Was the study population defined entirely on the basis of information that was available before treatment was initiated?
- If not, could the treatment itself have modified eligibility for inclusion in the study population? Did each treatment that is to be compared have a similar effect on eligibility for inclusion in the study population, or was there a differential effect[a]?
- Did evaluation of the outcome of interest begin only after the treatment groups had been defined?

Comparability of the treatment groups

- Were the major prognostic factors known to be associated with the outcome of interest measured and reported?
- Were those measures sufficiently sensitive and specific (for example, comorbidities)?
- How complete were the fields, and how accurate were the variables used to measure prognostic factors? Was this information estimated or audited?
- Were the treatment groups balanced with respect to major prognostic factors?

Comparison of the outcomes between treatment groups

- Were any differences in the characteristics of the treatment groups controlled for in the analysis?
- Were continuous variables (for example, age) compressed into inappropriately broad categories?
- What general methodological approach was used to evaluate the treatment effect (for example, traditional adjusted analyses, propensity score analysis or instrumental variable analysis)?

Interpretation[b]

- Has the treatment comparison been previously addressed in a randomized controlled trial (RCT)? If yes, were superior outcomes observed in the treatment arm over the control arm?
- How does the observed magnitude of the treatment effect compare with the effect observed in RCTs?
- Does the observed treatment effect satisfy Hill's criteria for causation[128]? Are the results biologically plausible? Are the results consistent with the existing knowledge? Is the treatment effect specific to the primary outcome of interest?

[a]If the answer to this question is yes, the treatment assignment might have a differential effect on the composition of the treatment groups. If the information available after treatment changes the eligibility of some patients, these patients will be allocated to the wrong treatment groups. [b]If pre-existing data from RCTs have shown that the treatment lacks efficacy, any effect observed in routine practice should be viewed sceptically. Likewise, if RCT evidence supports delivery of the treatment, a larger magnitude of effect observed in routine practice than that observed in RCTs should be interpreted with awareness regarding residual confounding and/or bias.

and evaluate outcomes. The analysis must proceed in a logical temporal order: treatment should not affect the eligibility of patients for inclusion in a study, and outcomes should not influence their assignment to a particular treatment group. The study population must therefore be defined and locked before treatment groups are assigned, and treatment groups must be defined and locked before assessment of the outcome. This method sounds easy to implement, but treatment assignment can affect eligibility for inclusion in a study if, as is often the case, cancer stage is an eligibility criterion. This problem is largely a consequence of registries recording only collaborative-stage results, which incorporate both surgical and clinical findings of patients who undergo cancer surgery but only clinical information of those who do not undergo surgery. This approach results in a systematic up-staging of patients who have surgery and invalidates comparison of the outcomes between treatments that include surgery and those that do not. Stage details can also be modified in some registries by incorporating the results of investigations conducted after treatment initiation (such as a postoperative staging bone scan). The risk of patient outcomes affecting the treatment group to which the patients are assigned is easier to deal with, if recognized in advance. Consider a study in which the main outcome is survival and patients are assigned to two treatment groups according to whether they have undergone a given intervention within 3 months of diagnosis or not. Those who die within 3 months of diagnosis are assigned automatically to the no intervention group; if survival is measured from the date of diagnosis, the values recorded from these patients will lower the average survival of that group. This bias is defined as immortal time bias[136] because patients in the intervention group are at no risk of death until they undergo treatment, while other patients are at risk of death from the time of diagnosis. In RCTs, similar situations are dealt with by analysing outcomes in the intention-to-treat population, but this approach is not an option in observational studies because treatment intentions are unknown and treatment groups can be defined only by the treatment actually received. In the above example, however, this problem can be avoided simply by starting measurement of outcomes after 3 months and ignoring earlier deaths.

In evaluating claims that a given treatment is effective on the basis of observational data, clinicians should emulate the epidemiologists' approach to evaluating claims that a given exposure causes a particular disease. In studies using RWD, Bradford Hill's framework for establishing causation can guide the evaluation of treatment effects[137]. These criteria for causation can be applied to the totality of available evidence and include considerations of the strength of the association and its consistency across studies (strong and consistent associations are less likely to occur by chance or owing to confounding)[138]; the presence of a dose–response relationship; the specificity of exposure or treatment being associated with the expected outcome and not with other outcomes; and the plausibility of the effect and its coherence with related scientific knowledge of the treatment and disease settings. A lack of knowledge and understanding of mechanisms can limit plausibility and coherence and are not necessary for attributing causation but, if present, can increase confidence in the result[131,138]. We include questions based on Hill's criteria to assist readers in evaluating the results of a population-based outcome study and in determining the likelihood that one treatment is more effective than another on the basis of the totality of the available evidence (BOX 2). The authors of three commentaries published in the past 2 years have highlighted the methodological pitfalls inherent to specific studies using RWD that reported very large effect sizes almost certainly caused by confounding and bias[18,109,111].

Regulatory aspects

RWD are increasingly used to support regulatory drug approval and funding decisions. For regulatory approval purposes, in many countries, RWD are viewed as a supplement to data from clinical trials[139,140], and, thus, drug approval decisions are not taken solely on the basis of RWD. As part of the 21st Century Cures Act, the FDA will be required to develop a framework and guidance for the use of RWD in the drug approval process[141,142]. Indeed, the FDA has formed a partnership with CancerLinQ, an initiative of ASCO, to coordinate the sharing of patient data and incorporate that information into the FDA regulatory process[143].

The increasing focus on value-based health care has stimulated interest in RWD as a tool to assist with reimbursement decisions. The authorities that conduct health technology assessments (HTAs) to evaluate new therapeutic agents are demanding RWD; these authorities are seeking evidence that clinical interventions provide meaningful benefits (improved survival, quality of life and cost) to patients[144]. RWD can be used after regulatory approval in discussions of drug reimbursement, including conditional reimbursement schemes, and for pharmaco-economic analyses[139].

At the time of initial HTAs, uncertainties often exist as to whether the efficacy demonstrated in phase III clinical trials will translate into effectiveness when the drug is used in routine practice. A common use of RWD after regulatory approval is to define conditional reimbursement schemes or managed-entry agreements. These agreements include risk sharing (between payers and industry), payment for outcomes, performance-based reimbursement schemes and various forms of coverage dependent on the capture of RWD for evidence development[145].

In the UK, the National Institute for Health and Care Excellence (NICE) initially rejected the reimbursement of bortezomib for the treatment of multiple myeloma but eventually made a deal with the manufacturer to reimburse the National Health Service for the cost of the drug for those patients who have only a limited response or no response to therapy. Other European countries have adopted similar risk-sharing and managed-entry agreements for oncology drugs[146]. In Japan, the approval of all therapeutic agents is temporary and subject to review 4–10 years after initial approval[147]. On the basis of this review, the initial approval for the drug can be cancelled, the criteria for using the drug can be modified or, otherwise, no action is taken. A critical issue that needs

to be carefully considered in the future is which group will be responsible for funding and conducting RWD studies for the purposes of regulators and payers (that is, will it be the sponsoring pharmaceutical company or an arm's-length HTA organization?).

Conclusions

The cornerstone of cancer control planning is the availability of population-based cancer registries. Existing registries require ongoing support and quality assurance, whereas emerging health-care systems should invest in a cancer registry. The linkage of claims-based data sources to cancer registries enables the identification of gaps in care and offers the potential for short-term improvement in public health by helping health-care systems to achieve the achievable. Emerging sources of RWD can offer improved granularity over traditional claims-based data sets, but in many cases their internal validity remains unknown. Ultimately, population research needs to move beyond simply describing gaps in care towards population-level intervention studies to change practice and improve outcomes. To date, this has been the greatest shortfall in the use of RWD; efforts in this regard are needed to derive the maximum societal benefit from these data sources. Studies using RWD should not be used as substitutes for RCTs: RCTs remain the gold standard for establishing the efficacy of new treatments for patients with cancer, although many clinical and policy questions are best addressed using RWD. However, it behoves the research community (including journal editors and peer reviewers) to apply a high level of insight into the quality of the data and methodology used in such studies. This rigour is critical in studies of comparative effectiveness because clinicians are less likely to be as familiar with the potential pitfalls of RWD studies as they are with those of RCTs. The widespread availability of large electronic databases presents an opportunity for the research community to gain important insights into the burden of cancer, management of the disease and patient outcomes in routine practice. By asking the right questions with the right data and the right methods, RWD has the potential to close gaps in our knowledge of cancer and to identify gaps between evidence and practice that will ultimately lead to improved patient care and outcomes.

Published online 30 January 2019

1. Goldberg, R. M., Wei, L. & Fernandez, S. The evolution of clinical trials in oncology: defining who benefits from new drugs using innovative study designs. *Oncologist* **22**, 1015–1019 (2017).
2. Sherman, R. E. et al. Real-world evidence — what is it and what can it tell us? *N. Engl. J. Med.* **375**, 2293–2297 (2016).
3. Potosky, A. L., Riley, G. F., Lubitz, J. D., Mentnech, R. M. & Kessler, L. G. Potential for cancer related health services research using a linked Medicare-tumor registry database. *Med. Care* **31**, 732–748 (1993).
4. Mackillop, W. J. et al. Waiting for radiotherapy in Ontario. *Int. J. Radiat. Oncol. Biol. Phys.* **30**, 221–228 (1994).
5. Parkin, D. M. The evolution of the population-based cancer registry. *Nat. Rev. Cancer* **6**, 603–612 (2006).
6. Friends of Cancer Research. Establishing a framework to evaluate real-world endpoints. *FOCR* https://www.focr.org/sites/default/files/pdf/RWE_FINAL%207.6.18.pdf (2018).
7. Cook, J. A. & Collins, G. S. The rise of big clinical databases. *Br. J. Surg.* **102**, e93–e101 (2015).
8. Khozin, S. et al. Characteristics of real-world metastatic non-small cell lung cancer patients treated with nivolumab and pembrolizumab during the year following approval. *Oncologist* **23**, 328–336 (2018).
9. Harrison, L. D. et al. Comparing effectiveness with efficacy: outcomes of palliative chemotherapy for non-small-cell lung cancer in routine practice. *Curr. Oncol.* **22**, 184–191 (2015).
10. Lorenzi, M. et al. Educational outcomes among survivors of childhood cancer in British Columbia, Canada: report of the Childhood/Adolescent/Young Adult Cancer Survivors (CAYACS) Program. *Cancer* **115**, 2234–2245 (2009).
11. Krzyzanowska, M. K. et al. Can chemotherapy-related acute care visits be accurately identified in administrative data? *J. Oncol. Pract.* **14**, e51–e58 (2018).
12. Satkunam, N., Wei, X., Biagi, J. J., Nanji, S. & Booth, C. M. Delivery of adjuvant oxaliplatin for colon cancer: insights from routine clinical practice. *J. Natl Compr. Canc. Netw.* **14**, 1548–1554 (2016).
13. Booth, C. M. et al. Perioperative chemotherapy for muscle-invasive bladder cancer: a population-based outcomes study. *Cancer* **120**, 1630–1638 (2014).
14. Fellegi, I. P. & Sunter, A. B. A theory of record linkage. *J. Am. Stat. Assoc.* **61**, 1183–1210 (1969).
15. Zhu, Y., Matsuyama, Y., Ohashi, Y. & Setoguchi, S. When to conduct probabilistic linkage versus deterministic linkage? A simulation study. *J. Biomed. Inform.* **56**, 80–86 (2015).
16. Booth, C. M. et al. Adoption of adjuvant chemotherapy for non-small-cell lung cancer: a population-based outcomes study. *J. Clin. Oncol.* **28**, 3472–3478 (2010).
17. Seisen, T. 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.* **34**, 3529–3536 (2016).
18. Booth, C. M. et al. Radical treatment of the primary tumor in metastatic bladder cancer: potentially dangerous findings from observational data. *J. Clin. Oncol.* **36**, 533–535 (2017).
19. Black, N. & Payne, M. Directory of clinical databases: improving and promoting their use. *Qual. Saf. Health Care* **12**, 348–352 (2003).
20. Raftery, J., Roderick, P. & Stevens, A. Potential use of routine databases in health technology assessment. *Health Technol. Assess.* **9**, 1–92 (2005).
21. Haider, A. H., Bilimoria, K. Y. & Kibbe, M. R. A checklist to elevate the science of surgical database research. *JAMA Surg.* **153**, 505–507 (2018).
22. Doll, K. M., Rademaker, A. & Sosa, J. A. Practical guide to surgical data sets: surveillance, epidemiology, and end results (SEER) database. *JAMA Surg.* **153**, 588–589 (2018).
23. Merkow, R. P., Rademaker, A. W. & Bilimoria, K. Y. Practical guide to surgical data sets: national cancer database (NCDB). *JAMA Surg.* **153**, 850–851 (2018).
24. Kaji, A. H., Rademaker, A. W. & Hyslop, T. Tips for analyzing large data sets from the JAMA surgery statistical editors. *JAMA Surg.* **153**, 508–509 (2018).
25. Bray, F., Jemal, A., Grey, N., Ferlay, J. & Forman, D. Global cancer transitions according to the Human Development Index (2008–2030): a population-based study. *Lancet Oncol.* **13**, 790–801 (2012).
26. Rabkin, C. S., Biggar, R. J. & Horm, J. W. Increasing incidence of cancers associated with the human immunodeficiency virus epidemic. *Int. J. Cancer* **47**, 692–696 (1991).
27. Chaturvedi, A. K., Engels, E. A., Anderson, W. F. & Gillison, M. L. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *J. Clin. Oncol.* **26**, 612–619 (2008).
28. Mackillop, W. J., Zhang-Salomons, J., Boyd, C. J. & Groome, P. A. Associations between community income and cancer incidence in Canada and the United States. *Cancer* **89**, 901–912 (2000).
29. Booth, C. M. & Tannock, I. F. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *Br. J. Cancer* **110**, 551–555 (2014).
30. Mitchell, A. P. et al. Clinical trial participants with metastatic renal cell carcinoma differ from patients treated in real-world practice. *J. Oncol. Pract.* **11**, 491–497 (2015).
31. Seow, H. et al. Trajectory of performance status and symptom scores for patients with cancer during the last six months of life. *J. Clin. Oncol.* **29**, 1151–1158 (2011).
32. Mackillop, W. J. et al. Does a centralized radiotherapy system provide adequate access to care? *J. Clin. Oncol.* **15**, 1261–1271 (1997).
33. Schrag, D., Cramer, L. D., Bach, P. B. & Begg, C. B. Age and adjuvant chemotherapy use after surgery for stage III colon cancer. *J. Natl Cancer Inst.* **93**, 850–857 (2001).
34. Goossens-Laan, C. A. et al. Effects of age and comorbidity on treatment and survival of patients with muscle-invasive bladder cancer. *Int. J. Cancer* **135**, 905–912 (2014).
35. Booth, C. M., Siemens, D. R., Peng, Y. & Mackillop, W. J. Patterns of referral for perioperative chemotherapy among patients with muscle-invasive bladder cancer: a population-based study. *Urol. Oncol.* **32**, 1200–1208 (2014).
36. Chandhoke, G. et al. Patterns of referral for adjuvant chemotherapy for stage II and III colon cancer: a population-based study. *Ann. Surg. Oncol.* **23**, 2529–2538 (2016).
37. Booth, C. M. et al. Perioperative chemotherapy for bladder cancer in the general population: are practice patterns finally changing? *Urol. Oncol.* **36**, 89.e13–89.e20 (2017).
38. Cuffe, S. et al. Adjuvant chemotherapy for non-small-cell lung cancer in the elderly: a population-based study in Ontario, Canada. *J. Clin. Oncol.* **30**, 1813–1821 (2012).
39. Kankesan, J. 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.* **20**, 30–37 (2013).
40. Quirt, J. S., Siemens, D. R., Zaza, K., Mackillop, W. J. & Booth, C. M. Patterns of referral to radiation oncology among patients with bladder cancer: a population-based study. *Clin. Oncol.* **29**, 171–179 (2017).
41. Kumachev, A., Trudeau, M. E. & Chan, K. K. Associations among socioeconomic status, patterns of care and outcomes in breast cancer patients in a universal health care system: Ontario's experience. *Cancer* **122**, 893–898 (2016).
42. Deb, S. et al. The effect of socioeconomic status on gross total resection, radiation therapy and overall survival in patients with gliomas. *J. Neurooncol.* **132**, 447–453 (2017).
43. Huang, J. et al. Factors affecting the use of palliative radiotherapy in Ontario. *J. Clin. Oncol.* **19**, 137–144 (2001).

44. Jacobs, L. K., Kelley, K. A., Rosson, G. D., Detrani, M. E. & Chang, D. C. Disparities in urban and rural mastectomy populations: the effects of patient- and county-level factors on likelihood of receipt of mastectomy. *Ann. Surg. Oncol.* **15**, 2644–2652 (2008).
45. Atkins, G. T., Kim, T. & Munson, J. Residence in rural areas of the United States and lung cancer mortality. Disease incidence, treatment disparities, and stage-specific survival. *Ann. Am. Thorac. Soc.* **14**, 403–411 (2017).
46. Xu, L., Kim, Y., Spolverato, G., Gani, F. & Pawlik, T. M. Racial disparities in treatment and survival of patients with hepatocellular carcinoma in the United States. *Hepatobiliary Surg. Nutr.* **5**, 43–52 (2016).
47. Lawrenson, R. et al. Treatment and survival disparities by ethnicity in New Zealand women with stage I-III breast cancer tumour subtypes. *Cancer Causes Control* **28**, 1417–1427 (2017).
48. Tyldesley, S. et al. Association between age and the utilization of radiotherapy in Ontario. *Int. J. Radiat. Oncol. Biol. Phys.* **47**, 469–480 (2000).
49. Magrath, I. et al. Paediatric cancer in low-income and middle-income countries. *Lancet Oncol.* **14**, e104–e116 (2013).
50. Duthey, B. & Scholten, W. Adequacy of opioid analgesic consumption at country, global, and regional levels in 2010, its relationship with development level, and changes compared with 2006. *J. Pain Symptom Manage.* **47**, 283–297 (2014).
51. McLaughlin, P. Y. et al. Do radiation oncology outreach clinics affect the use of radiotherapy? *Radiother. Oncol.* **127**, 143–149 (2018).
52. Robinson, D. et al. Waiting times for radiotherapy: variation over time and between cancer networks in southeast England. *Br. J. Cancer* **92**, 1201–1208 (2005).
53. Hershman, D. et al. Timing of adjuvant chemotherapy initiation after surgery for stage III colon cancer. *Cancer* **107**, 2581–2588 (2006).
54. Wasserman, D. W. et al. Reasons for delay in time to initiation of adjuvant chemotherapy for colon cancer. *J. Oncol. Pract.* **11**, e28–e35 (2014).
55. Arndt, V. et al. Patient delay and stage of diagnosis among breast cancer patients in Germany — a population based study. *Br. J. Cancer* **86**, 1034–1040 (2002).
56. McKenzie, F. et al. Drivers of advanced stage at breast cancer diagnosis in the multicountry African breast cancer — disparities in outcomes (ABC-DO) study. *Int. J. Cancer* **142**, 1568–1579 (2018).
57. Simos, D. et al. Imaging for distant metastases in women with early-stage breast cancer: a population-based cohort study. *CMAJ* **187**, E387–E397 (2015).
58. Parmar, A. D. et al. Quality of post-treatment surveillance of early stage breast cancer in Texas. *Surgery* **154**, 214–225 (2013).
59. Arnaout, A. et al. Use of preoperative magnetic resonance imaging for breast cancer: a Canadian population-based study. *JAMA Oncol.* **1**, 1238–1250 (2015).
60. Ashworth, A., Kong, W., Chow, E. & Mackillop, W. J. Fractionation of palliative radiation therapy for bone metastases in Ontario: do practice guidelines guide practice? *Int. J. Radiat. Oncol. Biol. Phys.* **94**, 31–39 (2016).
61. Ashworth, A., Kong, W., Whelan, T. & Mackillop, W. J. A population-based study of the fractionation of postlumpectomy breast radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **86**, 51–57 (2013).
62. Schnipper, L. E. 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.* **30**, 1715–1724 (2012).
63. Mitera, G. et al. Choosing Wisely Canada cancer list: ten low-value or harmful practices that should be avoided in cancer care. *J. Oncol. Pract.* **11**, e296–e303 (2015).
64. Olson, R. A. et al. Impact of using audit data to improve the evidence-based use of single-fraction radiation therapy for bone metastases in British Columbia. *Int. J. Radiat. Oncol. Biol. Phys.* **94**, 40–47 (2016).
65. Tyldesley, S., Boyd, C., Schulze, K., Walker, H. & Mackillop, W. J. Estimating the need for radiotherapy for lung cancer: an evidence-based, epidemiologic approach. *Int. J. Radiat. Oncol. Biol. Phys.* **49**, 973–985 (2001).
66. Round, C. E. et al. Radiotherapy demand and activity in England 2006–2020. *Clin. Oncol.* **25**, 522–530 (2013).
67. Mackillop, W. J. et al. A comparison of evidence-based estimates and empirical benchmarks of the appropriate rate of use of radiation therapy in Ontario. *Int. J. Radiat. Oncol. Biol. Phys.* **91**, 1099–1107 (2015).
68. Barbera, L., Zhang-Salomons, J., Huang, J., Tyldesley, S. & Mackillop, W. Defining the need for radiotherapy for lung cancer in the general population: a criterion-based, benchmarking approach. *Med. Care* **41**, 1074–1085 (2003).
69. Cancer Quality Council of Ontario. Access to radiation treatment. *Cancer System Quality Index* https://www.csqi.on.ca/indicators/access-radiation-treatment (2017).
70. Rahal, R., Klein-Geltink, J., Forte, T., Lockwood, G. & Bryant, H. Measuring concordance with guidelines for the diagnosis and treatment of colon cancer. *Curr. Oncol.* **20**, 227–229 (2013).
71. Salloum, R. G., Smith, T. J., Jensen, G. A. & Lafata, J. E. Factors associated with adherence to chemotherapy guidelines in patients with non-small cell lung cancer. *Lung Cancer* **75**, 255–260 (2012).
72. Young, J. M. et al. Concordance with national guidelines for colorectal cancer care in New South Wales: a population-based patterns of care study. *Med. J. Aust.* **186**, 292–295 (2007).
73. Wong, S. J. et al. Longitudinal Oncology Registry of Head and Neck Carcinoma (LORHAN): analysis of chemoradiation treatment approaches in the United States. *Cancer* **117**, 1679–1686 (2011).
74. Nandi, M., Mandal, A. & Asthana, A. K. Audit of cancer patients from eastern Uttar Pradesh (UP), India: a university hospital based two year retrospective analysis. *Asian Pac. J. Cancer Prev.* **14**, 4993–4998 (2013).
75. Porter, G. A., Urquhart, R., Bu, J., Johnson, P. & Grunfeld, E. The impact of audit and feedback on nodal harvest in colorectal cancer. *BMC Cancer* **11**, 2 (2011).
76. Alvarez, E. et al. Improvement in treatment abandonment in pediatric patients with cancer in Guatemala. *Pediatr. Blood Cancer* **64**, 10 (2017).
77. Gadalla, S. M. et al. A population-based assessment of mortality and morbidity patterns among patients with thymoma. *Int. J. Cancer* **128**, 2688–2694 (2011).
78. Lee, Y. Y. et al. Incidence and outcomes of pregnancy-associated cancer in Australia, 1994-2008: a population-based linkage study. *BJOG* **119**, 1572–1582 (2012).
79. Chu, C. N. et al. Increase in stroke risk in patients with head and neck cancer: a retrospective cohort study. *Br. J. Cancer* **105**, 1419–1423 (2011).
80. Marmor, S., Burke, E. E., Virnig, B. A., Jensen, E. H. & Tuttle, T. M. A comparative analysis of survival outcomes between pancreatectomy and chemotherapy for elderly patients with adenocarcinoma of the pancreas. *Cancer* **122**, 3378–3385 (2016).
81. Paszat, L. F., Mackillop, W. J., Groome, P. A., Schulze, K. & Holowaty, E. Mortality from myocardial infarction following postlumpectomy radiotherapy for breast cancer: a population-based study in Ontario, Canada. *Int. J. Radiat. Oncol. Biol. Phys.* **43**, 755–762 (1999).
82. Seruga, B., Sterling, L., Wang, L. & Tannock, I. F. Reporting of serious adverse drug reactions of targeted anticancer agents in pivotal phase III clinical trials. *J. Clin. Oncol.* **29**, 174–185 (2011).
83. Groome, P. A. et al. Management and outcome of glottic cancer: a population-based comparison between Ontario, Canada and the SEER areas of the United States. Surveillance, epidemiology and end results. *J. Otolaryngol.* **29**, 67–77 (2000).
84. Capri, S. & Russo, A. Cost of breast cancer based on real-world data: a cancer registry study in Italy. *BMC Health Serv. Res.* **17**, 84 (2017).
85. Siemens, D. R. et al. Processes of care and the impact of surgical volumes on cancer-specific survival: a population-based study in bladder cancer. *Urology* **84**, 1049–1057 (2014).
86. Booth, C. M., Nanji, S., Wei, X. & Mackillop, W. J. Management and outcome of colorectal cancer liver metastases in elderly patients: a population-based study. *JAMA Oncol.* **1**, 1111–1119 (2015).
87. Patel, M. I., Bang, A., Gillett, D., Cheluvappa, R. & Smith, D. P. 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.* **4**, 1145–1152 (2015).
88. Walker, M. et al. A call for theory-informed approaches to knowledge translation studies: an example of chemotherapy for bladder cancer. *Curr. Oncol.* **22**, 178–181 (2015).
89. Templeton, A. J. et al. Translating clinical trials to clinical practice: outcomes of men with metastatic castration resistant prostate cancer treated with docetaxel and prednisone in and out of clinical trials. *Ann. Oncol.* **24**, 2972–2977 (2013).
90. Booth, C. M. et al. Curative therapy for bladder cancer in routine clinical practice: a population-based outcomes study. *Clin. Oncol.* **26**, 506–514 (2014).
91. Stein, J. P. et al. Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1,054 patients. *J. Clin. Oncol.* **19**, 666–675 (2001).
92. Lau, K. et al. The effect of a regional hepatopancreaticobiliary surgical program on clinical volume, quality of cancer care, and outcomes in the Veterans Affairs system. *JAMA Surg.* **149**, 1153–1161 (2014).
93. Simunovic, M. et al. Assessing the volume-outcome hypothesis and region-level quality improvement interventions: pancreas cancer surgery in two Canadian Provinces. *Ann. Surg. Oncol.* **17**, 2537–2544 (2010).
94. van Roest, M. H., van der Aa, M. A., van der Geest, L. G. & de Jong, K. P. The impact of socioeconomic status, surgical resection and type of hospital on survival in patients with pancreatic cancer. A population-based study in The Netherlands. *PLOS ONE* **11**, e0166449 (2016).
95. Bendzsak, A. M., Baxter, N. N., Darling, G. E., Austin, P. C. & Urbach, D. R. Regionalization and outcomes of lung cancer surgery in Ontario, Canada. *J. Clin. Oncol.* **35**, 2772–2780 (2017).
96. Nienhuijs, S. W. et al. Reduction of in-hospital mortality following regionalisation of pancreatic surgery in the south-east of the Netherlands. *Eur. J. Surg. Oncol.* **36**, 652–656 (2010).
97. De Angelis, R. et al. The EUROCARE-4 database on cancer survival in Europe: data standardisation, quality control and methods of statistical analysis. *Eur. J. Cancer* **45**, 909–930 (2009).
98. Thomson, C. S. & Forman, D. Cancer survival in England and the influence of early diagnosis: what can we learn from recent EUROCARE results? *Br. J. Cancer* **101** (Suppl. 2), 102–109 (2009).
99. Lozano, R. et al. Benchmarking of performance of Mexican states with effective coverage. *Lancet* **368**, 1729–1741 (2006).
100. Zubizarreta, E. H., Fidarova, E., Healy, B. & Rosenblatt, E. Need for radiotherapy in low and middle income countries — the silent crisis continues. *Clin. Oncol.* **27**, 107–114 (2015).
101. Nandakumar, A. et al. Concurrent chemoradiation for cancer of the cervix: results of a multi-institutional study from the setting of a developing country (India). *J. Glob. Oncol.* **1**, 11–22 (2015).
102. Veenstra, C. M. et al. Long-term economic and employment outcomes among partners of women with early-stage breast cancer. *J. Oncol. Pract.* **13**, e916–e926 (2017).
103. Lin, P. J. et al. Linking costs and survival in the treatment of older adults with chronic myeloid leukemia: an analysis of SEER-Medicare data from 1995 to 2007. *Med. Care* **54**, 380–385 (2016).
104. Roehrborn, C. G. & Black, L. K. The economic burden of prostate cancer. *BJU Int.* **108**, 806–813 (2011).
105. Mullins, C. D., Hsiao, F. Y., Onukwugha, E., Pandya, N. B. & Hanna, N. Comparative and cost-effectiveness of oxaliplatin-based or irinotecan-based regimens compared with 5-fluorouracil/leucovorin alone among US elderly stage IV colon cancer patients. *Cancer* **118**, 3173–3181 (2012).
106. Mackillop, W. J. in *Clinical Radiation Oncology* 3rd edn (eds Gunderson, L. & Tepper, J.) 203–222 (Churchill Livingstone, Philadelphia, 2012).
107. Ong, M. B. H. Flatiron, BMS form collaboration to curate regulatory-grade real-world data. *The Cancer Letter* https://cancerletter.com/articles/20180504_1/ (2018).
108. Goodwin, P. J., Ballman, K. V., Small, E. J. & Cannistra, S. A. Evaluation of treatment benefit in journal of clinical oncology. *J. Clin. Oncol.* **31**, 1123–1124 (2013).
109. Chang, G. J. Is there validity in propensity score-matched estimates of adjuvant chemotherapy effects for patients with rectal cancer? *JAMA Oncol.* **4**, 921–923 (2018).
110. Choudhury, A. & Hoskin, P. J. Bladder cancer and the National Cancer Data Base: new insight or misinformation? *Cancer* **124**, 1105–1107 (2018).
111. Unger, J. M., Cook, E., Tai, E. & Bleyer, A. The role of clinical trial participation in cancer research: barriers, evidence, and strategies. *Am. Soc. Clin. Oncol. Educ. Book* **35**, 185–198 (2016).

112. The Children's Oncology Group. What is a clincial trial? *COG* https://childrensoncologygroup.org/index.php/what-is-a-clinical-trial (2018).
113. Yusuf, S., Collins, R. & Peto, R. Why do we need some large, simple randomized trials? *Stat. Med.* **3**, 409–422 (1984).
114. Parmar, M. K. et al. Paclitaxel plus platinum-based chemotherapy versus conventional platinum-based chemotherapy in women with relapsed ovarian cancer: the ICON4/AGO-OVAR-2.2 trial. *Lancet* **361**, 2099–2106 (2003).
115. James, N. D. et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *Lancet* **387**, 1163–1177 (2016).
116. Khozin, S., Blumenthal, G. M. & Pazdur, R. Real-world data for clinical evidence generation in oncology. *J. Natl Cancer Inst.* **109**, djx187 (2017).
117. Corrigan-Curay, J., Sacks, L. & Woodcock, J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. *JAMA* **320**, 867–868 (2018).
118. Sackett, D. L. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* **95**, 2S–4S (1989).
119. Oxford Centre for Evidence-Based Medicine. OCEBM levels of evidence. *CEBM* https://www.cebm.net/2016/05/ocebm-levels-of-evidence/ (2018).
120. Stukel, T. A. et al. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. *JAMA* **297**, 278–285 (2007).
121. Schneeweiss, S. & Maclure, M. Use of comorbidity scores for control of confounding in studies using administrative databases. *Int. J. Epidemiol.* **29**, 891–898 (2000).
122. Groome, P. A. & Mackillop, W. J. Uses of ecologic studies in the assessment of intended treatment effects. *J. Clin. Epidemiol.* **52**, 903–904 (1999).
123. Pearcey, R., Miao, Q., Kong, W., Zhang-Salomons, J. & Mackillop, W. J. Impact of adoption of chemoradiotherapy on the outcome of cervical cancer in Ontario: results of a population-based cohort study. *J. Clin. Oncol.* **25**, 2383–2388 (2007).
124. Groome, P. A. et al. Management and outcome differences in supraglottic cancer between Ontario, Canada, and the surveillance, epidemiology, and end results areas of the United States. *J. Clin. Oncol.* **21**, 496–505 (2003).
125. Gupta, S., Kong, W., Booth, C. M. & Mackillop, W. J. Impact of concomitant chemotherapy on outcomes of radiation therapy for head-and-neck cancer: a population-based study. *Int. J. Radiat. Oncol. Biol. Phys.* **88**, 115–121 (2014).
126. Booth, C. M. et al. Adjuvant chemotherapy for non-small cell lung cancer: practice patterns and outcomes in the general population of Ontario, Canada. *J. Thorac. Oncol.* **7**, 559–566 (2012).
127. Erlichman, C. Efficacy of adjuvant fluorouracil and folinic acid in B2 colon cancer. *J. Clin. Oncol.* **17**, 1356–1363 (1999).
128. van Gijn, W. et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. *Lancet Oncol.* **12**, 575–582 (2011).
129. Casadaban, L. et al. Adjuvant chemotherapy is associated with improved survival in patients with stage II colon cancer. *Cancer* **122**, 3277–3287 (2016).
130. Freischlag, K. et al. Association between incomplete neoadjuvant radiotherapy and survival for patients with locally advanced rectal cancer. *JAMA Surg.* **152**, 558–564 (2017).
131. Glasziou, P., Vandenbroucke, J. P. & Chalmers, I. Assessing the quality of research. *BMJ* **328**, 39–41 (2004).
132. Sanderson, S., Tatt, I. D. & Higgins, J. P. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int. J. Epidemiol.* **36**, 666–676 (2007).
133. Vandenbroucke, J. P. et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Ann. Intern. Med.* **147**, W163–W194 (2007).
134. Wells, G. S. et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. *Ottawa Hospital Research Institute* http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm (2009).
135. Stang, A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur. J. Epidemiol.* **25**, 603–605 (2010).
136. Suissa, S. Immortal time bias in pharmaco-epidemiology. *Am. J. Epidemiol.* **167**, 492–499 (2008).
137. Hill, A. B. The environment and disease: association of causation. *Proc. R. Soc. Med.* **58**, 295–300 (1965).
138. Rothman, K. J., Greenland, S. & Lash, T. L. *Modern epidemiology* (Lippincott Williams & Wilkins, Philadelphia, 2008).
139. Makady, A. et al. Policies for use of real-world data in health technology assessment (HTA): a comparative study of six HTA agencies. *Value Health* **20**, 520–532 (2017).
140. Risk Sciences International. Evidence synthesis on post-approval surveillance of approved novel cancer drugs. Canadian Partnership Against Cancer https://content.cancerview.ca/download/cv/quality_and_planning/system_planning/quality_initiatives_microsite/documents/evidence_synthesis_post_approval_surveillance_cancer_drugs_enpdf?attachment=0 (2015).
141. Jarow, J. P., LaVange, L. & Woodcock, J. Multidimensional evidence generation and FDA regulatory decision making: defining and using "real-world" data. *JAMA* **318**, 703–704 (2017).
142. Food and Drug Administration. Statement from FDA Commissioner Scott Gottlieb, M. D., on FDA's new strategic framework to advance use of real-world evidence to support development of drugs and biologics. *FDA* https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm627760.htm (2018).
143. Hagen, T. ASCO partners with FDA to incorporate real-world evidence into drug decisions. *OncLive* http://www.onclive.com/conference-coverage/asco-2017/asco-partners-with-fda-to-incorporate-realworld-evidence-into-drug-decisions (2017).
144. Hall, P. S. Real-world data for efficient health technology assessment. *Eur. J. Cancer* **79**, 235–237 (2017).
145. Lewis, J. R., Kerridge, I. & Lipworth, W. Coverage with evidence development and managed entry in the funding of personalized medicine: practical and ethical challenges for oncology. *J. Clin. Oncol.* **33**, 4112–4117 (2015).
146. Espin, J., Rovira, J. & Garcia, L. Experiences and impact of European risk-sharing schemes focusing on oncology medicines. *European Commission* http://ec.europa.eu/DocsRoom/documents/7608 (2011).
147. Fujiwara, Y. & Kobayashi, K. Oncology drug clinical development and approval in Japan: the role of the pharmaceuticals and medical devices evaluation center (PMDEC). *Crit. Rev. Oncol. Hematol.* **42**, 145–155 (2002).
148. Mackillop, W. J. & Kong, W. Estimating the need for palliative radiation therapy: a benchmarking approach. *Int. J. Radiat. Oncol. Biol. Phys.* **94**, 51–59 (2016).

Acknowledgements
The work of C.M.B. is supported by his position as the Canada Research Chair in Population Cancer Care. All the authors thank I. Tannock and E. Eisenhauer for their comments on an earlier draft of this manuscript.

Author contributions
All authors made a substantial contribution to all aspects of the preparation of this manuscript.

Competing interests
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Reviewer information
Nature Reviews Clinical Oncology thanks A. Abernethy, C. Uyl-de-Groot and the other anonymous reviewer(s) for their contribution to the peer review of this work.