

# Real World Data Studies of Antineoplastic Drugs: How Can They Be Improved to Steer Everyday Use in the Clinic?

Gincy George<sup>1</sup>, Beth Russell<sup>1</sup>, Anne Rigg<sup>2</sup>, Anthony CC Coolen<sup>3</sup>, Mieke Van Hemelrijck<sup>1</sup> 

<sup>1</sup>Translational Oncology and Urology Research, King's College London, London, UK; <sup>2</sup>Guy's Cancer Centre, Guy's and St Thomas NHS Foundation Trust, London, UK; <sup>3</sup>Department of Biophysics, Radboud University, Nijmegen, the Netherlands

Correspondence: Gincy George, King's College London, Translational Oncology and Urology Research, London, SE1 9RT, United Kingdom, Email [gincy.e.george@kcl.ac.uk](mailto:gincy.e.george@kcl.ac.uk)

**Abstract:** There is a growing interest in real world evidence when developing antineoplastic drugs owing to the shorter length of time and low costs compared to randomised controlled trials. External validity of studies in the regulatory phase can be enhanced by complementing randomised controlled trials with real world evidence. Furthermore, the use of real world evidence ensures the inclusion of patients often excluded from randomised controlled trials such as the elderly, certain ethnicities or those from certain geographical areas. This review explores approaches in which real world data may be integrated with randomised controlled trials. One approach is by using big data, especially when investigating drugs in the antineoplastic setting. This can even inform artificial intelligence thus ensuring faster and more precise diagnosis and treatment decisions. Pragmatic trials also offer an approach to examine the effectiveness of novel antineoplastic drugs without evading the benefits of randomised controlled trials. A well-designed pragmatic trial would yield results with high external validity by employing a simple study design with a large sample size and diverse settings. Although randomised controlled trials can determine efficacy of antineoplastic drugs, effectiveness in the real world may differ. The need for pragmatic trials to help guide healthcare decision-making led to the development of trials within cohorts (TWICs). TWICs make use of cohorts to conduct multiple randomised controlled trials while maintaining characteristics of real world data in routine clinical practice. Although real world data is often affected by incomplete data and biases such as selection and unmeasured biases, the use of big data and pragmatic approaches can improve the use of real world data in the development of antineoplastic drugs that can in turn steer decision-making in clinical practice.

**Keywords:** real world data, randomised controlled trials, big data, pragmatic trials, TWICs

## Gold Standard – A Constantly Evolving Perceptive In Research

The translation of real world data to real world evidence has been a hot topic for several years now. However, randomised controlled trials remain the “gold standard” especially in the drug development phase of antineoplastic or anticancer drugs that include cytotoxic drugs, hormones, and signal transduction inhibitors.<sup>1</sup> Whereas efficacy and safety assessments of most antineoplastic drugs rely heavily on randomised controlled trials, there is a growing interest in the use of observational data or real world evidence owing to the low costs and shorter length of time required by the latter in comparison to the traditional drug approval processes.<sup>2</sup>

Although randomised controlled trials remain the gold standard during the drug development process, complementing randomised controlled trials with real world evidence especially in the regulatory phase can vastly increase the external validity and generate evidence on treatment effects observed in the real world.<sup>2</sup>

There is a need to reconsider what is defined as the gold standard in all phases of drug development. We need to evolve our perceptions and educate ourselves to apply real world evidence into practice as many (especially within the clinical community) still do not feel confident in real world evidence outcomes other than results from registry data.<sup>3–5</sup> Decision-making in the drug development phase needs to translate real world data into real world evidence which should

be integrated with randomised controlled trials.<sup>2</sup> This should be defined as the new “gold” standard that we aspire towards and the gold standard in this context should be a constantly evolving perspective across both the research and clinical communities. The aim of this review will be to discuss and explore how real world data may be integrated with randomised controlled trials.

## Big Data In Healthcare Research

The term “big data” has been used exponentially over the past few decades. Big data is not only relevant in the healthcare setting but in other industries including finance, advertising, and retail. There exist numerous definitions of big data, however, essentially, it refers to the collection of a large volume of data, which, in the healthcare setting, can be used to facilitate decision-making on the diagnosis and treatment of patients. In theory, big data is the future of medicine as it can inform artificial intelligence (AI) models thus saving clinician time and ensuring more accurate diagnosis and treatment decisions, particularly in the field of cancer. However, in practice, the innate heterogeneity in data poses a significant issue.

The collection of healthcare data takes many forms from electronic healthcare records (EHRs), claims data and registries. Within each of these methods of data collection, the variables themselves can differ substantially and the way in which the variables are coded can also differ. Even when using the data, or in randomised controlled trials, studies themselves are at times poorly designed and can have ill-defined endpoints leading to heterogenous outcomes which are not comparable to other studies.<sup>6,7</sup> To help overcome these issues, core outcome sets (COS) and common data models (CDM) may be applied. COS are an agreed standardised set of outcomes which should be reported as a minimum within a clinical trial.<sup>8,9</sup> Initiatives like core outcome measures in effectiveness trials (COMET) are paving the way to the development and application of COS in the real world.<sup>10</sup> Moreover, other cancer specific initiatives like PIONEER, which is part of the innovative medicine initiative’s (IMI’s) “Big Data for Better Outcomes” (BD4BO), have built the development of COS into their projects. It is encouraging to see that the COS within PIONEER are endorsed by the European Association of Urology, a reflection of the importance of COS in such projects.<sup>11</sup> COS are therefore a useful tool to ensure that the correct data is collected. However, the coding of such data can differ enormously from dataset to dataset.

CDMs offer a possible solution to standardise the structure of observational data to aid in the analysis of multiple large-scale datasets. CDMs like the one developed by The observational health data sciences and informatics (OHDSI), called the observational medical outcomes partnership (OMOP) CDM are just one example of such an idea.<sup>12</sup> The idea behind the OMOP CDM is to facilitate the analysis of data from datasets originally built for different purposes, eg, claims data and EHR data, for one common purpose of generating evidence for clinical research.

## Artificial Intelligence – Is It Worth The Hype?

Although big data can provide the platform for more standardised and structured data, heterogeneity and ethical considerations are still an issue when it comes to real world data. AI techniques have been emerging to combat the challenges of traditional methods for analysis of biological data, as it can often result in overfitting in clinical datasets. AI using machine learning approaches in biomedical research can include genetic algorithms to calibrate complex models of clinical data while preserving biological heterogeneity.<sup>13</sup> The enormity of the multi-omics and biological sequencing data has opened AI to a field of genomic and proteomic data analysis. The Basic Local Alignment Search Tool (BLAST) is an example of a robust algorithm for initial analysis of each DNA sequence, identifying areas of similarity, and mapping sequences to specific sites in a reference genome. The application even calculates the statistical significance by comparing nucleotide or protein sequences to sequence databases.<sup>14</sup> Other such examples include the expressed sequence tag (EST)<sup>15</sup> which is comprised of raw transcriptome data and the Sequence Read Archive (SRA)<sup>16</sup> database which holds raw DNA, complementary DNA, and RNA sequencing data along with alignment data.

However, AI has so far failed in risk prediction models due to its struggle to distinguish between real and random data patterns when data sets are modest in size, its inability to incorporate subtleties of medical data such as the effects of informative missingness, censoring by competing risks, or confounding factors, and by its limited ability to quantify the reliability of its predictions. Moreover, extensive algorithm refinement on a small sample of data (particularly in the real

world) can result in a sample-specific refined algorithm with limited generalisability to larger, diverse datasets.<sup>17</sup> An AI algorithm developed in this manner would, therefore, memorise the noise and statistical variations of the restricted sample leading to overfitting. Overfitting occurs when the number of parameters is too large compared to the number of data points available to determine clinical outcome predictions.

Modern cancer data is high-dimensional and multi-modal which requires complex mathematical models that specifically target clinical outcome prediction. Coolen et al suggests replica analysis for undoing the effects of overfitting in studies that focus on risk and outcome prediction, including (in ongoing work) for competing risk and time-to-event analyses.<sup>18–20</sup> When studying the risk of developing a disease, Bayesian models that allow for latent heterogeneity enable one to identify classes of individuals that differ in their susceptibility to a particular disease, how they react to a specific risk factor, or how they respond to a particular therapy. Further investigation of these individuals can therefore help to uncover indicators for either preventative or therapeutic measures.<sup>21</sup> Therefore, replica analysis and further AI techniques may be useful for prediction models in the real world settings.

## Efficacy And Effectiveness In Real World Data

Whilst randomised controlled trials are the gold standard in healthcare research, particularly when looking at new antineoplastic treatments, they are conducted in a highly controlled setting with strict inclusion and exclusion criteria. Once a drug or new treatment has been deemed beneficial through sufficient randomised controlled trials, the performance of that drug can differ greatly in the real world. This is because, patients who are otherwise not eligible to take part in or are under-represented within clinical trials are suddenly given a treatment which has not been tested on their demographic. Patients of ethnic minorities, older patients or patients with disabilities are just some of the examples of patients not sufficiently represented within clinical trials, which has also been acknowledged by the Food and Drug Administration in its draft guidance to enrol more patients from underrepresented racial and ethnic populations.<sup>22,23</sup> This is why we must consider both the efficacy and effectiveness of any drug or new treatment. The efficacy of a treatment relates to how well the treatment performs when the ideal conditions are met ie, in a randomised controlled trial.<sup>24,25</sup> Effectiveness, however, is how well the treatment performs in a real-world setting. Pragmatic trials offer a way to look at the effectiveness of new treatments, without losing all the benefits of randomised controlled trials.

## Need for Pragmatic Trials – Trials Within Cohort (TWICs)

The need for pragmatic trials arose because many of the randomised controlled trials failed to adequately inform routine clinical practice since they were optimised to assess efficacy. A pragmatic trial aims to determine the ideal care option as opposed to an explanatory trial where causal hypotheses are investigated. Whereas explanatory trials focus on homogeneity to reduce errors and biases, pragmatic trials aim to maximise heterogeneity in all aspects. A well-designed pragmatic trial would have the following characteristics: high external validity, large sample size, simple design, diverse settings and would include mostly Phase IV trials.<sup>26</sup> Both pragmatic and explanatory aspects exist in most randomised clinical trials. This review will explore two such methodologies for conducting pragmatic trials – pragmatic-explanatory continuum indicator summary (PRECIS) and trials within cohorts (TWICs). The PRECIS tool with 10 domains from eligibility criteria to primary outcomes was developed by Thorpe et al to help researchers design trials that take both aspects into account.<sup>27</sup>

TWICs evolved from the need for pragmatic trials to better inform decision-making in healthcare. TWICs was first introduced as “cohort multiple randomised controlled trials” designed by Clare Relton et al in 2010.<sup>28</sup> TWICs make use of cohorts to conduct multiple randomised controlled trials without compromising the precision of randomisation while retaining the characteristics of real world data in routine clinical practice. It was developed to address shortcomings relating to recruitment, treatment comparisons and ethics associated with existing randomised controlled trials design.<sup>28</sup>

The Graham Roberts Study was the first single-centre TWICs study in bladder cancer that created an observational cohort of bladder cancer patients with longitudinal measurements of cancer characteristics and outcomes recorded.<sup>29</sup> Following consent to the study, patients were given validated self-administered questionnaires which aimed to collect patient-reported outcome measures including quality of life, fatigue, anxiety and depression, physical activity, and dietary habits. This information was collected at baseline and every 12 months thereafter for at least 10 years.

The Utrecht cohort for Multiple BREast cancer intervention studies and Long-term evaLuAtion (UMBRELLA) Fit trial within the UMBRELLA breast cancer cohort also used the TWICs framework. At cohort entry, breast cancer patients provided consent for randomisation in future trials. For the UMBRELLA Fit trial, inactive patients were randomised 12–18 months after cohort enrolment. Whereas the intervention group ( $n = 130$ ) were offered a 12-week exercise intervention, the control group ( $n = 130$ ) received usual care.<sup>30</sup> Although the offer of an exercise intervention resulted in no improvement on quality of life, it gave more of an insight into exercise intervention in this group as 52% of inactive women with breast cancer accepted the intervention. A modest reduction in fatigue levels was also noted in the intervention group.

Although the TWICs design enhances study recruitment logistics, there are some limitations due to non-compliance owing to refusal of an intervention when offered one at the start of a study, which attenuates the estimated effect of the intervention under investigation.<sup>31</sup> Moreover, if an intervention measurement coincides or overlaps with a cohort measurement as was the case in the UMBRELLA study,<sup>30</sup> self-reported measurements may be completed with trial participation in mind thus diluting the effect size of the intervention. To employ cohort measurements as pre- and post-intervention measurements, thorough intervention planning is therefore necessary in between cohort assessments.

## Quest For The Perfect Study Design

One of the biggest challenges that affect real world evidence are biases including unmeasured confounding and confounding by indication.<sup>32</sup> Confounding by indication or channeling bias is a major concern when physician prescription behaviour is influenced by patient characteristics rather than random allocation. Moreover, choosing the wrong controls in a case-control study and accounting for the wrong factors or confounders in a study can result in reporting non-existent associations.<sup>33</sup> A randomised controlled trial can account for these biases to a certain extent by using methods such as randomisation, blinding and concealment. Identification and investigation of treatment effects and its extent are feasible in a randomised controlled trial without knowing the risk factors associated with the treatment. However, randomised controlled trials are not without their own limitations.

Randomised controlled trials and real world evidence must complement one another. However, the biggest challenge with this is that findings of randomised controlled trials and real world evidence vary significantly in some cases. Whereas real world data showed a lower risk of developing cardiovascular disease in men with prostate cancer on long-term androgen deprivation therapy, no such associations were found in randomised controlled trials.<sup>34,35</sup> Such contradictory findings largely depend on the varying study populations and metrics used in the two designs. The Food and Drug Administration Demonstration Project therefore seeks to gain insights on factors that define clinical questions that can be evaluated using real world data, which will lay the foundation for future substitutions of randomised controlled trials with real world evidence.<sup>36</sup>

## Conclusions

Real world data is often limited as it is affected by incomplete data, selection bias and residual or unmeasured confounding due to a lack of randomisation, which means they are far from being classed as substitutes for randomised controlled trials in studies of antineoplastic drugs. However, large, well-established patient cohort studies with consistent methodological rigour and AI methods designed to accommodate disease risk prediction models can complement pragmatically designed randomised controlled trials and can ultimately help bridge the gap between randomised controlled trials and clinical practice.

## Abbreviations

AI, artificial intelligence; EHR, electronic healthcare record; COS, core outcome sets; CDM, common data models; COMET, core outcome measures in effectiveness trials; IMI, innovative medicine initiative; BD4BO, big data for better outcomes; OHDSI, observational health data sciences and informatics; OMOP, observational medical outcomes partnership; TWICs, trials within cohorts; PRECIS, pragmatic-explanatory continuum indicator summary; UMBRELLA, Utrecht cohort for Multiple BREast cancer intervention studies and Long-term evaLuAtion.

## Disclosure

The authors report no conflicts of interest in this work. This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## References

- National Institute for Occupational Safety and Health. Antineoplastic (Chemotherapy) drugs – reproductive health [Internet]; 2022. Available from: <https://www.cdc.gov/niosh/topics/repro/antineoplastic.html#:~:text=Antineoplasticdrugsaremedicationsused,forms%2Cincludingliquidsorpills>. Accessed August 29, 2023.
- Skovlund E, Leufkens HGM, Smyth JF. The use of real-world data in cancer drug development. *Eur J Cancer*. 2018;101:69–76. doi:10.1016/j.ejca.2018.06.036
- Villines TC, Cziraky MJ, Amin AN. Awareness, knowledge, and utility of RCT data vs RWE: results from a survey of US cardiologists: real-world evidence in clinical decision making. *Clin Med Insights Cardiol*. 2020;14:1179546820953410. doi:10.1177/1179546820953410
- Saesen R, Lacombe D, Huys I. Real-world data in oncology: a questionnaire-based analysis of the academic research landscape examining the policies and experiences of the cancer cooperative groups. *ESMO Open*. 2023;8(2):100878. doi:10.1016/j.esmoop.2023.100878
- Saesen R, Kantidakis G, Marinus A, et al. 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;13:1–19. doi:10.3389/fphar.2022.969778
- Delgado A, Guddati AK. Clinical endpoints in oncology - a primer. *Am J Cancer Res*. 2021;11(4):1121–1131.
- Khambholja K, Gehani M. Use of structured template and reporting tool for real-world evidence for critical appraisal of the quality of reporting of real-world evidence studies: a systematic review. *Value Heal*. 2023;26(3):427–434. doi:10.1016/j.jval.2022.09.003
- Webbe J, Sinha I, Gale C. Core outcome sets. *Arch Dis Child Educ Pract Ed*. 2018;103(3):163–166. doi:10.1136/archdischild-2016-312117
- Williamson PR, Altman DG, Blazeby JM, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials*. 2012;13(1):1–8. doi:10.1186/1745-6215-13-132
- Initiative C. COMET initiative - core outcome measures in effectiveness trials [Internet]. Available from: <https://www.comet-initiative.org/>. Accessed August 29, 2023.
- PIONEER - European Network of Excellence for big data in prostate cancer [Internet]; 2019. Available from: <https://prostate-pioneer.eu/>. Accessed August 29, 2023.
- OHDSI. Standardized data: the OMOP common data model [Internet]; 2023. Available from: <https://www.ohdsi.org/data-standardization/>. Accessed August 29, 2023.
- Cockrell C, An G. Utilizing the heterogeneity of clinical data for model refinement and rule discovery through the application of genetic algorithms to calibrate a high-dimensional agent-based model of systemic inflammation. *Front Physiol*. 2021;12:662845. doi:10.3389/fphys.2021.662845
- Basic local alignment search tool [Internet]. Available from: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>. Accessed August 29, 2023.
- Jongeneel CV. Searching the expressed sequence tag (EST) databases: panning for genes. *Brief Bioinform*. 2000;1(1):76–92. doi:10.1093/bib/1.1.76
- Leinonen R, Sugawara H, Shumway M. The sequence read archive. *Nucleic Acids Res*. 2011;39(Database issue):D19–D21. doi:10.1093/nar/gkq1019
- Kulkarni S, Seneviratne N, Baig MS, et al. Artificial intelligence in medicine: where are we now? *Acad Radiol*. 2020;27(1):62–70. doi:10.1016/j.acra.2019.10.001
- Coolen ACC, Sheikh M, Mozeika A, Aguirre-Lopez F, Antenucci F. Replica analysis of overfitting in generalized linear models. *J Phys a Math Theor*. 2020;53(365001):365001. doi:10.1088/1751-8121/aba028
- Sheikh M, Coolen ACC. Analysis of overfitting in the regularized Cox model. *J Phys a Math Theor*. 2019;52(38):384002. doi:10.1088/1751-8121/ab375c
- Coolen ACC, Barrett JE, Paga P, Perez-Vicente CJ. Replica analysis of overfitting in regression models for time-to-event data. *J Phys a Math Theor*. 2017;50(375001):38. doi:10.1088/1751-8121/aa812f
- Hägström C, Van Hemelrijck M, Garmo H, et al. Heterogeneity in risk of prostate cancer: a Swedish population-based cohort study of competing risks and Type 2 diabetes mellitus. *Int J Cancer*. 2018;143(8):1868–1875. doi:10.1002/ijc.31587
- Bleyer A. In and out, good and bad news, of generalizability of SWOG treatment trial results. *J Natl Cancer Inst*. 2014;106(3):dju027. doi:10.1093/jnci/dju027
- Administration USF& D. FDA takes important steps to increase racial and ethnic diversity in clinical trials; 2022. Available from: <https://www.fda.gov/news-events/press-announcements/fda-takes-important-steps-increase-racial-and-ethnic-diversity-clinical-trials>. Accessed August 29, 2023.
- Galsky MD, Oh WK. Mind the gap: efficacy versus effectiveness and pivotal prostate cancer clinical trial demographics. *Cancer*. 2014;120(19):2944–2945. doi:10.1002/cncr.28808
- Nilsson MP, Winter C, Kristofferson U, et al. Efficacy versus effectiveness of clinical genetic testing criteria for BRCA1 and BRCA2 hereditary mutations in incident breast cancer. *Fam Cancer*. 2017;16(2):187–193. doi:10.1007/s10689-016-9953-x
- Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci*. 2011;13(2):217–224. doi:10.31887/DCNS.2011.13.2/natsopoulos
- Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol*. 2009;62(5):464–475. doi:10.1016/j.jclinepi.2008.12.011
- Relton C, Torgerson D, O’Cathain A, et al. Rethinking pragmatic randomised controlled trials: introducing the “cohort multiple randomised controlled trial” design. *BMJ*. 2010;340(7753):963–967. doi:10.1136/bmj.c1066
- Wylie H, Cahill F, Santaolalla A, et al. Graham Roberts Study protocol: first “trials within cohort study” for bladder cancer. *BMJ Open*. 2019;9(9):1–5. doi:10.1136/bmjopen-2019-029468
- Gal R, Monnkhof EM, van Gils CH, et al. Effects of exercise in breast cancer patients: implications of the trials within cohorts (TwICs) design in the UMBRELLA Fit trial. *Breast Cancer Res Treat*. 2021;190(1):89–101. doi:10.1007/s10549-021-06363-9

31. Gal R, Monninkhof EM, van Gils CH, et al. The Trials within Cohorts design faced methodological advantages and disadvantages in the exercise oncology setting. *J Clin Epidemiol.* 2019;113:137–146. doi:10.1016/j.jclinepi.2019.05.017
32. Blais L, Ernst P, Suissa S. Confounding by indication and channeling over time: the risks of beta 2-agonists. *Am J Epidemiol.* 1996;144(12):1161–1169. doi:10.1093/oxfordjournals.aje.a008895
33. Cinelli C, Forney A, Pearl J. A crash course in good and bad controls. *Sociol Methods Res.* 2020. doi:10.2139/ssrn.3689437
34. Bosco C, Bosnyak Z, Malmberg A, et al. Quantifying observational evidence for risk of fatal and nonfatal cardiovascular disease following androgen deprivation therapy for prostate cancer: a meta-analysis. *Eur Urol.* 2015;68(3):386–396. doi:10.1016/j.eururo.2014.11.039
35. Nguyen PL, Je Y, Schutz FAB, et al. Association of androgen deprivation therapy with cardiovascular death in patients with prostate cancer: a meta-analysis of randomized trials. *JAMA.* 2011;306(21):2359–2366. doi:10.1001/jama.2011.1745
36. Franklin JM, Patorno E, Desai RJ, et al. Emulating randomized clinical trials with nonrandomized real-world evidence studies. *Circulation.* 2021;143(10):1002–1013. doi:10.1161/CIRCULATIONAHA.120.051718

## Pragmatic and Observational Research

Dovepress

### Publish your work in this journal

Pragmatic and Observational Research is an international, peer-reviewed, open access journal that publishes data from studies designed to reflect more closely medical interventions in real-world clinical practice compared with classical randomized controlled trials (RCTs). The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/pragmatic-and-observational-research-journal>