

# A Process for the Emulation of Comparative Oncology Trials with Real-world Evidence (ENCORE)

**Authors:** Janick Weberpals<sup>1</sup>, Kenneth L. Kehl<sup>2</sup>, Donna R. Rivera<sup>3</sup>, ..., Sebastian Schneeweiss<sup>1</sup>, Shirley V. Wang<sup>1</sup>

Author affiliations:

<sup>1</sup> Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

<sup>2</sup> Dana-Farber, ..., USA

<sup>3</sup> Oncology Center of Excellence, US Food and Drug Administration, Silver Spring, MD, USA

Correspondence:

Shirley V. Wang, PhD

Division of Pharmacoepidemiology and Pharmacoeconomics,

Department of Medicine, Brigham and Women's Hospital, Harvard Medical School,

1620 Tremont Street, Suite 3030-R, Boston, MA 02120, USA

Phone: +1 617-278-0932

Fax: + 1 617-232-8602

Email: [SWANG1@BWH.HARVARD.EDU](mailto:SWANG1@BWH.HARVARD.EDU)

Article type: Review

Word count: xxx words / xxx words

Tables: x

Figures: x

Supplementary material: Supplementary tables and figures

**Short running title:** ...

**Keywords:** Oncology, Real-World Evidence, Trial emulation, EHR

**Funding Statement:** This project was supported by ...

**Competing Interests Statement:** Dr. Weberpals is now an employee of AstraZeneca and owns stocks in AstraZeneca. Dr. Wang has consulted ad hoc for Exponent Inc. and MITRE a federally funded research center for the Centers for Medicare and Medicaid Services on unrelated work. Dr. Schneeweiss is participating in investigator-initiated grants to the Brigham and Women's Hospital from Boehringer Ingelheim, Takeda, and UCB unrelated to the topic of this study. He owns equity in Aetion Inc., a software manufacturer. He is an advisor to Temedica GmbH, a patient-oriented data generation company. His interests were declared, reviewed, and approved by the Brigham and Women's Hospital in accordance with their institutional compliance policies.

**Data sharing statement:** ...

**Analytic code sharing statement:** ...

**Proposed target journals:** *JCO CCI => JAMA Network Open => CPT*

*Manuscript last updated: 2024-12-04 14:18:23.087217*

## **Abstract**

xxx words/xxx words

...

## Background

Randomized controlled trials (RCTs) have been the gold standard for establishing the efficacy and safety of medical products. With the advent of the the 21<sup>st</sup> Century Cures Act directive<sup>1</sup>, the Food and Drug Administration (FDA) established a framework to increasingly consider real-world evidence (RWE) generated from routine-care health data such as electronic health records (EHR) to evaluate and contextualize the comparative safety and effectiveness of novel cancer therapies.<sup>2</sup> With 21% of all approvals, oncology was the disease area with the most FDA drug approvals in 2023,<sup>3</sup> and especially in the field of precision oncology, RWE has a large potential to complement evidence coming from RCTs. Potential use cases comprise the assessment of effectiveness in patient populations that are underrepresented in RCTs, the construction of external control arms in single-arm trials where active recruitment may not be feasible or the discovery of biomarkers among pan-tumor populations that harbor specific genomic and immuno-pathological signatures.

However, the validity and transportability of results derived between RWE studies and RCTs can depend on many factors and frequently referenced limitations include missing data, small sample sizes, data discontinuity<sup>4, 5</sup>, rapid changes in guideline treatment patterns and the inability to measure and emulate common eligibility criteria and prognostic factors in real-world data (RWD).<sup>6</sup> While there are already published examples of oncology trial emulations<sup>6-8</sup>, a systematic and scaled approach to emulate a diverse set of different oncology trials in various heterogeneous databases is necessary to gain confidence in the accuracy of RWE studies and to provide an answer as to which questions can be validly answered.

The RCT DUPLICATE initiative<sup>9</sup> increased our understanding of when RWE studies can come to causal conclusions on treatment effects by comparing results against RCTs under the assumption that each RCT finding reflects a causal treatment effect. In settings where the RCT designs could be emulated well, RWE studies came to the same conclusions.<sup>10</sup> However, prior work from RCT-DUPLICATE has focused primarily on emulating trials in the cardio-metabolic, renal, and pulmonary clinical areas using claims databases.

The *Emulation of Comparative Oncology Trials with Real-world Evidence* (**ENCORE**) project<sup>11</sup> aims to extend this work to the field of oncology which comes with its own unique set of challenges which must be systematically explored and understood. Building on a process co-developed with the FDA through RCT DUPLICATE<sup>9</sup>, this expansion to oncology is going to emulate 12 randomized clinical trials using multiple EHR data sources. The process includes an emphasis on transparency with documented assessment of data fitness of the RWD source for each trial<sup>12, 13</sup> and the conduct of extensive sensitivity analyses to assess robustness of findings and trial eligibility criteria.

The objectives of this project are to develop state-of-the-art methodological approaches and apply these to create insights that may provide guidance on the potential use of RWE for regulatory science in oncology. This includes the systematic evaluation of the suitability of data in relation to the study design and statistical analysis by emulating 12 oncology trials

across four cancers and assessing the agreement of treatment effect estimates between RCTs and their respective emulations.

In this process paper, we describe the design and process for the selection of the 12 oncology RCTs, the assessment of the database quality and selection, protocol development, study design and statistical analysis and final agreement metrics to evaluate the concordance between RCTs and emulations.

## Methods

A visual summary of the entire systematic process from trial selection to final results is provided in Figure 1.

### Trial selection

The focus of ENCORE is to maximize potential learnings on when RWE studies can or cannot yield similar results compared to RCTs. To that end, the emphasis of the project is on trials of therapies for the most common cancers and/or cancers for which there has been substantial therapeutic development in recent years. After careful review and exchange with clinical and regulatory experts, four cancer indications were identified including lung cancer, breast cancer, colorectal cancer and multiple myeloma. For each cancer we aim to conduct three trial emulations which will be implemented using multiple databases accessible for the scope of this project (i.e. the total number of emulations will equal 12 trials x  $n$  databases which are found fit-for-purpose for each trial).

The trial selection will follow a semi-automated process for which we will document the eligibility criteria resulting in a CONSORT diagram showing reasons for excluding RCTs. The search will be conducted using the AACT database which is a publicly available relational database developed and maintained by the Clinical Trials Transformation Initiative (CTTI) which contains all information (protocol and result data elements) about every study registered on ClinicalTrials.gov.<sup>14</sup> To identify eligible trials, we will use a combined search query strategy of the National Library of Medicine (NLM)-controlled *MeSH* term and a free keyword search for the respective cancer indication in the *conditions*, *studies* and *detailed\_descriptions* fields of each trial entry on ClinicalTrials.gov.

Eligible trials need to fulfill the following basic criteria:

- Interventional
- Randomized
- Intervention model: parallel assignment
- Industry-sponsored

- Trial start in 2011 or later
- Primary purpose was to study treatment effects
- Overall survival must be one of the endpoints reported
- Recruitment status: ‘Completed’ or ‘Active, not recruiting’
- Feasibility and clinical relevance

The rationale and operationalization of each criterion is listed in detail in Table 1. We will mainly consider pivotal interventional, randomized trials after 2011 since treatment guidelines among studied cancer indications have undergone significant changes in recent years. Due to the rapid adoption of new breakthrough therapies in routine care, it is unlikely to find patients who may be still treated with outdated treatment regimens in the real-world. In parallel, trial results should have also not been published too recently in order to allow for enough data and follow-up time accrual in databases used for this project. An further focus is on trials that have reported overall survival (OS) as one of the pre-specified endpoints in the protocol. Although there have been substantial methodological advancements to increase our understanding on the emulation and comparison of real-world progression-free survival (PFS) and objective response rates (ORR) to a RECISTv1.1<sup>15</sup>-based PFS and ORR assessment in RCTs<sup>16, 17</sup>, imaging-based evaluations still hold a level of granularity which may not be necessarily reflected in chart-abstracted assessments of a patient’s progression in routine care.<sup>18, 19</sup> Given the large number of other methodological challenges like missing data, small sample sizes, data discontinuity and rapidly changing guideline treatments, the scope of this project was to focus on the emulation of OS as the endpoint of interest.

While most trial-eligibility criteria can be operationalized in an automated fashion, the last criterion on emulation feasibility and clinical relevance involves extensive human review. The critical points considered in this step include a thorough feasibility assessment of the data fitness<sup>19</sup>, sample size considerations and the assessment if critical eligibility criteria (e.g., biomarker status) and prognostic factors (e.g., ECOG performance score) are measurable and can be balanced using propensity score matching or weighting methods.<sup>20</sup> Lastly, trial candidates are ranked and shortlisted into primary and runner-up candidates based on their clinical and regulatory relevance.

A list of tentative, shortlisted primary candidates is presented in Table 2 and the corresponding selection process is illustrated in the CONSORT diagrams (Supplementary Figures 1-4). Naturally, the majority trials will cover advanced or metastatic cancer populations since a large proportion of drug development efforts have focused on these settings in recent years. A key learning that we aim to foster with the shortlisted trials is to achieve a better understanding how different disease settings (early, late), line settings ([neo]adjuvant, first line, advanced lines of therapy), therapy protocols (monotherapy, combination therapy) and population characteristics (simple versus complex genetic or immunological signatures) can be emulated using RWD. If ongoing feasibility assessments indicate that these trials cannot be emulated with high enough confidence, runner-up candidates will be considered instead.

## Databases

The ENCORE project will utilize data from a total four different oncology-specific electronic health records (EHR)-derived data sources: ConcertAI, COTA, Flatiron Health, McKesson/Ontada. All available databases draw from a comprehensive national sample of patients with cancer in the US with detailed EHR-derived information on the information necessary to study medication effectiveness in oncology. For ENCORE, not all databases will be available for each cancer indication and the names of the databases will be blinded and referred to as ENCORE DataBase (EDB) 1, 2, 3 and 4 for the final reporting of results. If more than one database is considered fit-for-purpose for a respective trial emulation, the best possible analytic model will be employed for each database separately and final treatment effect estimates will be pooled using a meta-analytic approach.

## Protocol development

For each shortlisted and selected RCT, a detailed protocol, pre-specifying key elements of the trial emulation, will be developed following the HARPER protocol template<sup>21</sup> and registered on ClinicalTrials.gov. Following the target trial emulation framework, we will provide an explicit statement and rationale on how each element will be emulated including database selection, covariate measurement, operationalization of key eligibility criteria, study design, data analysis and causal contrasts of interest.<sup>22, 23</sup> Since it is common that oncology RCTs update OS estimates periodically based on accrued follow-up time, the protocol will give a brief summary of each emulated RCT and specify which target OS estimates will be used to compare agreement metrics to. All eligibility criteria will be extracted based on publicly available protocols and statistical analysis plans of the selected RCT. According to the *Structured Process to Identify Fit-For-Purpose Data* (SPIFD) framework<sup>13</sup>, tables will outline how eligibility criteria will be measured and operationalized and a color-coded heatmap will indicate the level of confidence on how well each criterion can be emulated in each selected database. As there are general eligibility criteria in oncology trials which either won't be possible to emulate (e.g., physician-assessed survival prognosis of xy months) or that are clinically not relevant for the emulation of the trial (e.g., male patients should be willing to use barrier contraception), the study team will decide on key eligibility criteria for the emulation of the trial. We will additionally provide a definition on how exposure and outcomes are defined and ascertained in each respective database. For the considered databases, the OS endpoint is a composite that is derived from different sources like EHR flags, social security death index, obituary and other linkages. Given that not all relevant sources that provide mortality data are synchronized and updated uniformly, sensitivity analyses with more conservative (i.e., earlier) censoring dates will be considered for each trial emulation to mitigate the potential impact of ghost-time bias.<sup>24</sup> Each final protocol draft will be reviewed by a clinical and FDA regulatory expert panel before registration on ClinicalTrials.gov.

Critical aspects when emulating oncology trials are the choice and estimation of the appropriate

estimand of interest.<sup>25</sup> Particularly when emulating pivotal trials of paradigm-changing treatments, multiple aspects need to be considered such as the contemporaneity of the (historical) control cohort, the adoption rate of the novel intervention in routine care, the magnitude of the clinical treatment benefit and the rate in which (particularly patients in the control arm) discontinue or cross-over to the interventional treatment, which could lastly bias emulated treatment effects towards the null. To that end, comprehensive diagnostics will be performed to contextualize these parameters and (if reported) draw comparisons to the emulated trial.

## **gn and statistical analysis**

26

## **Agreement metrics**



## Discussion

...

## Conclusions

...

## References

1. Framework for FDA’s real-world evidence program (last accessed 11/28/2024) [Internet], 2018 Available from: <https://www.fda.gov/media/120060/download?attachment>
2. Purpura CA, Garry EM, Honig N, et al: The role of real-world evidence in FDA-approved new drug and biologics license applications. *Clinical Pharmacology & Therapeutics* 111:135–144, 2022
3. Senior M: Fresh from the biotech pipeline: Record-breaking FDA approvals. *Nature Biotechnology*, 2024
4. Merola D, Schneeweiss S, Schrag D, et al: An algorithm to predict data completeness in oncology electronic medical records for comparative effectiveness research [Internet]. *Annals of Epidemiology* 76:143–149, 2022 Available from: <http://dx.doi.org/10.1016/j.annepidem.2022.07.007>
5. Joshua Lin K, Jin Y, Gagne J, et al: Longitudinal data discontinuity in electronic health records and consequences for medication effectiveness studies. *Clinical Pharmacology & Therapeutics* 111:243–251, 2022
6. Rider JR, Wasserman A, Slipski L, et al: Emulations of oncology trials using real-world data: A systematic literature review. *American journal of epidemiology* kwae346, 2024
7. Merola D, Campbell U, Gautam N, et al: The action coalition to advance real-world evidence through randomized controlled trial emulation initiative: oncology. *Clinical Pharmacology & Therapeutics* 113:1217–1222, 2023
8. Merola D, Campbell U, Lenis D, et al: Calibrating observational health record data against a randomized trial. *JAMA Network Open* 7:e2436535–e2436535, 2024
9. Wang SV, Schneeweiss S, Franklin JM, et al: [Emulation of randomized clinical trials with nonrandomized database analyses: Results of 32 clinical trials](#). *Jama* 329:1376–1385, 2023
10. Heyard R, Held L, Schneeweiss S, et al: Design differences and variation in results between randomised trials and non-randomised emulations: Meta-analysis of RCT-DUPLICATE data [Internet]. *BMJ medicine* 3, 2024 Available from: <https://doi.org/10.1136/bmjmed-2023-000709>
11. Calibrating real-world evidence studies in oncology against randomized trials: ENCORE (last accessed 11/28/2024) [Internet], 2024 Available from: <https://www.fda.gov/about-fda/oncology-center-excellence/calibrating-real-world-evidence-studies-oncology-against->

- 12.** Rivera DR, Eckert JC, Rodriguez-Watson C, et al: The oncology QCARD initiative: Fostering efficient evaluation of initial real-world data proposals. *Pharmacoepidemiology and Drug Safety* 33:e5818, 2024
- 13.** Gatto NM, Campbell UB, Rubinstein E, et al: The structured process to identify fit-for-purpose data: A data feasibility assessment framework. *Clinical Pharmacology & Therapeutics* 111:122–134, 2022
- 14.** Tasneem A, Aberle L, Ananth H, et al: The database for aggregate analysis of ClinicalTrials. Gov (AACT) and subsequent regrouping by clinical specialty. *PloS one* 7:e33677, 2012
- 15.** Eisenhauer EA, Therasse P, Bogaerts J, et al: New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European journal of cancer* 45:228–247, 2009
- 16.** Ton TGN, Pal N, Trinh H, et al: Replication of overall survival, progression-free survival, and overall response in chemotherapy arms of nonâ small cell lung cancer trials using real-world data [Internet]. *Clinical Cancer Research* 28:2844–2853, 2022Available from: <https://doi.org/10.1158/1078-0432.CCR-22-0471>
- 17.** McKelvey BA, Garrett-Mayer E, Rivera DR, et al: Evaluation of real-world tumor response derived from electronic health record data sources: A feasibility analysis in patients with metastatic non-small cell lung cancer treated with chemotherapy. *JCO Clinical Cancer Informatics* 8:e2400091, 2024
- 18.** Chen L, Davis R, Lee J, et al: Comparison of response from RECIST1.1 and abstraction in real-world patients with lung cancer. [Internet]. *Journal of Clinical Oncology* 41:e21194–e21194, 2023Available from: [https://ascopubs.org/doi/abs/10.1200/JCO.2023.41.16\\_suppl.e21194](https://ascopubs.org/doi/abs/10.1200/JCO.2023.41.16_suppl.e21194)
- 19.** Rivera DR, Henk HJ, Garrett-Mayer E, et al: The friends of cancer research real-world data collaboration pilot 2.0: Methodological recommendations from oncology case studies. *Clinical Pharmacology & Therapeutics* 111:283–292, 2022
- 20.** Brookhart MA, Schneeweiss S, Rothman KJ, et al: Variable selection for propensity score models. *American journal of epidemiology* 163:1149–1156, 2006
- 21.** Wang SV, Pottegård A, Crown W, et al: HARmonized protocol template to enhance reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE/ISPOR task force [Internet]. *Value in Health* 25:1663–

1672, 2022 Available from: <https://doi.org/10.1016/j.jval.2022.09.001>

**22.** Hernán MA, Wang W, Leaf DE: Target trial emulation: A framework for causal inference from observational data. *Jama* 328:2446–2447, 2022

**23.** Hernán MA, Sauer BC, Hernández-Díaz S, et al: [Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses](#). *Journal of clinical epidemiology* 79:70–75, 2016

**24.** Meyer A-M, Davies J, Taylor M, et al: Open cohorts and ghost-time bias in real world data, in *PHARMACOEPIDEMIOLOGY AND DRUG SAFETY*. WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2020, pp 426–426

**25.** Rufibach K: Treatment effect quantification for time-to-event endpoints—Estimands, analysis strategies, and beyond [Internet]. *Pharmaceutical Statistics* 18:145–165, 2018 Available from: <http://dx.doi.org/10.1002/pst.1917>

**26.** Weberpals J, Raman SR, Shaw PA, et al: A principled approach to characterize and analyze partially observed confounder data from electronic health records [Internet]. *Clinical Epidemiology* 16:329–343, 2024 Available from: <https://www.tandfonline.com/doi/abs/10.2147/CLEP.S436131>

## Tables

file:///tmp/RtmpM017SF/file14985621d2c8.html screenshot completed

Table 1: Criteria to select eligible trials for emulation in ENCORE.

Criteria	Definition	Eligible
Interventional study	The nature of the investigation or investigational use for which clinical study information is being submitted	Interventional (clinical trial): Participants are assigned prospectively to an intervention or interventions according to a protocol to evaluate the effect of the intervention(s) on biomedical or other health related outcomes.
Randomized allocation	The method by which participants are assigned to arms in a clinical trial.	Randomized: Participants are assigned to intervention groups by chance
Interventional study model	The strategy for assigning interventions to participants.	Parallel: Participants are assigned to one of two or more groups in parallel for the duration of the study
Sponsor/source	The entity (for example, corporation or agency) that initiates the study	Industry
Study start date	The estimated date on which the clinical study will be open for recruitment of participants, or the actual date on which the first participant was enrolled.	2011 or later
Primary purpose	The main objective of the intervention(s) being evaluated by the clinical trial.	Treatment: One or more interventions are being evaluated for treating a disease, syndrome, or condition.
Primary outcome	A description of each primary outcome measure (or for observational studies, specific key measurement[s] or observation[s] used to describe patterns of diseases or traits or associations with exposures, risk factors or treatment).	Primary or secondary outcome needs to include overall survival
Overall Recruitment Status	The recruitment status for the clinical study as a whole, based upon the status of the individual sites. If at least one facility in a multi-site clinical study has an Individual Site Status of "Recruiting," then the Overall Recruitment Status for the study must be "Recruiting."	Completed: The study has concluded normally; participants are no longer receiving an intervention or being examined (that is, last participant's last visit has occurred) OR Active, not recruiting: Study is continuing, meaning participants are receiving an intervention or being examined, but new participants are not currently being recruited or enrolled
Feasibility and clinical relevance	Are all key variables available to emulate the clinical trial at hand and is the clinical trial considered clinically relevant?	Trials for which there is reasonable believe that key study parameters can be emulated and there is a high enough clinical relevance (e.g., paradigm-changing trials)

file:///tmp/RtmpM017SF/file1498cc15e28.html screenshot completed

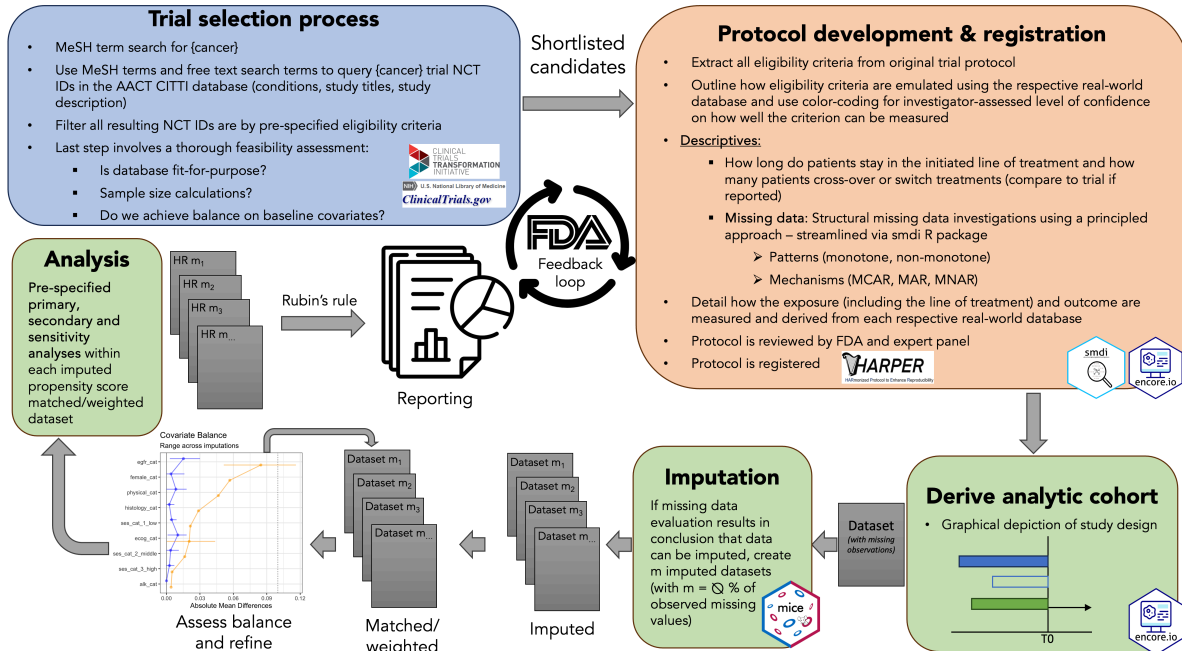
Table 2: Tentative list of randomized controlled trials (RCTs) considered for emulation.

NCTID	Acronym	Clinical setting	Line of therapy	Treatment comparison
<b>Non-small cell lung cancer</b>				
NCT02296125	FLAURA	Advanced/metastatic EGFRm+	1L	osimertinib versus erlotinib or gefitinib
NCT01673867	CheckMate017/057	Metastatic squamous/non-squamous	2L	nivolumab versus docetaxel
NCT03215706	CheckMate9LA	Metastatic	1L	nivolumab, ipilimumab, chemotherapy versus chemotherapy alone
<b>Breast cancer</b>				
NCT01740427	PALOMA-2	Advanced postmenopausal ER-positive and HER2-negative	1L	palbociclib, letrozole versus letrozole
NCT02819518	KEYNOTE-355	Locally recurrent inoperable or metastatic triple negative	1L	pembrolizumab, chemotherapy vs. placebo, chemotherapy
NCT01772472	KATHERINE	HER2-positive	Adjuvant	trastuzumab emtansine versus trastuzumab
<b>Colorectal cancer</b>				
NCT04737187	SUNLIGHT	Refractory metastatic	3L	trifluridine, tipiracil, bevacizumab versus trifluridine, tipiracil
NCT01374425	MAVERICC	Metastatic	1L	bevacizumab, mFOLFOX6 versus bevacizumab, FOLFIRI
NCT02563002	KEYNOTE-177	Metastatic microsatellite instability-high (MSI-H) or mismatch repair deficient (dMMR)	2L+	pembrolizumab versus standard of care
<b>Multiple Myeloma</b>				
NCT01568866	ENDEAVOR	Relapsing or progressing disease	2L/3L	carfilzomib, dexamethasone versus bortezomib, dexamethasone
NCT02252172	MAIA	Newly diagnosed	1L	daratumumab, lenalidomide, dexamethasone versus lenalidomide, dexamethasone
NCT01239797	ELOQUENT - 2	Relapsed or refractory	2L+	elotuzumab, lenalidomide, dexamethasone versus lenalidomide, dexamethasone



## Figures

Figure 1: Systematic process to understand effectiveness claims of oncology trials using real-world evidence.



[View figure in higher resolution here](#)

Warning in `get_plot_component(plot, "guide-box")`: Multiple components found; returning the first one. To return all, use ``return_all = TRUE``.

Saving 5.5 x 7 in image

Figure 2: Descriptive drug utilization analyses.

