European Journal of Cancer 186 (2023) 52–61

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Current Perspective

Defining the role of real-world data in cancer clinical research: The position of the European Organisation for Research and Treatment of Cancer

Robbe Saesen [a,*], Mieke Van Hemelrijck [b,c], Jan Bogaerts [a], Christopher M. Booth [d], Jan J. Cornelissen [e], Andre Dekker [f], Elizabeth A. Eisenhauer [d], André Freitas [g,h,i], Alessandro Gronchi [j], Miguel A. Hernán [k], Frank Hulstaert [l], Piet Ost [m,n], Petr Szturz [o], Helena M. Verkooijen [p,q], Michael Weller [r], Roger Wilson [s,t], Denis Lacombe [a,l], Winette T. van der Graaf [u,v,l]

[a] *European Organisation for Research and Treatment of Cancer (EORTC), Brussels, Belgium*
[b] *Translational Oncology and Urology Research, School of Cancer and Pharmaceutical Sciences, King's College London, London, United Kingdom*
[c] *Guy's Cancer Real World Evidence, Guy's and St Thomas NHS Foundation Trust, London, United Kingdom*
[d] *Department of Oncology, Queen's University, Kingston, Ontario, Canada*
[e] *Department of Hematology, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, The Netherlands*
[f] *Department of Radiation Oncology (Maastro), GROW School for Oncology and Reproduction, Maastricht University Medical Center+, Maastricht, The Netherlands*
[g] *Department of Computer Science, University of Manchester, Manchester, United Kingdom*
[h] *Cancer Biomarker Centre, CRUK Manchester Institute, University of Manchester, Manchester, United Kingdom*
[i] *Idiap Research Institute, Martigny, Switzerland*
[j] *Department of Surgery, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*
[k] *CAUSALab and Departments of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States*
[l] *Belgian Health Care Knowledge Centre (KCE), Brussels, Belgium*
[m] *Department of Human Structure and Repair, Ghent University, Ghent, Belgium*
[n] *Department of Radiation Oncology, Iridium Network, Antwerp, Belgium*
[o] *Department of Oncology, University of Lausanne (UNIL) and Lausanne University Hospital (CHUV), Lausanne, Switzerland*
[p] *Division of Imaging and Oncology, University Medical Center Utrecht, Utrecht, The Netherlands*
[q] *University of Utrecht, Utrecht, The Netherlands*
[r] *Department of Neurology, University Hospital and University of Zurich, Zurich, Switzerland*
[s] *Sarcoma Patients Advocacy Global Network (SPAGN), Wölfersheim, Germany*
[t] *EORTC Patient Panel, Brussels, Belgium*
[u] *Medical Oncology Department, Netherlands Cancer Institute, Amsterdam, The Netherlands*
[v] *Medical Oncology Department, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, The Netherlands*

[*] *Corresponding author.*
E-mail address: robbe.saesen@eortc.org (R. Saesen).
[1] Shared last authors.

https://doi.org/10.1016/j.ejca.2023.03.013

Received 28 February 2023; Received in revised form 9 March 2023; Accepted 10 March 2023
Available online 21 March 2023

KEYWORDS
Real-world data;
Real-world evidence;
Randomised controlled trials;
Pragmatic trials;
Cohort multiple randomised controlled trials;
Trials within cohorts;
Add-on cohorts;
Clinical research;
Cancer;
Oncology

Abstract The emergence of the precision medicine paradigm in oncology has led to increasing interest in the integration of real-world data (RWD) into cancer clinical research. As sources of real-world evidence (RWE), such data could potentially help address the uncertainties that surround the adoption of novel anticancer therapies into the clinic following their investigation in clinical trials. At present, RWE-generating studies which investigate antitumour interventions seem to primarily focus on collecting and analysing observational RWD, typically forgoing the use of randomisation despite its methodological benefits. This is appropriate in situations where randomised controlled trials (RCTs) are not feasible and non-randomised RWD analyses can offer valuable insights. Nevertheless, depending on how they are designed, RCTs have the potential to produce strong and actionable RWE themselves. The choice of which methodology to employ for RWD studies should be guided by the nature of the research question they are intended to answer. Here, we attempt to define some of the questions that do not necessarily require the conduct of RCTs. Moreover, we outline the strategy of the European Organisation for Research and Treatment of Cancer (EORTC) to contribute to the generation of robust and high-quality RWE by prioritising the execution of pragmatic trials and studies set up according to the trials-within-cohorts approach. If treatment allocation cannot be left up to random chance due to practical or ethical concerns, the EORTC will consider undertaking observational RWD research based on the target trial principle. New EORTC-sponsored RCTs may also feature concurrent prospective cohorts composed of off-trial patients.

1. Setting the scene: an introduction to real-world data

Although various interpretations of the term 'real-world data' (RWD) seem to exist, it is most frequently used to denote health data that have been collected outside of traditional randomised controlled trials (RCTs), typically as part of routine clinical practice [1]. If analysed appropriately, such data can give rise to so-called real-world evidence (RWE), which can offer insights into the benefits and risks of therapeutic interventions as observed in a real-life environment as opposed to a trial setting [2,3].

RWD can originate from a wide variety of sources, with some of the most common ones being patient registries, electronic health records (EHRs) and administrative claims [4,5]. Patient registries are organised systems that apply observational methods to longitudinally capture specific information on patients characterised by their condition (disease registries) or by their receipt of a particular treatment (treatment registries) [4,5]. Guidance on how RWD extracted from registries can help inform the decisions made by regulatory authorities has been published by both the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA) [6,7]. EHRs are digital versions of patients' medical records, documenting their encounters with healthcare providers, whether clinical in nature or not [4,5]. The RWD contained within EHRs can be de-identified and aggregated for research purposes. Administrative claims are claims that healthcare providers submit to payers in order to get compensated for health services they provided to their patients [4,5]. They are stored in databases maintained by insurers from the public or private sector, by pharmacies, or by hospitals. Additional sources of RWD include patient questionnaires, wearable devices, smartphones and social media [4,5].

Studies that produce RWE confer several important advantages over conventional clinical trials. For instance, the size and scale of the underlying datasets can be much bigger, covering a greater and more representative proportion of the population under investigation, and improving the generalisability of the findings obtained [8]. Furthermore, since RWD have often already been generated prior to the start of a study in which they are examined, the research objectives tackled can be achieved more quickly, requiring just a few weeks or months instead of multiple years [8]. In addition, because RWD studies may take less time and involve fewer procedures, they are normally not as expensive to perform [8]. They are sometimes also the only viable means of studying treatments for rare diseases [9]. However, their conduct is impeded by a number of challenges. On an operational level, it can be difficult to gain access to the data and to protect them in a manner that is compliant with the laws and regulations

governing their use by researchers [10]. From a technical perspective, the RWD coming from different hospitals and institutions may not be interoperable due to formatting variations and differences in the recording of variables rendering them incompatible [10]. The EHDEN project running under the auspices of the Innovative Medicines Initiative (IMI) is attempting to reduce heterogeneity in this regard by mapping institutional data to the Observational Medical Outcomes Partnership (OMOP) common data model of the Observational Health Data Sciences and Informatics (OHDSI) collaborative [11]. Methodologically speaking, selection and information biases as well as confounding can undermine the internal and external validity of RWD analyses [10]. These methodological flaws lead to misleading and artifactual results which may harm patients by unduly influencing clinical care.

In recent years, regulators and payers have increasingly incorporated RWE into their decision-making [12,13], as the efforts of the EMA to establish the Data Analysis and Real-World Interrogation Network of the European Union (DARWIN EU) [14] have clearly demonstrated. Pharmaceutical companies have also rapidly expanded their reliance on RWE as a tool to support the market access of the drugs they develop [15]. Additionally, many initiatives have been launched to harness the knowledge embedded within RWD. For example, under IMI's Big Data For Better Outcomes (BD4BO) programme [16] and the Horizon Europe framework, numerous projects have been set up to stimulate and facilitate the exploitation of these data, including, besides EHDEN [11], PIONEER [17], HARMONY [18] and IDEA4RC [19]. The GetReal Institute [20], an offshoot of another IMI project, strives to enhance the credibility of RWE by identifying best practices for the collection and analysis of RWD and by fostering collaboration between stakeholders.

In the field of oncology, where the emergence of the precision medicine model has complicated the setup and execution of large and rigorously designed RCTs, regulatory agencies like EMA and FDA have resorted to granting marketing authorisations to novel anticancer agents based on their performance in relatively small trials, which may not employ randomisation or feature any comparator arms at all [21]. More concretely, over the past 20 years, the FDA has approved 176 indications for such agents exclusively on the basis of insights acquired from single-arm studies [22]. As a result, uncertainties usually remain about the effects of these products on patients' quality of life and overall survival at the time of their approval by regulators [23]. In addition, questions surrounding their optimal integration into the available armamentarium of antineoplastic treatments are generally not addressed in the registrational trials carried out by industry [24–26]. While RWD analyses continue to spark controversy with regularity, it has been suggested that they can help fill at least some of the evidence gaps with which downstream decision-makers such as health technology assessment (HTA) bodies and clinicians are currently faced [27–29]. Moreover, patients broadly appear to have favourable opinions of RWE, expressing willingness to contribute to its generation, under certain conditions [30]. In light of the growing importance of RWD for the development of antitumor therapies of all types (not just pharmacotherapy, but surgery and radiotherapy too) [31,32], the role that data of this kind can play in cancer clinical research warrants further scrutiny. Here, we aim to define that role from an academic viewpoint, namely the one of the European Organisation for Research and Treatment of Cancer (EORTC).

2. RWD versus RCT-derived data: a false dichotomy

RWD studies are commonly juxtaposed against RCTs in the literature [33–36], with comparisons between the two mainly centring around the relative strength of the evidence they yield. In these comparisons, RCTs invariably prevail due to their inherent methodological robustness. By including a control group and leaving treatment allocation up to random chance rather than deliberate choice, the RCT design eliminates the need to manage the impact of confounders, no matter whether they are known or not [33,37]. Even though treatment effects can be validly estimated from non-randomised data when confounding factors are correctly adjusted for, there are no guarantees that all of these factors will be identified and measured [33,37], and erroneous conclusions may therefore be drawn from RWD analyses. Additionally, in RCTs, the research data are standardly collected in a prospective manner, which ensures that they are fit for purpose and allows for their quality to be safeguarded [38]. As RWD are oftentimes sourced from existing datasets that were only created to support the provision of care to patients, they may be of insufficient evidentiary value to deliver actionable RWE [2]. For instance, although regulators require information on both the benefits and the risks of the drugs they evaluate [39], many registries, unlike RCT-linked databases, do not record in detail the toxicities experienced by the individuals whose data they capture, nor the comorbidities these people suffer from [40]. This also makes it challenging to offer tailored therapies to patients, especially if they are elderly, as the case of acute myeloid leukaemia shows [41].

However, the supposed dichotomy between RWD and RCT-derived data should be considered a false one: in reality, depending on how they are set up, RCTs can be RWD studies themselves and produce RWE [42]. For example, pragmatic trials try to emulate the circumstances under which an investigational intervention would be administered in clinical practice by, among other things, employing more relaxed eligibility criteria to recruit a heterogeneous sample of participants,

reducing the number of study procedures to a minimum, and relying on healthcare infrastructure that is already in place to monitor therapeutic outcomes [43–46]. Consequently, such trials measure how well the intervention in question works when it is applied in a real-life setting, providing a generalisable estimate of its effectiveness in lieu of its efficacy. When RWE is obtained from comparative studies in which randomisation is used, it can be labelled randomised RWE, or R²WE for short. This R²WE may be seen as an enhanced version of RWE, representing the product of trials that combine the analytical rigour of RCTs with the inclusiveness of RWD research [37,47,48]. Similarly, randomised RWD or R²WD could be regarded as an augmented form of RWD. It should be noted here that R²WD studies may be subject to protocol deviations (e.g. non-adherence to treatment, loss to follow-up). When such deviations occur, the analysis of the data generated in these studies requires assumptions and adjustments similar to those needed for RWD analyses.

Nevertheless, RCTs are not always ethically or practically feasible to undertake, or may still be under way at the time decisions need to be made about the treatments they investigate. In situations where there is no R²WE to fall back on, RWE arising from observational RWD studies can be a valuable alternative [49,50], with the caveat that one should be cautious about adopting new therapies based on such evidence alone, given its susceptibility to bias [48,49]. The notion of abandoning RCTs as the gold standard methodology in this respect must thus be dismissed [37,48]. In fact, it can be argued that instead of doing away with RCTs, the focus should shift towards conducting more trials that are capable of R²WE generation [37,47,48]. On the other hand, in a scenario where an RCT has been performed and its results have been published, these results can be used as benchmarks for non-interventional and non-randomised RWD analyses [51]. Failure of such analyses to replicate the findings of the RCT, even after explicitly emulating it, likely means insufficient data are available to correctly estimate the effect of interest [51]. Conversely, when the outcomes observed in the RWD studies and in the RCT approximate each other, the inferences from the trial can cautiously be extended to other patient populations or to longer follow-up schedules [51]. The use of RWD to extend inferences from RCTs after benchmarking is key because it is not achievable to run long-term RCTs in all patient populations [51].

3. Defining the scope of applicability of RWD and R²WD

When a decision needs to be made on which methodological setup would be most suitable for answering a specific research question, the choice of study design should primarily be determined by the nature of this question. Different questions necessitate different approaches: while RWD analyses can for instance shed light on the performance of treatments in particular subpopulations of patients that were excluded from participating in any preceding trials [38,52], they may not be appropriate for assessing the effectiveness of these treatments relative to established standards of care [27,49]. As explained above, to replace an intervention that is routinely administered in clinical practice with a novel therapy, evidence from randomised head-to-head comparison studies justifying this substitution should ideally be available [48,49], potentially in the form of R²WE. It may not be possible though to deliver such evidence in situations where equipoise is no longer present and randomisation would therefore not be feasible, for example when a new treatment shows very high levels of activity in a specific subpopulation of patients who would otherwise receive the standard of care.

The selection of a methodology that is congruent with the research question to be tackled can be guided by the following four factors (non-exhaustive list):

- The study sample: is a population-level dataset required, or does a subset of patients suffice?
- The temporal character of the study: can data collection predate the design of the study by leveraging existing datasets, or should the data be generated from scratch and collected only after the study has been set up?
- The delivery of the study treatment: do the conditions under which the treatment is delivered by investigators have to be controlled, or is it not necessary to intervene?
- The study purpose: does the study aim to formulate hypotheses, or to test them?

Table 1 details the scope of applicability of RWD and R²WD for a number of research questions based on these four factors.

4. The strategy of the EORTC

The EORTC is a not-for-profit organisation that has been carrying out independent and multidisciplinary clinical research in oncology for over 60 years, with the goal of improving cancer patients' quality of life and overall survival. Since its founding, the EORTC has executed in excess of 1400 studies through its network of more than 3100 scientists working across 760 distinct institutions in 48 different countries [53]. Many EORTC-sponsored trials have led to major changes in oncological practice [54], and the instruments developed by the organisation to measure health-related quality of life in a standardised and validated manner have been widely adopted by stakeholders in the field [55].

Throughout its history, the EORTC has focused on making causal inferences about the effects of anti-neoplastic therapies on patient outcomes by performing comparative studies [56,57], and more specifically, by setting up RCTs in all tumour types. Its trials, which are

Table 1
Overview of the optimal methodologies to address specific research questions.

Research question	Study sample	Temporal character of study	Delivery of study treatment	Study purpose	Optimal methodology
Is treatment A better than treatment B in terms of improving patient outcomes?	Subset	Study design predates data collection	Interventional	Hypothesis-testing/ confirmatory	Traditional RCT or R²WD study
Is treatment A given in a lower dose equally effective as treatment A given in a standard dose in terms of improving patient outcomes?	Subset	Study design predates data collection	Interventional	Hypothesis-testing/ confirmatory	Traditional RCT or R²WD study
Is treatment A given intermittently equally effective as treatment A given continuously in terms of improving patient outcomes?	Subset	Study design predates data collection	Interventional	Hypothesis-testing/ confirmatory	Traditional RCT or R²WD study
Is treatment A given for a limited duration of time equally effective as treatment A given lifelong in terms of improving patient outcomes?	Subset	Study design predates data collection	Interventional	Hypothesis-testing/ confirmatory	Traditional RCT or R²WD study
To what extent has treatment A been taken up by clinicians?	Population	Data collection predates study design	Observational	Hypothesis-generating/ exploratory	RWD study
What has been the impact of treatment A on patient outcomes in the general population?	Population	Data collection predates study design	Observational	Hypothesis-generating/ exploratory	RWD study
To what extent does treatment A improve patient outcomes in a specific subpopulation?	Subset	Study design predates data collection or data collection predates study design	Observational or interventional	Hypothesis-generating/ exploratory	RWD study or R²WD study

designed to establish new standards of care, investigate not only pharmacological treatments like chemotherapeutic and targeted agents, but also surgical and radiotherapeutic interventions. While the organisation maintains several cohorts as part of the SPECTA [58] and E²-RADIatE [59] platforms, it has refrained from regularly engaging in RWD research, mostly because its purview does not extend to purely descriptive or predictive analyses [56,57].

However, recent innovations in data science (e.g. federated data infrastructures, artificial intelligence and machine learning applications, causal models) and the quickly evolving RWD landscape across most cancer domains have prompted the EORTC to rethink its position in this area and devise a strategy for contributing to the development of actionable RWE in the future, drawing on its rich experience with conducting academic RCTs. After surveying its members [29] and hosting a multidisciplinary workshop with participation of international subject matter experts and patient representatives, the organisation has decided to prioritise the execution of clinical trials that produce R²WE. The emphasis of its RWD research plans will therefore lie on randomised, interventional RWD studies which are undertaken in subsets of patients for hypothesis-testing purposes and whose design predates their collection of data (Table 2). Only when randomisation is impossible to implement, will the EORTC consider running non-randomised or observational RWD analyses.

Regardless of whether they are capable of yielding RWE, EORTC trials will continue to be initiated with the chief aim of delivering evidence that can underpin the adoption of novel therapeutic approaches into the clinic. The organisation may nonetheless launch RWD studies whose results would provide additional information on treatments that are already in common use if these studies are intended to answer a clinically relevant, patient-centred research question which an

Table 2
Focus of the EORTC's future RWD research.

Factor	Focus of EORTC RWD research
Study sample	Subset
Temporal character of study	Study design predates data collection
Delivery of study treatment	Interventional (preferably) or observational (if interventional not possible)
Study purpose	Hypothesis-testing/ confirmatory
Type of RWD	R²WD (preferably) or RWD (if R²WD not possible)

RCT would be unable to tackle. In the absence of such a question, the EORTC will likely forgo collecting RWD, unless the knowledge that would be acquired from their analysis is needed to sustain its scientific agenda. In general, the organisation will exclusively work with high-quality and auditable RWD.

More concretely, when designing a study in which RWD will be analysed, the EORTC will give precedence to the research methodologies shown in Fig. 1. Trials that can generate R²WE will play a key role in its RWD strategy. These include both pragmatic trials (Fig. 1B) [43–46] and cohort multiple RCTs (cmRCTs), which are studies that employ the trials-within-cohorts (TwiCs) setup (Fig. 1C) [60–62]. In cmRCTs or TwiCs studies, a random selection of eligible patients that are part of a large prospective cohort are offered the choice to receive the investigational intervention, while the other participants satisfying the eligibility criteria are given usual care and constitute the study's control group, not being informed of the fact that a trial is in progress. Since the cohort is integrated into clinical practice, the evidence that is derived from cmRCTs can be categorised as R²WE. The EORTC is currently planning to undertake a TwiCs study that builds on its OligoCare cohort [63].

Besides interventional R²WD studies, the EORTC may elect to carry out observational RWD analyses when treatment allocation cannot feasibly be randomised. In such circumstances, the target trial principle will be followed, meaning that the design of the RWD study to be performed will mimic that of a hypothetical RCT which addresses the same research question (Fig. 1D) [64–66]. This involves preparing an extensive protocol beforehand that outlines the study's

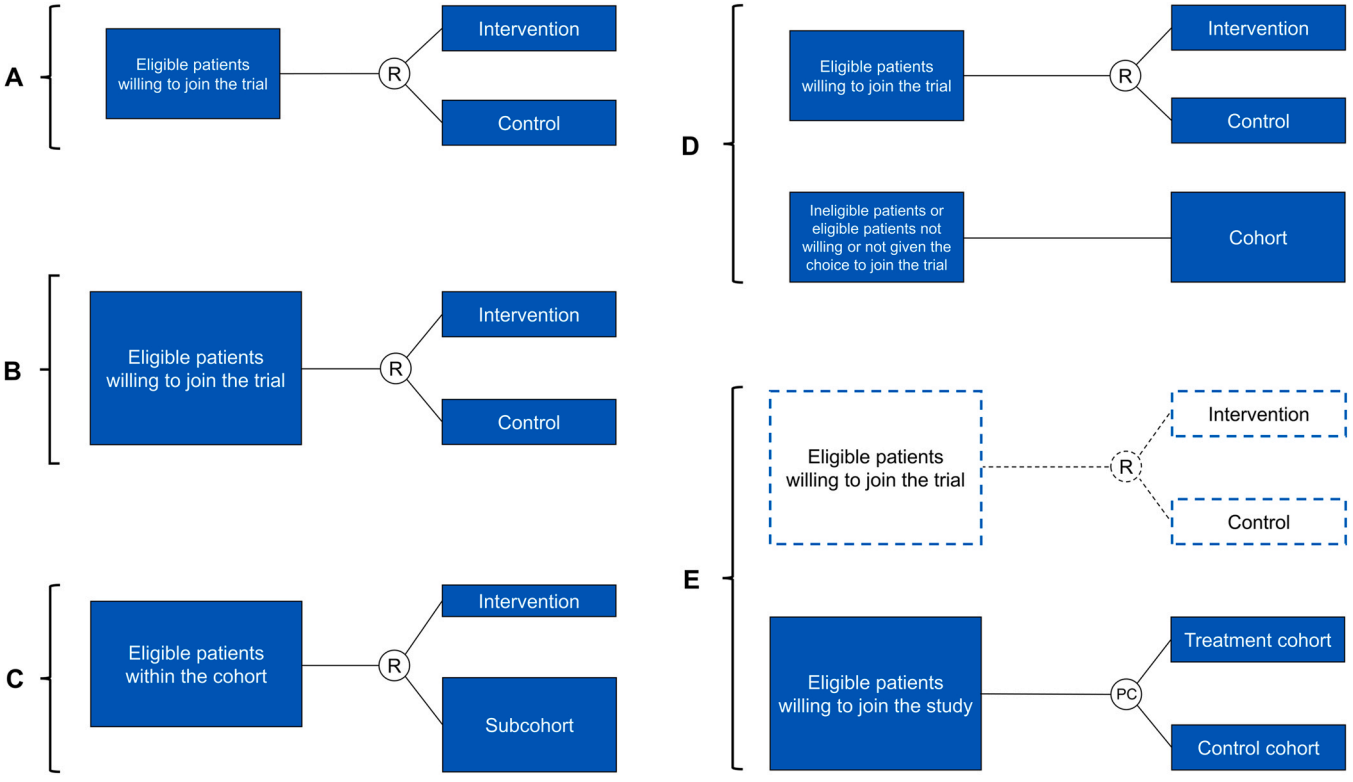

Fig. 1. Schematic overview of the methodological designs of (A) a traditional RCT, (B) a pragmatic trial, (C) a cmRCT, (D) an RCT with an add-on cohort, and (E) an observational study emulating a hypothetical target trial. R = randomisation; PC = physician's choice. Note the difference in the size of the eligible population between (A) and (B).

methodological characteristics, including the inclusion and exclusion criteria used, the end-points assessed, and the statistical tests conducted. The emulation of the target trial, which should be pragmatic in nature, requires that (A) the start of follow-up (also known as time zero) coincides with the determination of patient eligibility and the assignment of the intervention (to prevent immortal time bias and selection bias), and (B) confounders are adjusted for as much as possible [64–66].

In addition, the EORTC will evaluate the possibility of recruiting individuals who do not take part in its future trials into concurrent add-on cohorts which are monitored alongside the participants of each new study (Fig. 1E). Patients in these cohorts are subjected to the same or similar procedures by the same doctors in the same institutions across the same time interval as their counterparts in the RCT that is taking place in parallel. Consequently, such cohorts represent a source of contemporary RWD that despite originating from a highly specialised setting can contextualise the data from the accompanying trial, which could then be labelled as an augmented RCT [67]. For example, the investigators of the EORTC-sponsored STRASS study [68], which examined whether adding preoperative radiotherapy to surgery can improve abdominal recurrence-free survival compared to surgery alone in primary retroperitoneal sarcoma (RPS), retrospectively assembled a cohort of patients (STREXIT) who did not join the trial but were administered care identical to that received by STRASS participants, with many being treated according to the radiation schedule studied in the RCT [69]. This allowed for comparisons to be made between the outcomes of on- and off-trial patients. Furthermore, upon merging the STRASS and adjusted STREXIT datasets, it became possible to execute relevant subgroup analyses due to the cohort's large size. Going forward, add-on cohorts like STREXIT will be set up prospectively by the EORTC. Recently, the organisation has acquired European funding for the launch of the STREXIT2 prospective add-on cohort within the context of the STRASS2 trial, which investigates whether neoadjuvant chemotherapy can benefit high-risk RPS patients [70]. It should be highlighted that securing sufficient funds to run large-scale, international RWD or R²WD studies can be extremely challenging.

The EORTC operates in a complex, multi-stakeholder environment, and its clinical research portfolio reflects this by comprising trials that are designed to resolve uncertainties with which patients, clinicians and HTA bodies are confronted. RWD studies undertaken by the organisation will also be initiated with this objective in mind, irrespective of the methodology employed.

5. Conclusions

RWD are increasingly being incorporated into cancer clinical research, often being characterised as an alternative source of evidence to support the decision-making of regulators, payers and clinicians. Until now, RWD studies in oncology have mainly relied on the collection and analysis of observational data, which can offer valuable insights into the real-world conditions in which antitumor treatments are applied. However, given the inherent methodological limitations of such data, their quality may sometimes be too low to yield actionable RWE. Certain research questions require a different approach, as they can likely only be conclusively answered through the conduct of RCTs. Hybrid methodologies that combine the strengths of both RCTs and RWD studies may be able to address these questions by producing R²WE, which can be regarded as an enhanced form of RWE. Recognising this, the EORTC has devised a strategy for contributing to the generation of R²WE by prioritising the execution of pragmatic trials and cmRCTs in the future. In situations where randomisation cannot be implemented for practical or ethical reasons, the organisation will consider setting up non-randomised RWD studies adhering to the target trial principle, if circumstances permit. Moreover, the possibility of assembling add-on cohorts running synchronously with new EORTC trials will be explored on a case-by-case basis. Ultimately, the study design used should always be determined by the nature of the research question to be tackled.

Funding statement

Robbe Saesen's work as a Fellow at EORTC Headquarters was supported by a grant from the EORTC Cancer Research Fund (ECRF). The funder had no role in the conceptualisation nor the writing of this manuscript.

CRediT authorship contribution statement

All authors: Conceptualisation, Writing – Review & Editing. **R.S.**: Writing – Original Draft, Visualisation. **D.L., W.T.V.D.G.**: Supervision, Project administration.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: R.S.: Nothing to declare. M.V.H.: Nothing to declare. J.B.: Nothing to declare. C.M.B.: Nothing to declare. J.J.C.: Nothing to declare. A.D.: I have received institutional research funding from Janssen and IQVIA, speaker honoraria from Janssen and Medtronic, and am founder, shareholder and employee of Medical Data Works B.V. E.A.E.: Nothing to declare. A.F.: Nothing to declare. A.G.: Nothing to declare. M.A.H.: My research is supported by funding from NIH, the US Veterans Administration, and PCORI. I am a consultant for Cytel, data science adviser for ProPublica,

and member of the scientific advisory board of ADIA Lab. None of these interests or relationships have influenced my contribution to this paper. F.H.: I report that a party related to me is employed by Janssen-Cilag GmbH as principal data scientist. P.O.: I have performed consultancy work for AAA, Bayer, Curium, MSD, Janssen and have received research grants from Bayer. P.S.: I have had advisory relationships with Merck-Serono, Servier, and Merck Sharp & Dohme Corp in the last three years. H.M.V.: Nothing to declare. M.W.: I have received research grants from Quercis and Versameb, and honoraria for lectures or advisory board participation or consulting from Bayer, Curevac, Medac, Novartis, Novocure, Orbus, Philogen, Roche and Sandoz. R.W.: Nothing to declare. D.L.: I am member of the EMA Management Board, member of the EMA Health Care Professionals Working Party and of the EMA Healthcare Professionals Policy Officers' Group, and co-chair of the EMA Cancer Medicines Forum. W.T.V.D.G.: I have performed advisory work for SpringWorks, PTC Therapeutics, Agenus and received research funding from Eli Lilly, all fees going to the institute.

References

[1] Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What is real-world data? A review of definitions based on literature and stakeholder interviews. Value Health 2017;20:858–65. https://doi.org/10.1016/j.jval.2017.03.008.

[2] Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, et al. Real-world evidence – what is it and what can it tell us? N Engl J Med 2016;375:2293–7. https://doi.org/10.1056/nejmsb1609216.

[3] GetReal consortium. RWE navigator – real-world evidence. Available from: ⟨https://rwe-navigator.eu/use-real-world-evidence/rwe-importance-in-medicine-development/⟩. [Accessed 19 October 2022].

[4] GetReal consortium. RWE navigator – sources of real-world data. Available from: ⟨https://rwe-navigator.eu/use-real-world-evidence/sources-of-real-world-data/⟩. [Accessed 19 October 2022].

[5] National Institutes of Health (NIH) Pragmatic Trials Collaboratory. Common real-world data sources. Available from: ⟨https://rethinkingclinicaltrials.org/chapters/conduct/acquiring-real-world-data/common-real-world-data-sources/⟩. [Accessed 19 October 2022].

[6] European Medicines Agency (EMA). Guideline on registry-based studies. 2021 Available from: [Accessed 19 October 2022]. ⟨https://www.ema.europa.eu/documents/scientific-guideline/guideline-registry-based-studies_en-0.pdf⟩.

[7] Food and Drug Administration (FDA). Real-world data: assessing registries to support regulatory decision-making for drug and biological products – guidance for industry. 2021 Available from: [Accessed 19 October 2022]. ⟨https://www.fda.gov/media/154449/download⟩.

[8] De Lusignan S, Crawford L, Munro N. Creating and using real-world evidence to answer questions about clinical effectiveness. BMJ Health Care Inform 2015;22:368–73. https://doi.org/10.14236/jhi.v22i3.177.

[9] Liu J, Barrett JS, Leonardi ET, Lee L, Roychoudhury S, Chen Y, et al. Natural history and real-world data in rare diseases: applications, limitations, and future perspectives. J Clin Pharmacol 2022;62:S38–55. https://doi.org/10.1002/jcph.2134.

[10] Cave A, Kurz X, Arlett P. Real-world data for regulatory decision making: challenges and possible solutions for Europe. Clin Pharmacol Ther 2019;106:36–9. https://doi.org/10.1002/cpt.1426.

[11] Innovative Medicines Initiative (IMI). EHDEN project factsheet. Available from: ⟨https://www.imi.europa.eu/projects-results/project-factsheets/ehden⟩. [Accessed 19 October 2022].

[12] Pulini AA, Caetano GM, Clautiaux H, Vergeron L, Pitts PJ, Katz G. Impact of real-world data on market authorization, reimbursement decision & price negotiation. Ther Innov Regul Sci 2021;55:228–38. https://doi.org/10.1007/s43441-020-00208-1.

[13] Flynn R, Plueschke K, Quinten C, Strassmann V, Duijnhoven RG, Gordillo-Marañon M, et al. Marketing authorization applications made to the European Medicines Agency in 2018–2019: what was the contribution of real-world evidence? Clin Pharmacol Ther 2022;111:90–7. https://doi.org/10.1002/cpt.2461.

[14] European Medicines Agency (EMA). Data Analysis and Real World Interrogation Network (DARWIN EU). Available from: ⟨https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu⟩. [Accessed 19 October 2022].

[15] Deloitte Center for Health Solutions. RWE focus is shifting to R &D, early investments begin to pay off. 2020 Available from: [Accessed 19 October 2022]. ⟨https://www2.deloitte.com/content/dam/insights/us/articles/6578_CHS-RWE-benchmarking-survey/DI_RWE%20benchmarking%20survey%20(SECURED).pdf⟩.

[16] Innovative Medicines Initiative (IMI). BD4BO programme factsheet. Available from: ⟨https://www.imi.europa.eu/projects-results/project-factsheets/bd4bo⟩. [Accessed 19 October 2022].

[17] Innovative Medicines Initiative (IMI). PIONEER project factsheet. Available from: ⟨https://www.imi.europa.eu/projects-results/project-factsheets/pioneer⟩. [Accessed 19 October 2022].

[18] Innovative Medicines Initiative (IMI). HARMONY project factsheet. Available from: ⟨https://www.imi.europa.eu/projects-results/project-factsheets/harmony⟩. [Accessed 19 October 2022].

[19] European Commission. IDEA4RC project factsheet. Available from: ⟨https://cordis.europa.eu/project/id/101057048⟩. [Accessed 30 January 2023].

[20] GetReal Institute. About us – facilitating the adoption and implementation of RWE in health care decision-making in Europe. Available from: ⟨https://www.getreal-institute.org/about-us/⟩. [Accessed 19 October 2022].

[21] Pregelj L, Hwang TJ, Hine DC, Siegel EB, Barnard RT, Darrow JJ, et al. Precision medicines have faster approvals based on fewer and smaller trials than other medicines. Health Aff (Millwood) 2018;37:724–31. https://doi.org/10.1377/hlthaff.2017.1580.

[22] Agrawal S, Arora S, Amiri-Kordestani L, de Claro RA, Fashoyin-Aje L, Gormley N, et al. Use of single-arm trials for US Food and Drug Administration drug approval in oncology, 2002–2021. JAMA Oncol 2023;9:266–72. https://doi.org/10.1001/jamaoncol.2022.5985.

[23] Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009–13. BMJ 2017;359:j4530. https://doi.org/10.1136/bmj.j4530.

[24] Kempf E, Bogaerts J, Lacombe D, Liu L. "Mind the gap" between the development of therapeutic innovations and the clinical practice in oncology: a proposal of the European Organisation for Research and Treatment of Cancer (EORTC) to optimise cancer clinical research. Eur J Cancer 2017;86:143–9. https://doi.org/10.1016/j.ejca.2017.08.028.

[25] Saesen R, Lejeune S, Quaglio G, Lacombe D, Huys I. Views of European Drug Development Stakeholders on treatment optimization and its potential for use in decision-making. Front Pharmacol 2020;11:43. https://doi.org/10.3389/fphar.2020.00043.

[26] Saesen R, Lacombe D, Huys I. Design, organisation and impact of treatment optimisation studies in breast, lung and colorectal cancer: the experience of the European Organisation for Research and Treatment of Cancer. Eur J Cancer 2021;151:221–32. https://doi.org/10.1016/j.ejca.2021.04.012.
[27] Booth CM, Karim S, Mackillop WJ. Real-world data: towards achieving the achievable in cancer care. Nat Rev Clin Oncol 2019;16:312–25. https://doi.org/10.1038/s41571-019-0167-7.
[28] Hall PS. Real-world data for efficient health technology assessment. Eur J Cancer 2017;79:235–7. https://doi.org/10.1016/j.ejca.2017.04.003.
[29] Saesen R, Kantidakis G, Marinus A, Lacombe D, Huys I. How do cancer clinicians perceive real-world data and the evidence derived therefrom? Findings from an international survey of the European Organisation for Research and Treatment of Cancer. Front Pharmacol 2022;8:969778https://doi.org/10.3389/fphar.2022.969778.
[30] Oehrlein EM, Graff JS, Harris J, Perfetto EM. Patient-community perspectives on real-world evidence: enhancing engagement, understanding, and trust. Patient 2019;12:375–81. https://doi.org/10.1007/s40271-019-00356-z.
[31] Skovlund E, Leufkens HGM, Smyth JF. The use of real-world data in cancer drug development. Eur J Cancer 2018;101:69–76. https://doi.org/10.1016/j.ejca.2018.06.036.
[32] Greshock J, Lewi M, Hartog B, Tendler C. Harnessing real-world evidence for the development of novel cancer therapies. Trends Cancer 2020;6:907–9. https://doi.org/10.1016/j.trecan.2020.08.006.
[33] Collins R, Bowman L, Landray M, Peto R. The magic of randomization versus the myth of real-world evidence. N Engl J Med 2020;382:674–8. https://doi.org/10.1056/nejmsb1901642.
[34] Franklin JM, Schneeweiss S. When and how can real world data analyses substitute for randomized controlled trials? Clin Pharmacol Ther 2017;102:924–33. https://doi.org/10.1002/cpt.857.
[35] Ramagopalan SV, Simpson A, Sammon C. Can real-world data really replace randomised clinical trials? BMC Med 2020;18:13. https://doi.org/10.1186/s12916-019-1481-8.
[36] McGale P, Cutter D, Darby SC, Henson KE, Jagsi R, Taylor CW. Can observational data replace randomized trials? J Clin Oncol 2016;34:3355–7. https://doi.org/10.1200/jco.2016.68.8879.
[37] Gerstein HC, McMurray J, Holman RR. Real-world studies no substitute for RCTs in establishing efficacy. Lancet 2019;393:210–1. https://doi.org/10.1016/s0140-6736(18)32840-x.
[38] Di Maio M, Perrone F, Conte P. Real-world evidence in oncology: opportunities and limitations. Oncologist 2020;25:e746–52. https://doi.org/10.1634/theoncologist.2019-0647.
[39] Pignatti F, Wilking U, Postmus D, Wilking N, Delgado J, Bergh J. The value of anticancer drugs – a regulatory view. Nat Rev Clin Oncol 2022;19:207–15. https://doi.org/10.1038/s41571-021-00584-z.
[40] Pacurariu A, Plueschke K, McGettigan P, Morales DR, Slattery J, Vogl D, et al. Electronic healthcare databases in Europe: descriptive analysis of characteristics and potential for use in medicines regulation. BMJ Open 2018;8:e023090https://doi.org/10.1136/bmjopen-2018-023090.
[41] de Leeuw DC, Ossenkoppele GJ, Janssen JJWM. Older patients with acute myeloid leukemia deserve individualized treatment. Curr Oncol Rep 2022;24:1387–400. https://doi.org/10.1007/s11912-022-01299-9.
[42] Concato J, Corrigan-Curay J. Real-world evidence – where are we now? N Engl J Med 2022;386:1680–2. https://doi.org/10.1056/nejmp2200089.
[43] Ford I, Norrie J. Pragmatic trials. N Engl J Med 2016;375:454–63. https://doi.org/10.1056/nejmra1510059.
[44] Sedgwick P. Explanatory trials versus pragmatic trials. BMJ 2014;349:g6694. https://doi.org/10.1136/bmj.g6694.
[45] Patsopoulos NA. A pragmatic view on pragmatic trials. Dialogues Clin Neurosci 2011;13:217–24. https://doi.org/10.31887/dcns.2011.13.2/npatsopoulos.
[46] Giraudeau B, Caille A, Eldridge SM, Weijer C, Zwarenstein M, Taljaard M. Heterogeneity in pragmatic randomised trials: sources and management. BMC Med 2022;20:372. https://doi.org/10.1186/s12916-022-02569-w.
[47] Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology. J Natl Cancer Inst 2017;109:djx187. https://doi.org/10.1093/jnci/djx187.
[48] Gyawali B, Parsad S, Feinberg BA, Nabhan C. Real-world evidence and randomized studies in the precision oncology era: the right balance. JCO Precis Oncol 2017;1:1–5. https://doi.org/10.1200/po.17.00132.
[49] Karim S, Booth CM. Effectiveness in the absence of efficacy: cautionary tales from real-world evidence. J Clin Oncol 2019;37:1047–50. https://doi.org/10.1200/jco.18.02105.
[50] Muirhead R, Aggarwal A. Real world data – does it cut the mustard or should we take it with a pinch of salt? Clin Oncol (R Coll Radiol) 2022;35:15–9. https://doi.org/10.1016/j.clon.2022.09.058.
[51] Dahabreh IJ, Robins JM, Hernán MA. Benchmarking observational methods by comparing randomized trials and their emulations. Epidemiology 2020;31:614–9. https://doi.org/10.1097/ede.0000000000001231.
[52] Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. Br J Cancer 2014;110:551–5. https://doi.org/10.1038/bjc.2013.725.
[53] European Organisation for Research and Treatment of Cancer (EORTC). Annual report 2021. 2022 Available from: [Accessed 19 October 2022]. ⟨https://www.eortc.org/app/uploads/2022/07/EORTC-annual-report-online-version.pdf⟩.
[54] Bean J, Lacombe D, Meunier F. Fifty years of progress against cancer – the European Organisation for Research and Treatment of Cancer celebrates its achievements. Eur Oncol Haematol 2012;8:10–3. https://doi.org/10.17925/eoh.2012.08.01.10.
[55] Velikova G, Coens C, Efficace F, Greimel E, Groenvold M, Johnson C, et al. Health-related quality of life in EORTC clinical trials – 30 years of progress from methodological developments to making a real impact on oncology practice. Eur J Cancer Suppl 2012;10:141–9. https://doi.org/10.1016/s1359-6349(12)70023-x.
[56] Benbow JH, Rivera DR, Lund JL, Feldman JE, Kim ES. Increasing inclusiveness of patient-centric clinical evidence generation in oncology: real-world data and clinical trials. Am Soc Clin Oncol Educ Book 2022;42:116–26. https://doi.org/10.1200/edbk_350574.
[57] Hernán MA, Hsu J, Healy B. A second chance to get causal inference right: a classification of data science tasks. Chance 2019;32:42–9. https://doi.org/10.1080/09332480.2019.1579578.
[58] European Organisation for Research and Treatment of Cancer (EORTC). SPECTA platform. Available from: ⟨https://spectaplatform.org/platform/⟩. [Accessed 19 October 2022].
[59] European Organisation for Research and Treatment of Cancer (EORTC). E^2-RADIatE platform. Available from: ⟨https://project.eortc.org/e2-radiate/platform/⟩. [Accessed 19 October 2022].
[60] Relton C, Torgerson D, O'Cathain A, Nicholl J. Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design. BMJ 2010;340:963–7. https://doi.org/10.1136/bmj.c1066.
[61] van der Velden JM, Verkooijen HM, Ayoung-Afat D, Burbach JPM, van Vulpen M, Relton C, et al. The cohort multiple randomized controlled trial design: a valid and efficient alternative to pragmatic trials. Int J Epidemiol 2017;46:96–102. https://doi.org/10.1093/ije/dyw050.

[62] Couwenberg AM, Burbach JPM, May AM, Berbee M, Intven MPW, Verkooijen HM. The trials within cohorts design facilitated efficient patient enrollment and generalizability in oncology setting. J Clin Epidemiol 2020;120:33–9. https://doi.org/10.1016/j.jclinepi.2019.12.015.

[63] European Organisation for Research and Treatment of Cancer (EORTC). OligoCare (EORTC 1822-RP). Available from: ⟨https://project.eortc.org/e2-radiate/cohorts/#0ce6202b33b250989⟩. [Accessed 19 October 2022].

[64] Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol 2016;183:758–64. https://doi.org/10.1093/aje/kwv254.

[65] Hernán MA. Methods of public health research – strengthening causal inference from observational data. N Engl J Med 2021;385:1345–8. https://doi.org/10.1056/nejmp2113319.

[66] Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol 2016;79:70–5. https://doi.org/10.1016/j.jclinepi.2016.04.014.

[67] Karcher H, Fu S, Meng J, Ankarfeldt MZ, Efthimiou O, Belger M, et al. The "RCT augmentation": a novel simulation method to add patient heterogeneity into phase III trials. BMC Med Res Methodol 2018;18:75. https://doi.org/10.1186/s12874-018-0534-6.

[68] Bonvalot S, Gronchi A, Le Péchoux C, Swallow CJ, Strauss D, Meeus P, et al. Preoperative radiotherapy plus surgery versus surgery alone for patients with primary retroperitoneal sarcoma (EORTC-62092: STRASS): a multicentre, open-label, randomised, phase 3 trial. Lancet Oncol 2020;21:1366–77. https://doi.org/10.1016/s1470-2045(20)30446-0.

[69] Callegaro D, Raut CP, Ajayi T, Strauss D, Bonvalot S, Ng D, et al. Preoperative radiotherapy in patients with primary retroperitoneal sarcoma: EORTC-62092 trial (STRASS) versus off-trial (STREXIT) results. Ann Surg 2022. https://doi.org/10.1097/sla.0000000000005492. [Epub ahead of print].

[70] Istl AC, Gronchi A. Neoadjuvant therapy for primary resectable retroperitoneal sarcomas—looking forward. Cancers (Basel) 2022;14:1831. https://doi.org/10.3390/cancers14071831.