Transportability of overall survival estimates from US to Canadian patients with advanced non–small cell lung cancer with implications for regulatory and health technology assessment

CCHE Seminar Series

Paul Arora
University of Toronto

paul.arora@utoronto.ca

Friday September 29, 2023, 10am-12pm

HS Room 412 and Zoom

CCHE CCES
Canadian Centre for Health Economics
Centre canadien en économie de la santé

Agenda

CCHE Seminar Series:

Paul Arora
University of Toronto

paul.arora@utoronto.ca

Friday September 29, 2023, 10am-12pm

HS Room 412 and Zoom

- Speaker Introduction
- RWE in HTA
- External Validity
- Transportability & Generalizability
- Study Background
- Data Sources
- Methods
- Results
- Discussion
- Future Work

Speaker Introduction

MSc Epidemiology	2005
PhD Epidemiology	2013
PHAC	2007-2014
C-GCH Sickkids	2014-2015
Lighthouse Outcomes	2016-2020
Cytel	2020-2023
DLSPH	Current

JAMA Network Open

Original Investigation | Statistics and Research Methods

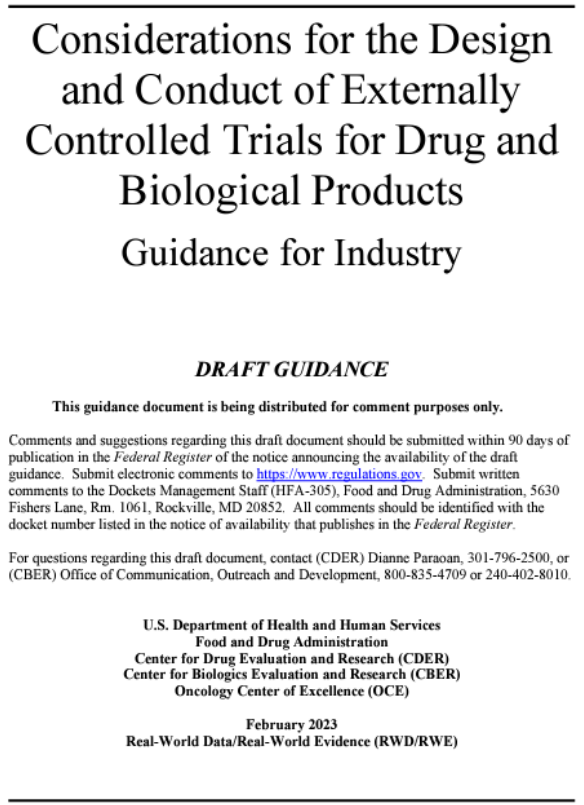
Transportability of Overall Survival Estimates From US to Canadian Patients With Advanced Non–Small Cell Lung Cancer With Implications for Regulatory and Health Technology Assessment

Sreeram V. Ramagopalan, PhD; Sanjay Popat, MD; Alind Gupta, PhD; Devon J. Boyne, PhD; Alexandre Lockhart, MSc; Grace Hsu, MSc; Dylan E. O'Sullivan, PhD; Jessica Inskip, PhD; Joshua Ray, MSc; Winson Y. Cheung, MD, MPH, FRCPC; Frank Griesinger, MD, PhD; Vivek Subbiah, MD

Real-world Evidence in HTA

Perspectives on RWD are Changing

Change	Nature of the role for RWD in drug development. Improved data access; updated guidance and advanced analytics have all opened doors.
Regulatory vs Payer	Felt the strongest in regulatory settings for targeted oncology and rare diseases. Areas of high unmet need, and challenging features for traditional trial design.
Specialization	As understanding of disease improves, many diseases become split into smaller subpopulations. Challenges for traditional trial evidence.
More evidence	Simultaneously, payers are requesting more evidence for comparative efficacy and decision making. RWD helping to bridge evidence gaps.



Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Dianne Paraoan, 301-796-2500, or (CBER) Office of Communication, Outreach and Development, 800-835-4709 or 240-402-8010.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

February 2023
Real-World Data/Real-World Evidence (RWD/RWE)

NICE real-world evidence framework

Corporate document

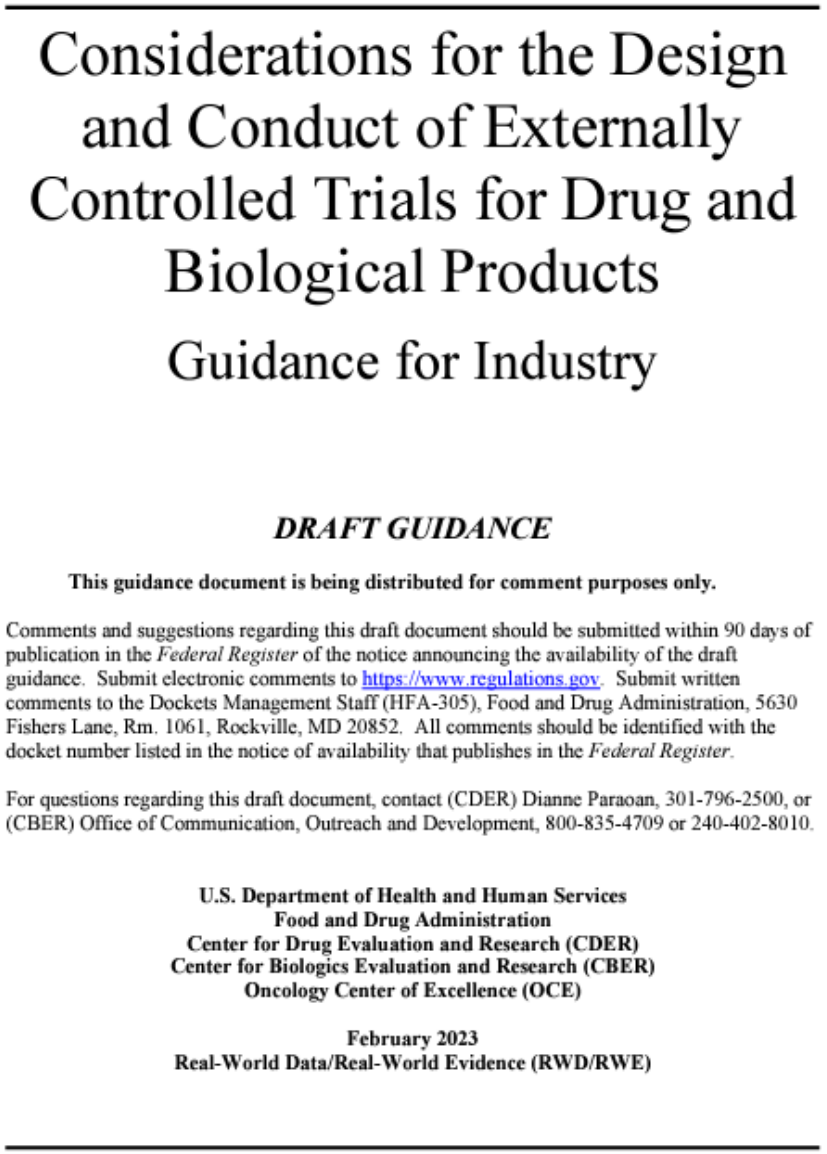
Published: 23 June 2022

www.nice.org.uk/corporate/ecd9

FDA Guidance

February 2023

- The FDA has issued *draft* guidance for sponsors to use data from registries and electronic health records in lieu of data from randomized controlled trials (RCTs).
- Sponsors should consider the *likelihood* that such a trial design would be able to distinguish the effect of a drug and meet regulatory requirements.
- **The suitability of using an externally controlled trial design warrants a case-by-case assessment.**
- Sponsors should consult with the FDA **early** on to determine whether it is reasonable to conduct an external control trial.



Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Dianne Paraoan, 301-796-2500, or (CBER) Office of Communication, Outreach and Development, 800-835-4709 or 240-402-8010.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

February 2023
Real-World Data/Real-World Evidence (RWD/RWE)

CADTH

- **RWE Guidance Working Group**
  - Reporting to CADTH's Real-World Evidence Steering Committee
  - Objective is to: *"…develop clear and comprehensive guidance on the conduct, reporting, and appraising of RWE studies concerning the safety and effectiveness of health technologies for the purpose of regulatory approval and health technology assessment (HTA) in Canada."*

Section 10 of **RWE Draft Reporting Guidance:** *Bias, Confounding, and Effect Modifiers/Subgroup Effects*

Recommendations:

- Report all procedures used to address potential sources of bias
- Specify how potential sources of bias could influence the outcomes of the analyses
- Report whether any potential confounders could not be measured and specify the anticipated impact of these confounders on study results
- Specify the methods used to conduct sensitivity analyses that test key assumptions and limitations of the data, and if no sensitivity analyses were conducted, explain why not

- *Guidance document now published*

Haute Autorité de santé (HAS)

A recently published article in BMJ Evidence-Based Medicine from the French National Authority for Health (Haute Autorité de Santé) draws attention to the importance of both target trial emulation and **quantitative bias analysis** as critical tools for including real-world evidence in submission packages.

- *"The discussion of...[residual bias]... should not be based on expert opinion only and should be documented, for example, using … **quantitative bias analysis**."*

- *"Residual confounding has been explored with analyses such as the use of …**quantitative bias analysis** and excludes a conclusion of no treatment effect."*

EBM analysis

OPEN ACCESS

Rapid access to innovative medicinal products while ensuring relevant health technology assessment. Position of the French National Authority for Health

Antoine Vanier,[1,2] Judith Fernandez ,[1] Sophie Kelley,[1] Lise Alter,[1] Patrick Semenzato,[1] Corinne Alberti,[3,4] Sylvie Chevret,[5] Dominique Costagliola,[6] Michel Cucherat,[7] Bruno Falissard,[8] François Gueyffier,[9] Jérôme Lambert,[5] Etienne Lengliné,[10] Clara Locher ,[11] Florian Naudet ,[12,13] Raphael Porcher,[14] Rodolphe Thiébaut,[15] Muriel Vray,[16] Sarah Zohar,[17,18] Pierre Cochat,[19] Dominique Le Guludec[19]

10.1136/bmjebm-2022-112091

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjebm-2022-112091).

For numbered affiliations see end of article.

Correspondence to:
Doctor Antoine Vanier, Health Technology Assessment Department, Haute Autorité de Santé, La Plaine Saint-Denis, 93210, France; a.vanier@has-sante.fr

The challenge of accelerated clinical developments

In France, decisions for reimbursement taken by the Ministry of Health are based on appraisal by an independent Health Technology Assessment body (HTAb): the 'Haute Autorité de santé' (HAS). HAS grades the clinical added value of any medicinal product for which a manufacturer seeks reimbursement. This appraisal considers different types of clinical and patient-centred outcomes, including patient-reported ones. Under certain conditions, a concomitant economic assessment which accounts for patients' preferences in the form of utility values is also performed.

As providing fast access to breakthrough therapies is a critical expectation from patients, clinicians and health policy makers, the European Medicines Agency and the Food and Drug Administration have established various acceler-

of relevant HTA in this context is highly challenging. Thus, the French Minister of Health requested HAS to provide recommendations. A consultation of patient associations, academics, manufacturers and various institutions was conducted from October 2021 to January 2022. With the support of an expert committee, a qualitative summary of the consultation has led to the prioritisation of recommendations, which are developed below (details on the consultation process are available in an online supplemental appendix 1).

Need for evidence from comparative designs allowing causal interpretation of treatment effect estimation

Performing relevant HTA requires that an unbiased estimate of the treatment effect is available. Thus, the additional effect must be disentangled beyond

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

- IQWiG has historically criticized the lack of relevant real-world comparator arms representing German routine practice [1,2] however, **recent benefit assessments in oncology showed a potential shift toward use of high quality non-German data**
    - In 2022, IQWiG suggested benefits of real-world ex-German comparator arms in assessments for lung cancer treatment[3,4]

> "…the company does not cite any reasons for […] using data only from the [German] CRISP registry for the comparison of individual arms […] despite the fact that further potentially relevant patient registries exist… For instance, the company itself mentioned the [United States] Flatiron Health database as a potential further data source…" (dossier assessment for sotorasib, 2022)

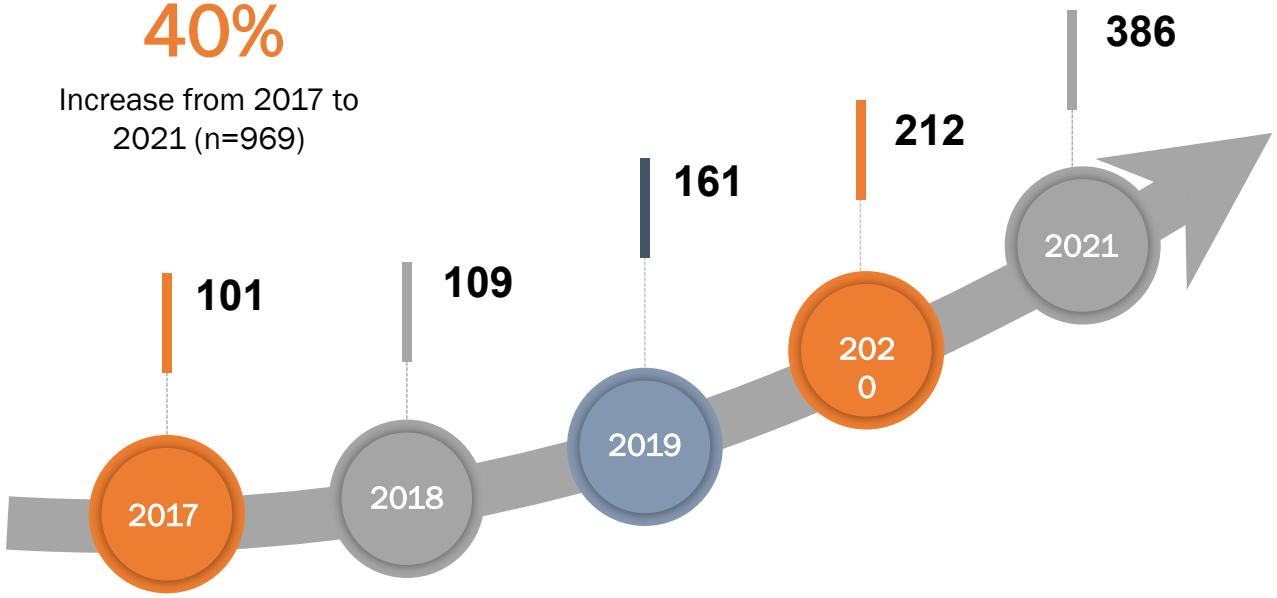
- Researchers may not be able to avoid the use of international data in German submissions if quality German RWD is not available. **Transportability methods can have a complementary role** to address common concerns about relevance of external evidence in the German setting.

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Supply-related data in manufacturer dossiers: things are not yet running smoothly. (2022). www.iqwig.de/en/presse/press-releases/press-releases-detailpage_67103.html
2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Registry-day – of sufficient quality – are suitable for the extended benefit assessment of drugs. https://www.iqwig.de/en/presse/press-releases/press-releases-detailpage_9982.html
3. Institute for Quality and Efficiency in Health Care. Sotorasib (NSCLC) - Nutzenbewertung gemäß § 35a SGB V. (2022). www.iqwig.de/download/a22-28_sotorasib_nutzenbewertung-35a-sgb-v_v1-0.pdf
4. Institute for Quality and Efficiency in Health Care. Amivantamab (NSCLC) - Nutzenbewertung gemäß § 35a SGB V. (2022). www.iqwig.de/download/a22-05_amivantamab_nutzenbewertung-35a-sgb-v_v1-0.pdf

Use of RWE in HTA and Regulation

Dynamics, drivers, and barriers to the use of RWE in HTA

HTA submissions with RWE are accelerating
(especially in UK, Germany, France, and Canada)



Submissions by country from 2017 to 2021

- Brazil = 152
- Canada = 816
- China = 4
- France = 728
- Germany = 749
- Italy = 153
- Japan = 13
- Spain = 269
- UK = 678
- US = 12

3,574
Total submissions

Source: IQVIA HTA Accelerator
Single Technology Assessment; original submissions, indication extensions and resubmissions between January 1, 2017 and December 2021 with RWE included and published by bodies
ISPOR Europe 2022 – Vienna: Use of Real-world Evidence to Support Health Technology Assessment in United States, Europe and Japan – A Brief Analysis

Use of RWE in HTA and regulation (cont.)

Dynamics, drivers, and barriers to the use of RWE in HTA

Top RWE research areas in accepted submissions	349 Safety	233 Epidemiology	149 Effectiveness
Main therapeutic areas with accepted submissions with RWE	280 Oncology	107 Endocrine and metabolic diseases	56 Infectious and parasitic diseases Central nervous system
Most frequent rationale for rejection	57 Risk of bias	25 Patient selection	17 Insufficient data Study population not well defined

Source: IQVIA HTA Accelerator
Single Technology Assessment; original submissions, indication extensions and resubmissions between January 1, 2017 and December 2021 with RWE included and published by bodies
ISPOR Europe 2022 – Vienna: Use of Real-world Evidence to Support Health Technology Assessment in United States, Europe and Japan – A Brief Analysis

Use of RWE in HTA and regulation

European payer, regulatory, and HTA perspectives on RWE

Payer perspective

- **RWE as a tool to supplement, rather than replace RCT**
- Payers with responsibility for assessment after market entry are more responsive to RWE vs payers focused entirely on initial assessment
- Advisable for manufacturers to obtain payer insights throughout the clinical development program

Suggested use cases

- (Ultra)orphan drugs and advanced medicinal products
- International approach to treatments for small patient populations
- Limited information on these products after market authorization
- Information on historical controls
- Direct collaboration between countries

HTA perspective

- Many widely accepted use cases for RWE in HTA, but comparative effectiveness remains contentious
- NICE 2021 to 2026 strategy underpinned by a broader use of data
- To build trust in RWE:
  - Build competence
  - Proactively address data gaps
  - Use NICE's RWE framework

Suggested use cases

- Populating and validating economic models
- Patient or user experience
- Impact of tests on decisions about care
- Impact of technologies on care delivery
- Understanding unmet
- Epidemiology of disease

Regulatory perspective

- RWE is a key aspect of the EMA clinical evidence vision for 2030
- Clinical evidence 2030 views on RWE:
  - Establish value across use cases
  - Build business processes
  - Set standards
  - Enable access
  - Validate methods
  - Train staff and stakeholders
  - Internationalize

Suggested use cases

- Support planning and validity of applicant studies: design and feasibility of planned studies, assess representative and validity of completed studies
- Understand clinical context: epidemiology, clinical management
- Investigate associations and impact: Effectiveness, safety, label expansion

Source: IQVIA HTA Accelerator
Single Technology Assessment; original submissions, indication extensions and resubmissions between January 1, 2017 and December 2021 with RWE included and published by bodies
ISPOR Europe 2022 – Vienna: Use of Real-world Evidence to Support Health Technology Assessment in United States, Europe and Japan – A Brief Analysis

Scrutiny of RWE

Common sources of bias in RWE

- Unmeasured variables of interest
- Lack of randomization
- Missing values in measured RWD
- Lack of harmonization
- Selection of eligible participants

Challenges for decision-makers

- Lack of trust in RWD (validity, geographical relevance)
- Concerns about residual bias
- Risk of bias assessments are often qualitative and based on heuristics
- Lack of pre-specified analysis

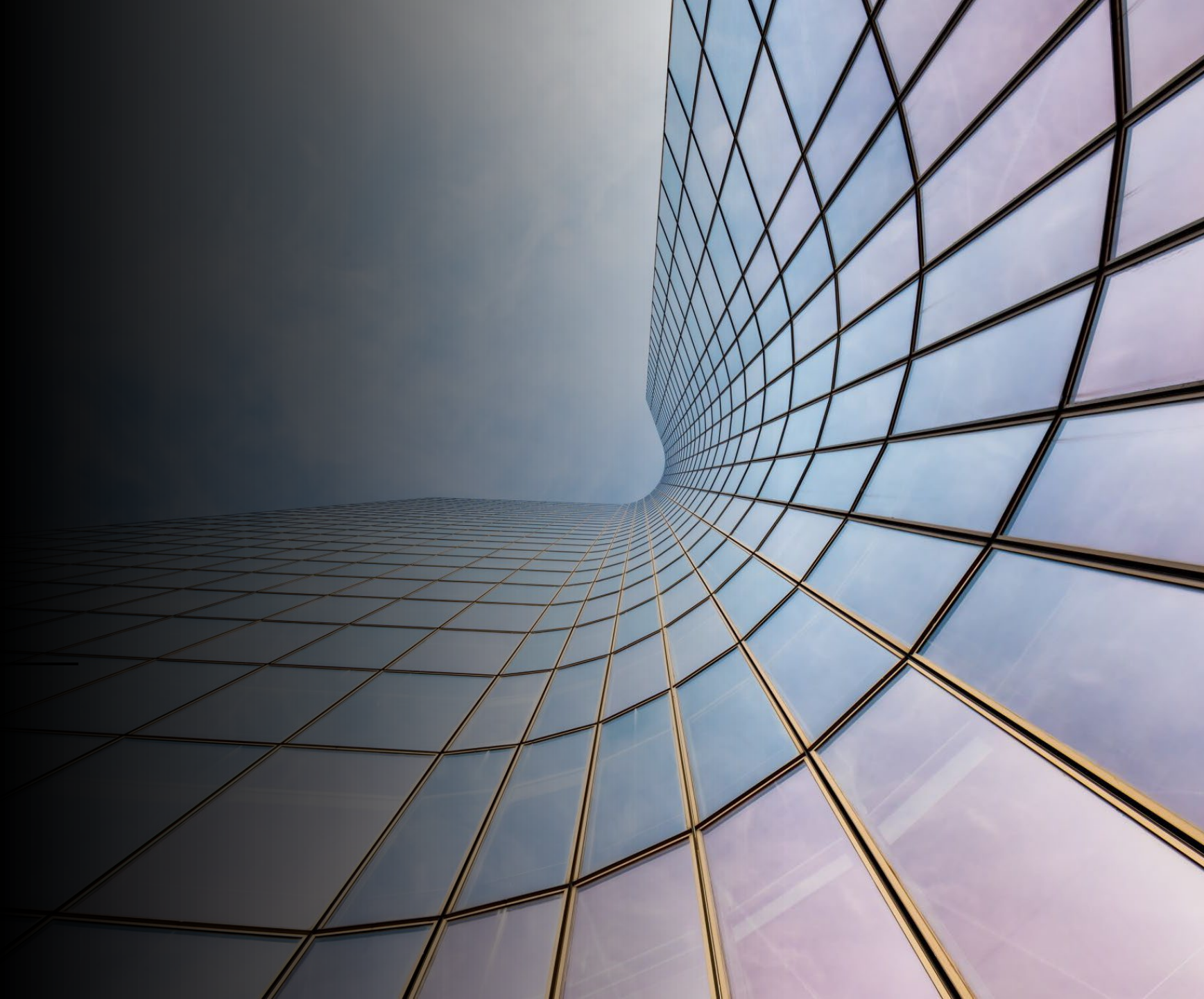
Confounding and missing data with external evidence

- Determining how “bad” these evidence gaps are is challenging, and context specific.
- Recently, efforts have been put towards attempting to quantify these limitations, acknowledging that no data source is perfect.
- For example, UK NICE specifically endorses **quantitative bias analysis** (QBA).

NICE Real-World Evidence Framework

Corporate document published 23 June 2022

External Validity



External Validity

- Correctly estimating therapeutic intervention effectiveness is critical
  - Internal validity
- Generalizability: From study sample to broader populations.
  - Study may show positive outcome but…
- Challenges: Differences across geographies and patient groups.
- Gap in practice: External validity often overlooked.
  - Inappropriate generalizations can potentially leading to suboptimal clinical decisions.

Potential Limitations of RCTs (6SC)

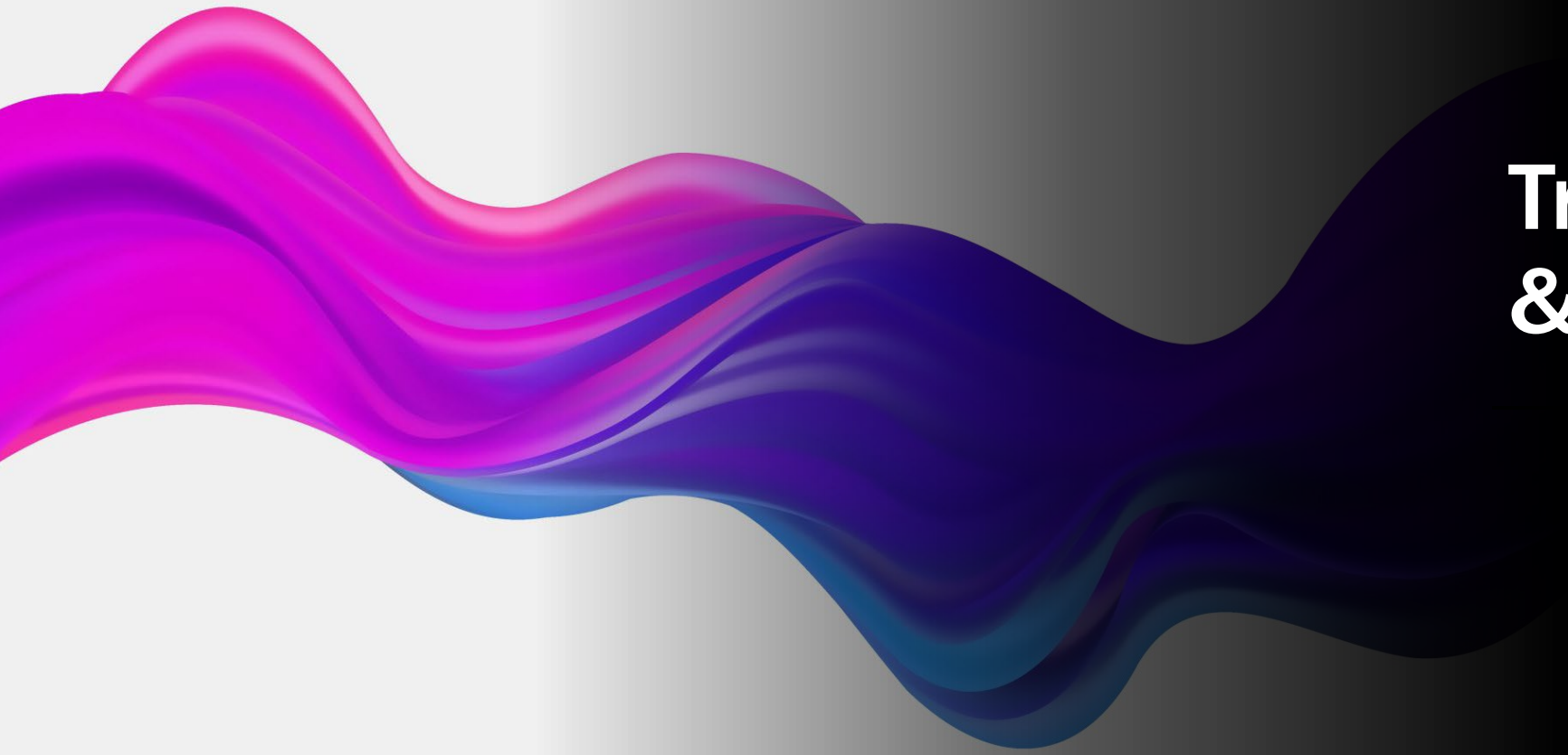
- Too Small: Limited size impacts rare outcome detection.
- Too Simple: Challenges in detecting interactions.
- Too Selected: Underrepresentation of key populations.
- Too Specific: Overly specific inclusion/exclusion criteria.
- Too Short: Short duration affects long-term outcome detection.
- Surrogate Measures: Efficacy may be based on indirect measures.
- Comparator Issues: Poor choice of comparator.

RCTs and Potential Threats to External Validity

- RCTs often lack representativeness of broader patient populations.
- Underrepresentation of certain at-risk demographic groups in RCTs.
- Trial participants often differ in age, health, and diversity from real-world clinics.
- Reduced applicability of RCTs' findings in real-world clinical decision-making.
- External factors, such as geography or changing medical standards, limit RCTs' usability.
- Greater adherence to medications observed in RCTs versus routine care.
- Meta-analyses of RCTs display heterogeneity in patient populations and treatment effects.

RWE to the rescue?

- Real-world evidence (RWE) can help address the generalizability limits of RCTs.
- However …
  - RWE sources: electronic health records, claims, registries.
    - with diverse ecosystem challenges.
  - Transferring RWD across countries needs local context understanding.
  - Need for definitive "decision-grade" RWD criteria.



Transportability & Generalizability

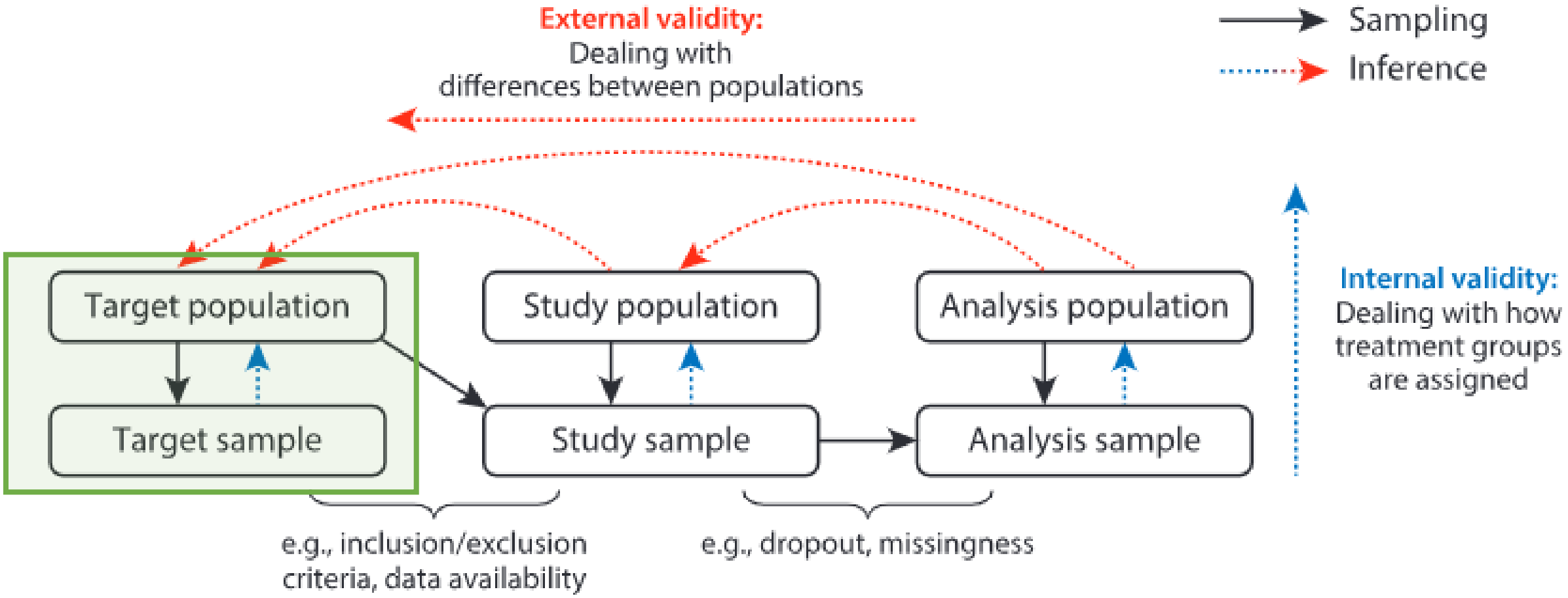
Internal Validity & External Validity

- Internal Validity
  - Ensures unbiased effect estimate within the study sample.
  - Main focus of most epidemiologic studies.

- External Validity (Generalizability & Transportability)
  - Ensures unbiased effect when applied to different settings.
  - Risks: Differences in subject characteristics, settings, treatments, outcome measures.

- Addressing Validity Concerns
  - Align study and real-world contexts (e.g., care standards, outcome measures).
  - Focus on enrollment variations, treatment effect differences, and correlations between them.
    - Study vs Target

Internal Validity & External Validity

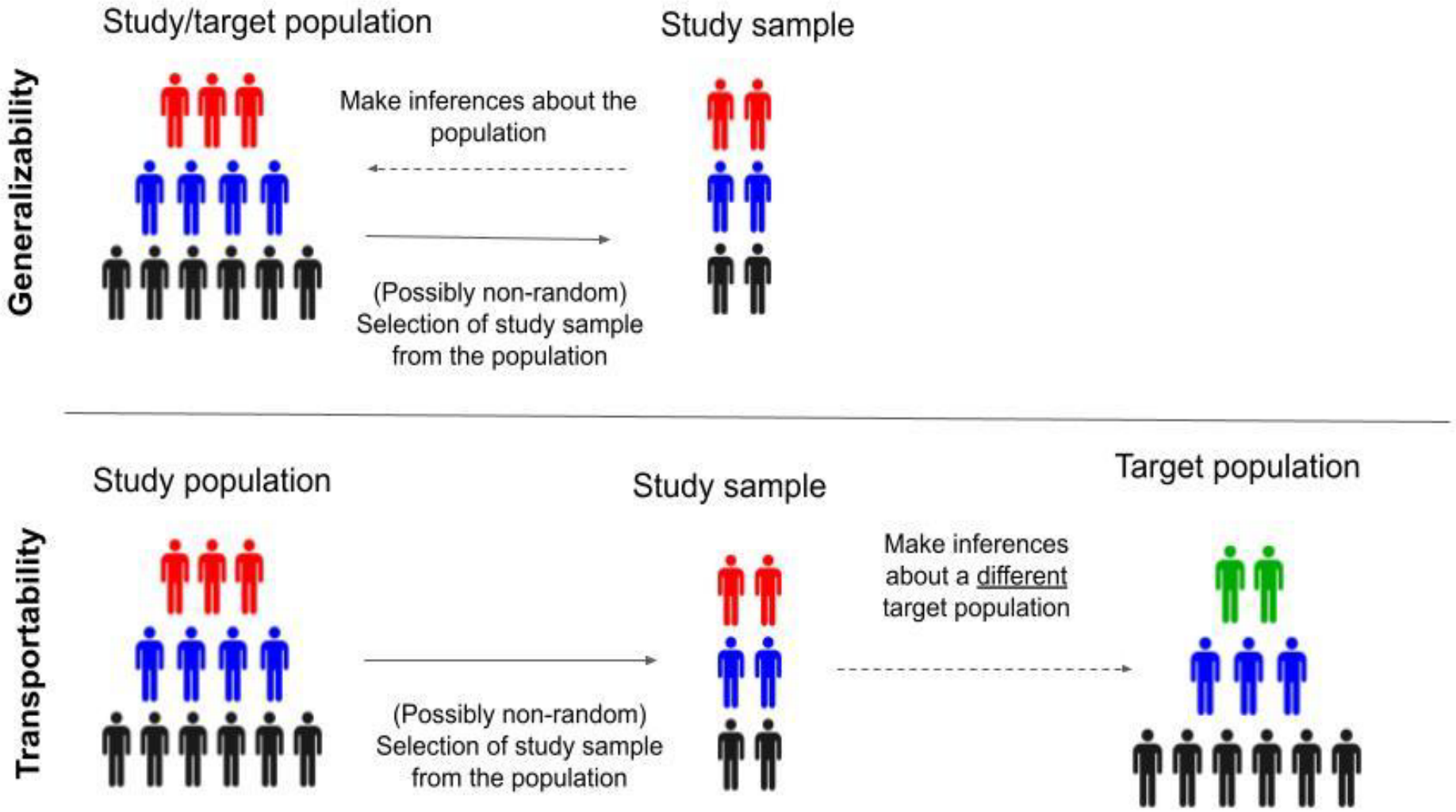


Degtiar, Irina, and Sherri Rose. "A review of generalizability and transportability." *Annual Review of Statistics and Its Application* 10 (2023): 501-524.

Generalizability and Transportability

- **Generalizability**
  - Extension of causal knowledge from study to target population.
  - Study population is a subset of the target.
  - E.g., From a specific region to a country.
- **Transportability**
  - Extends causal knowledge to a **distinct** target population.
  - Study population is **external** or distinct from the target.
  - E.g., Applying findings from one country to another.

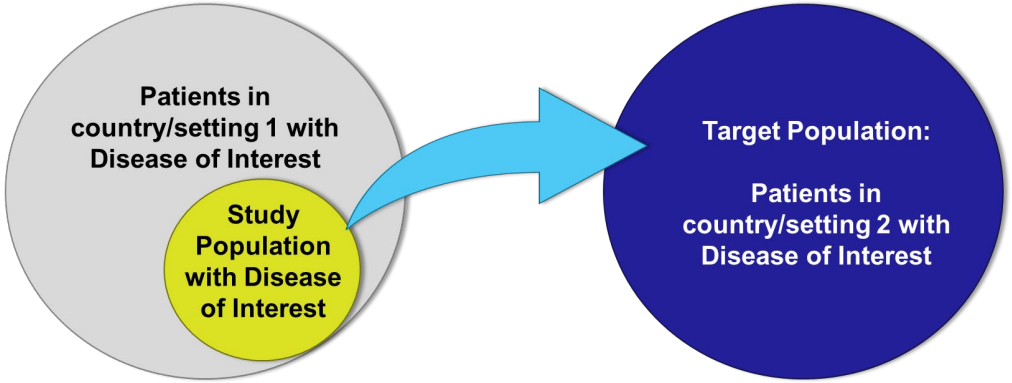
Generalizability and Transportability



Transportability Analysis

- Can a treatment effect estimated using data from population A be used to estimate the treatment effect in another population B for which we don't have treatment/outcome data?
- Transportability Analysis is a set of quantitative methods by which the extension of such effects can be estimated.
- Those working with regulators and payers can use transportability analysis to optimize data collected from trial participants, when interpreting real world evidence.
- Transportability analyses can provide evidence of **external validity**

Transportability analysis: what, where and how to use



- What is transportability analysis?
  - Quantitative methods to reliably extend conclusions made from one study population to an **external** *target population* (see Figure).
- Is transportability analysis restricted to real-world data (RWD)?
  - No. Transportability analysis can be conducted between any study population to any target population. It can involve extending conclusions from trial populations to real world target populations, or between non-overlapping real-world populations.
- How can you apply transportability analysis?
  - The use of external control arms (ECAs) from RWD to address comparative evidence gaps in single-arm submissions to reimbursement and regulatory bodies has been increasing rapidly. Variation in quality and availability of local RWD has led to use of RWD from outside of the country or setting of interest (the "target" population). **Key question for decision makers becomes how relevant the submitted evidence is for the target population and whether ECA conclusions can be reliably "transported" across countries or settings.**

Researchers can use transportability analysis methods (similar to those adjusting for confounding) to assess how ECA results from a specific population apply in the target population. Findings from transportability analyses can **enhance submissions** to regulatory and reimbursement stakeholders to address concerns over lack of transportability when using international RWD.

Technical Approaches

General

Matching
Pair individuals to achieve covariate balance between study and target population through distance metrics.

Weighting
Create a balanced pseudo-population using inverse probability of sampling.

Outcome Modeling (g-formula)
Model outcome conditional on covariates, then marginalize over (standardize to) target covariate distribution.

Doubly Robust Approaches
Combine models for sampling and outcome to provide robustness against potential misspecifications.

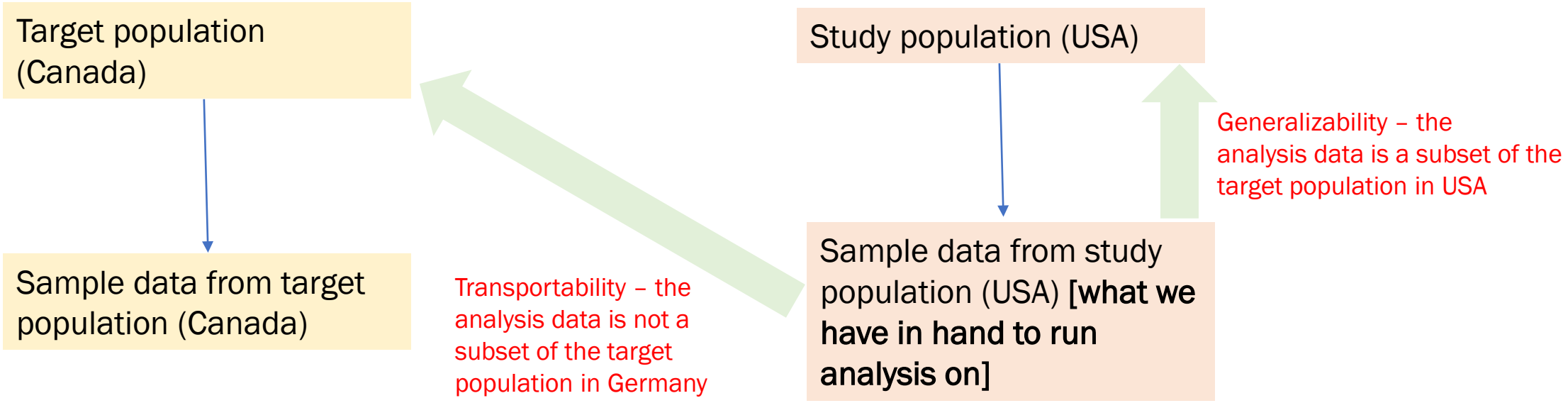
A Review of Generalizability and Transportability. Irina Degtiar and Sherri Rose. Annual Review of Statistics and Its Application 2023 10:1, 501-524

Study Motivation & Background

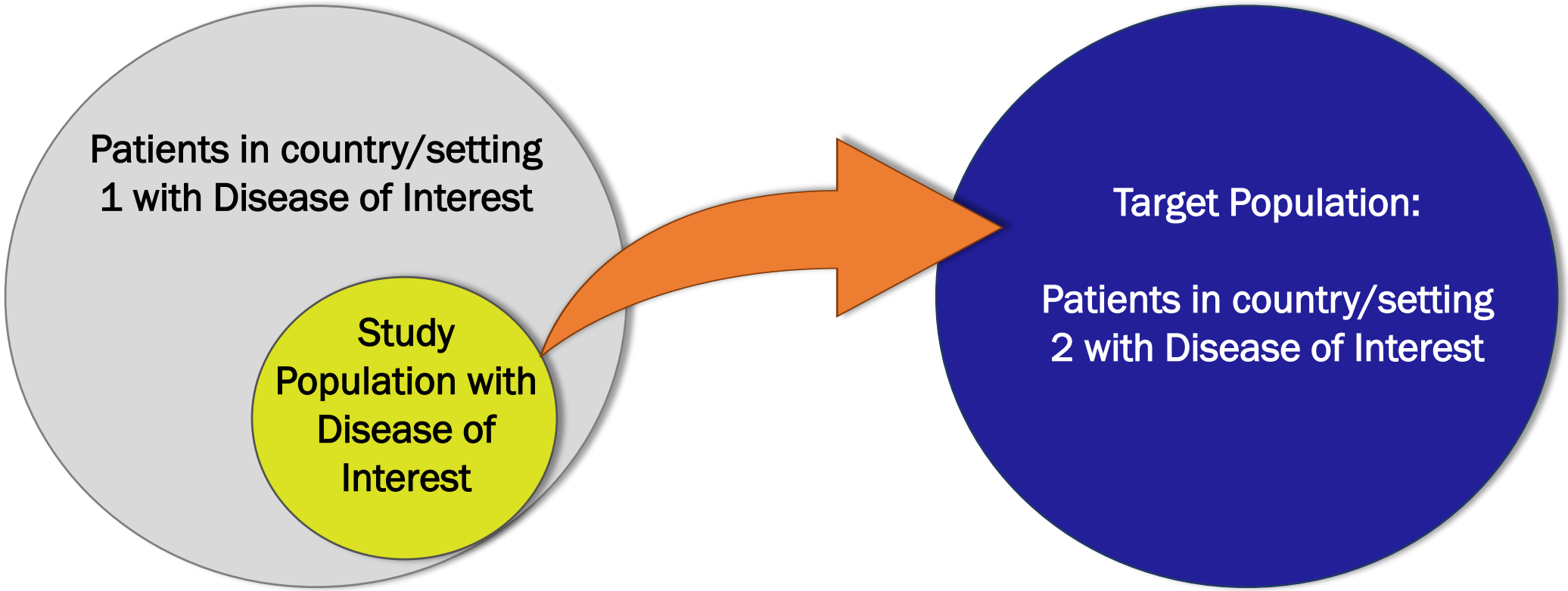
Background

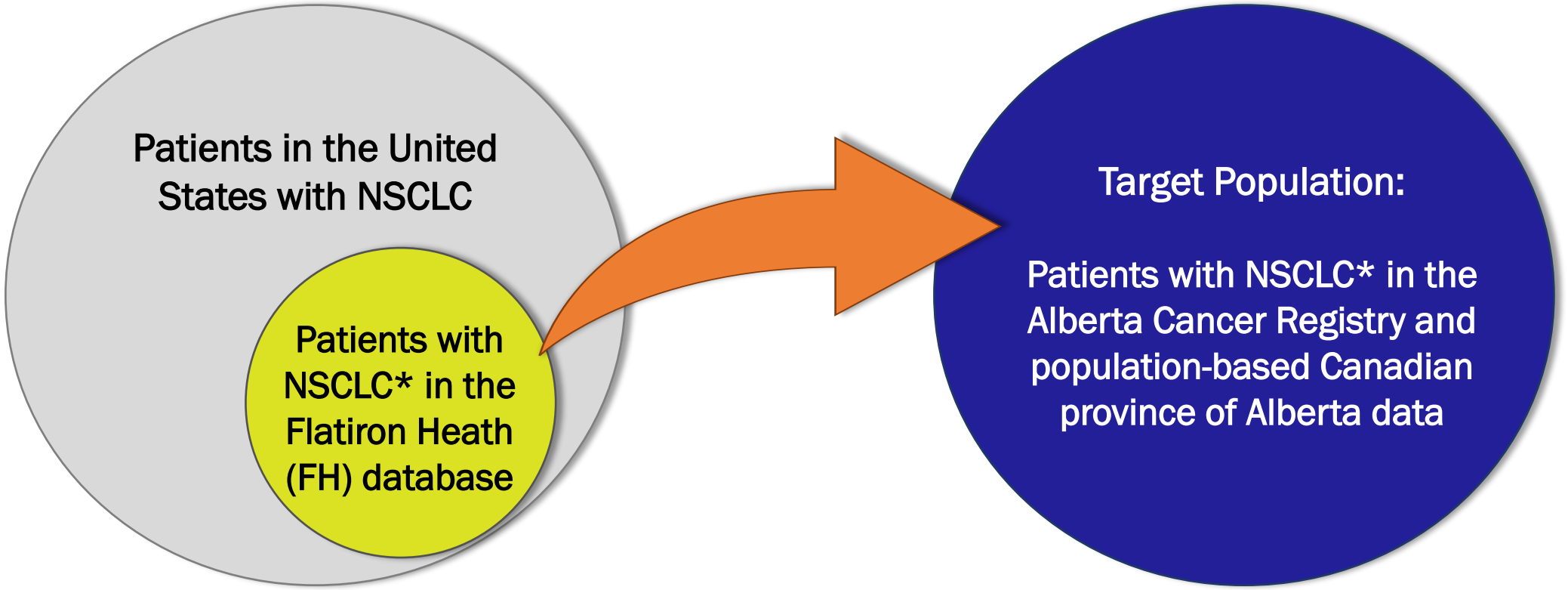
- Usage: Real-world data (RWD) increasingly supports regulatory submissions, especially for rare genetic cancers.
- Problem: Ensuring treatment effects from RWD are valid both for:
  - Original sample (generalizability)
  - Different populations (transportability)
- Study Focus: Assess RWD transportability of **survival estimates** for advanced non-small cell lung cancer (aNSCLC) between the US and Canada.
- Method: Use **transportability** analysis to evaluate if overall survival and treatment effect estimates from US RWD can be applied to Canada.

Visual representation of the study problem



- In both cases, we need to adjust the treatment outcomes estimated from sample data for effect modifiers (any variables that are imbalanced between sample data and target population that affect the treatment effect, e.g., if older patients do not respond to treatment as well as younger patients, and people in USA are younger than those in Canada)
- If we can adjust for all effect modifiers (unverifiable assumption*), then our transported effect estimate will be the same as the one that we would have estimated had we run the analysis on data from the target population


Patients in country/setting 1 with Disease of Interest
Study Population with Disease of Interest
Target Population:
Patients in country/setting 2 with Disease of Interest



* Patients on either first-line platinum-doublet chemotherapy or first-line pembrolizumab monotherapy

Project goals

Primary objective

Provide a demonstration of the application of **transportability methods** to transport overall survival estimates for aNSCLC patients who initiated 2L docetaxel and 1L platinum chemotherapy from the FHAD to O2

Secondary objective

Conduct a **quantitative bias analysis** to quantify the impact of unmeasured prognostic factors on any discrepancies between survival curves

We worked under the assumption that relative risks (e.g. hazard ratios) are transportable if absolute risks transport in the overall population.

Data Sources

RWE database analyses

Flatiron Health - USA

Derived from EHR (Electronic Health Records) data

- Longitudinal, demographically and geographically diverse
- Cutoff Date: September 30, 2020

Database Composition:

- Over 280 cancer clinics (approx. 800 sites of care) in the U.S.
- Represents >2 million active patients
- Majority are from community oncology settings

Study Approvals:

- Institutional Review Board approval obtained
- Informed patient consent waived (deidentified data)

Data Extraction:

- Includes patient-level demographic, clinical, and outcomes data
- Combination of structured data and elements from unstructured clinical documents

Data Processing:

- Structured data: aggregated, normalized, and harmonized across clinics
- Documents: classified into 24 standard categories
- Unstructured data: extracted via technology-enabled abstraction

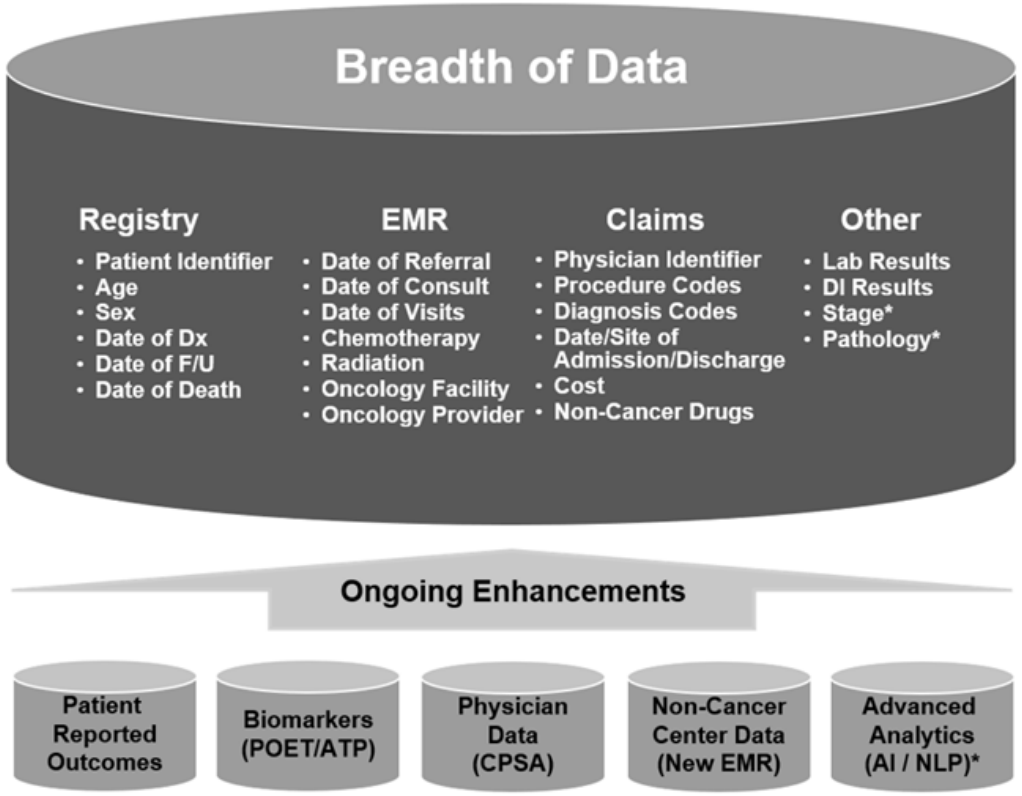
Specific Data Details:

- Dates of death: sourced from a composite mortality variable
- Lines of therapy: determined from drug order and administration, based on oncologist-defined rules

flatiron

RWE database analyses

Oncology outcomes - Canada



Population-based data
- Two tertiary centers
- Four regional centers
- 11 community centers

- Complete provincial population - Cancer Treatment and Outcomes Data for Province of Alberta (**4.5 million residents**)
- 2005 to 2022 (ongoing) (approximate six-month lag for a few of the data components)
- **>200,000 cases to date**
- **100% coverage of cancer cases** (through mandatory reporting in cancer registry)
- Lower numbers for less common cancers and by histology

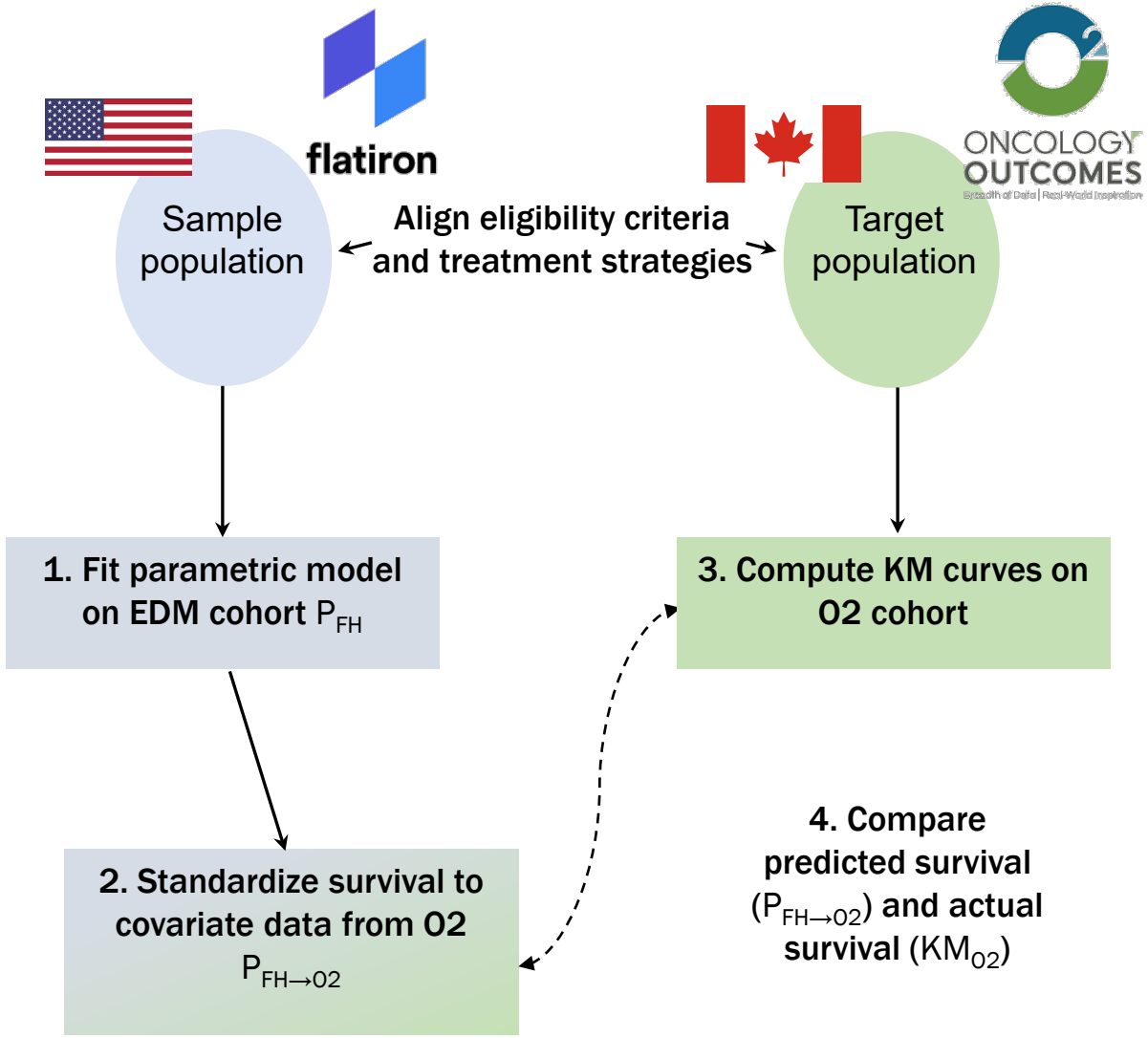
Possibility to extract additional clinical data from medical charts

Abbreviations: AI, artificial intelligence; ATP, Alberta's Tomorrow Project; CPSA, College of Physicians & Surgeons of Alberta; Dx, diagnosis; EMR, electronic medical record; F/U, follow-up; NLP, natural language processing; POET, Precision Oncology Experimental Therapeutics

Methods

Overview of transportability analysis

- Identical eligibility criteria are applied to select patient groups in sample and target populations
    - Kaplan-Meier estimates KM_{FH} and KM_{02} are unadjusted estimates of overall survival in the sample and target populations
- We first fit a parametric model P_{FH} of survival as a function of patient covariates on the sample patient group data
    - $P_{FH} \approx KM_{FH}$ if the parametric model fits well
- The transported curve $P_{FH \rightarrow 02}$ represents the model adjusted for individual-level baseline covariates from the target group 02
    - $P_{FH \rightarrow 02} \approx KM_{02}$ if transportability "holds". A threshold of <5% mean absolute difference between $P_{FH \rightarrow 02}$ and KM_{02} implied sufficient similarity for this study.



Eligibility Criteria

Harmonized between US and Canadian data sets.

Patients 18 years or older.

Diagnosed with advanced* NSCLC (stage IIIb, IIIc, or IV) on/after January 1, 2011.

Followed up until September 30, 2020.

Exclusions:

- US: >90-day gap between advanced NSCLC diagnosis and first recorded visit or medication.
- Canadian: No therapy initiation within 180 days of diagnosis.
- Tumor characteristics as "not otherwise specified".
- Missing data for baseline covariates (US)

* Data for patients with early-stage cancer progressing to advanced disease was unreliable in Alberta data set.

Treatment Regimens

- Two primary groups
  - First-line platinum-doublet chemotherapy after diagnosis (e.g., cisplatin + paclitaxel).
  - First-line pembrolizumab monotherapy.
- Exploratory analysis on third group
  - Second-line docetaxel after previous chemotherapy but no exposure to certain immunotherapies.
- Outcomes expected to be homogenous within each treatment group.
  - Any dose permitted.
- Limitation
  - Information on ECOG performance status and post-diagnosis metastases not available for Canadian data set.

Baseline Covariates

Age, sex, cancer stage at diagnosis, ECOG performance status.

Tumor histological characteristics, smoking history.

Time since diagnosis, time since January 1, 2011.

Comorbidities and metastases

- potentially recorded differently between US and Canadian samples.

*Race and ethnicity not analyzed

*Not available in Canadian Data; assumed expected similarity between US and Canadian patients.

Outcome

- **Overall survival**
  - Measured from index date to all-cause death.
  - For the FH data set, the 15th of each month was imputed as the date of death.
  - Patients with missing information were censored at last recorded activity or September 30, 2020.

Outcome Model & Approach

- Data pooling limitations between US and Canadian datasets.
- Prespecified outcome regression model used for survival as a function of patient-level covariates.
- Standardized using target population covariate distributions to obtain **marginal survival probabilities**.
- **Pooled logistic regression** model for transportability analysis:
  - Fitted on up to 60 months of US cohort follow-up data.
  - Modeled probability of survival based on baseline covariates.
  - Q model specification: no interaction terms; quadratic terms for continuous variables.
  - Time (in months) as a cubic spline with manually specified knot locations.
  - Coefficients equivalence checked against **Cox regression** for time-to-event data.

Estimation and Assessment of Transportability

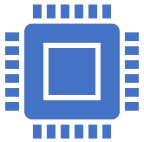
Individual-level survival probabilities estimated using fitted models for up to 60 months.

Used baseline covariates for either US or Canadian cohorts for analysis.

Cumulative mean survival probability by month derived.

Standardized parametric estimates of OS in Canadian cohorts compared with Kaplan-Meier estimates:

For sufficient similarity, ≤5% mean absolute difference chosen between model estimated & observed OS.



Percentile-based 95%CIs with 1000 iterations of nonparametric bootstrapping.

Resampling by patient, not observation (using patient-month unit).

Monthly survival probabilities plotted as a function of time.

Statistical Model

Pooled Logistic Regression

$$Ln\left(\frac{P(t_k, Y_i,\ X_i)}{1-P(t_k, Y_i,\ X_i)}\right) = \beta_o + Y_i(t_k)^T\boldsymbol{\gamma} + X_i{}^T\alpha + \theta_k$$

- **Definition:** PLR uses logistic regression to relate predictors to event outcomes within specific intervals.
- **Event Outcome:**
  - Indicates whether an event occurs in an interval.
  - Does not specify when the event occurs within that interval.
  - Events at start and end of the interval are treated equally.
- **Key Properties:**
  - No inflation of test statistics due to multiple interval records per individual.
  - Likelihood factors into a distinct term for each interval.
  - Treats all records within the person-period dataset as conditionally independent.
- **Estimations:**
  - Provides conditional odds ratios for event in an interval.
  - Direct estimates of the hazard rate with approximate standard errors.
- **Connections to Other Models:**
  - When follow-up is short and event is rare, approximates estimates from the Cox proportional hazards model.

β_o is the intercept for the logistic model.

$Y_i(t_k)$ represents the observed longitudinal measures for the interval;

θ_k denotes the effect of time t_k.

The time point t_k is an element of the vector representing when the longitudinal measures were recorded.

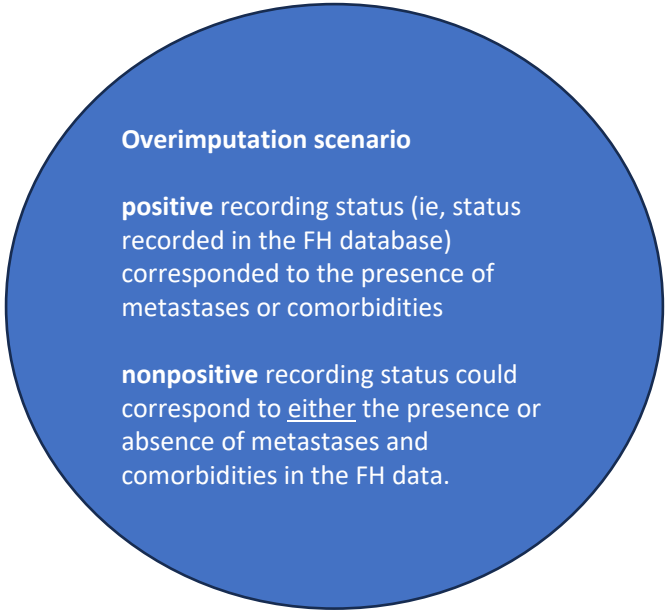
Ngwa, J.S., Cabral, H.J., Cheng, D.M. *et al.* A comparison of time dependent Cox regression, pooled logistic regression and cross sectional pooling with simulations and an application to the Framingham Heart Study. *BMC Med Res Methodol* **16**, 148 (2016). https://doi.org/10.1186/s12874-016-0248-6

Quantitative Bias Analysis

Tipping Point and Sensitivity Analyses

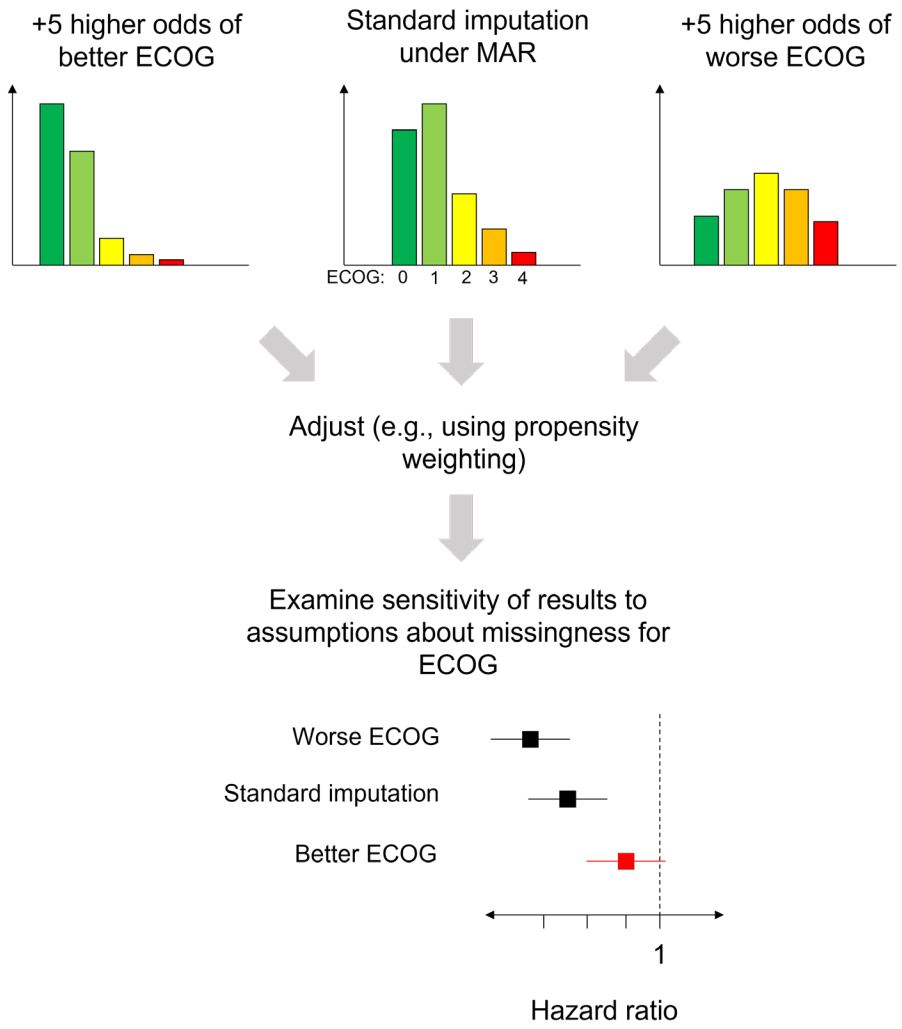
- **Objective**: Evaluate potential consequences of underrecorded **metastases** and **comorbidities** in the FH database for transportability results.

- **Methodology**:
  - Employed a **tipping point analysis** by imputing values for inaccurately measured metastases and comorbidities.
  - Used **logistic regression models** to determine metastases and comorbidities based on:
    - Survival time (months)
    - Event indicator at follow-up end
    - Baseline covariates.
  - Models helped in imputation for patients missing recorded data on conditions.
  - Introduced overimputation for bias analysis
  - Used δ adjustment to simulate prevalence increase until mean absolute difference was ≥ 5% (*“tipping point”*).



Handling missing values under different assumptions

δ-adjustment for MNAR



1. Imputation of missing values using different settings

2. Adjustment for each setting

3. Compare conclusions

Apply a shift value to predictions to simulate better- or worse-than-expected (given observed data) imputations **in one treatment group**

δ-adjustment for MNAR

- Apply a shift value δ to the imputation model (the interpretation of δ depends on the imputation model)

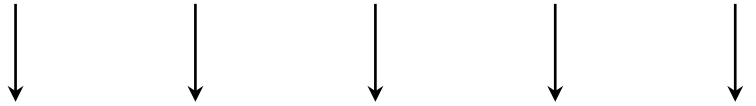
Impute Z with multiple imputation under MAR

$$\text{logit}\{P[Z|\boldsymbol{X}]\} = \gamma_0 + \boldsymbol{\gamma}^T\boldsymbol{X}$$

Run analysis (e.g. IPTW Cox) on each imputed dataset and get pooled effect estimate

Impute Z with multiple imputation under MNAR over a range of δ values

$$\text{logit}\{P[Z|\boldsymbol{X}]\} = \gamma_0 + \boldsymbol{\gamma}^T\boldsymbol{X} + (1-R)\delta$$



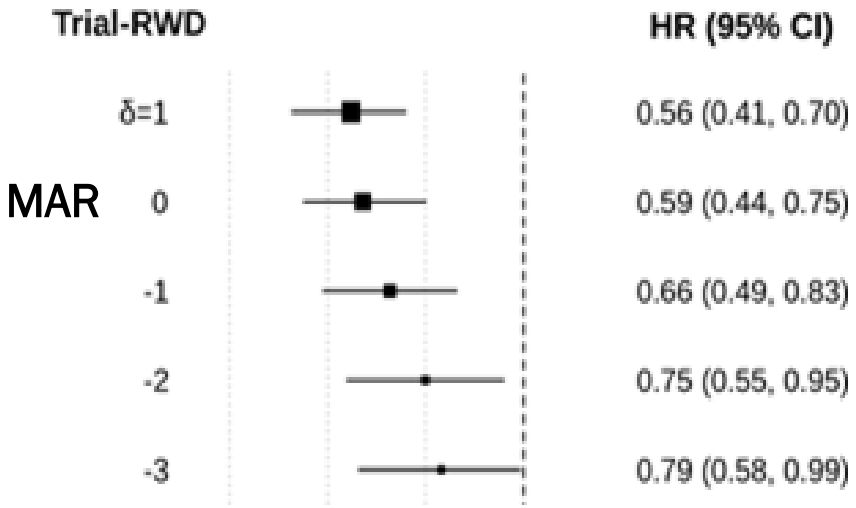
For each δ, run analysis model on each imputed dataset and get pooled effect estimates

Identify δ where conclusions change

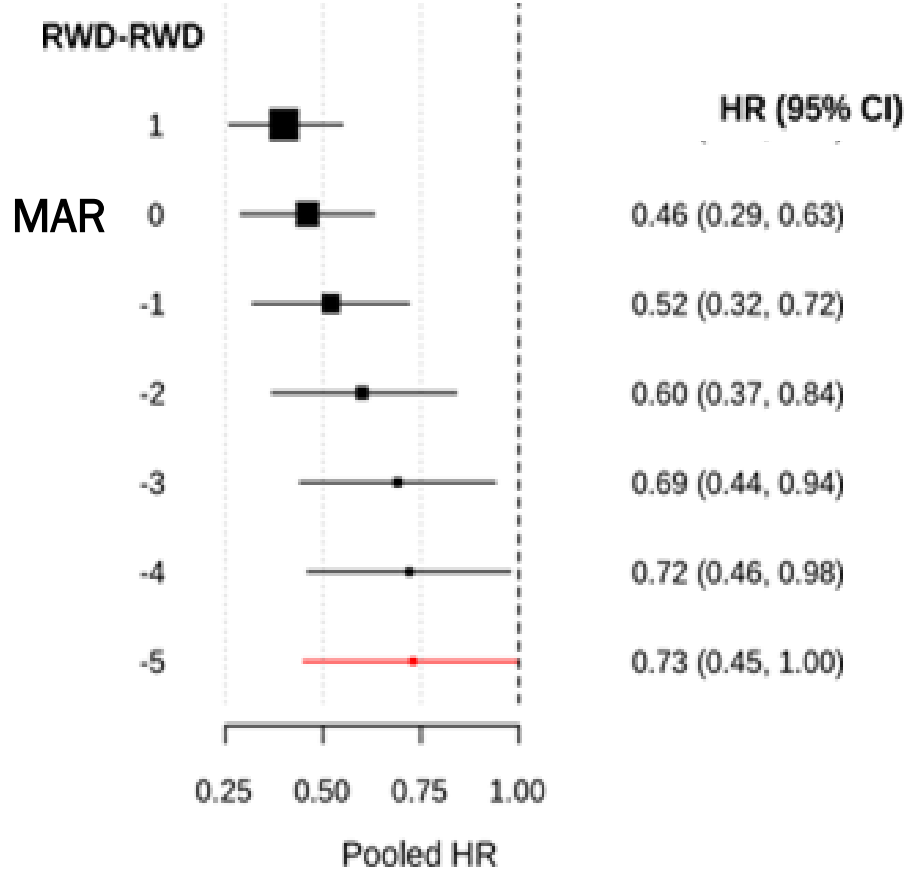
Modelling deviations from MAR: missing data

Results from Wilkinson et al

RCT vs. RWD
alectinib vs. ceritinib



RWD vs. RWD
alectinib vs. ceritinib



Wilkinson S, Gupta A, Scheuer N, Mackay E, Arora P, Thorlund K, et al. Assessment of Alectinib vs Ceritinib in ALK-Positive Non-Small Cell Lung Cancer in Phase 2 Trials and in Real-world Data. JAMA Netw Open. 2021 Oct 1;4(10):e2126306.

Sensitivity to Unadjusted Differences in 2L

US vs Canada

- Evaluated results' sensitivity concerning the prevalence of PD-L1 immunotherapy after disease progression from first-line platinum-doublet chemotherapy.
- Used **G computation** to estimate marginal risks under two hypothetical dynamic treatment regimens.
- Included time-varying cancer progression indicator and a 3-way interaction for time-varying hazards.
- Modeled two interventions based on transition between chemotherapy and immunotherapy:
  - Chemotherapy → Immunotherapy
  - Chemotherapy → Chemotherapy
- Represented hypothetical scenarios where patients discontinuing first-line treatment could either receive only immunotherapy or only chemotherapy.
- Calculated maximum risk differences using **nonparametric bootstrapping** with gfoRmula package for R, version 0.3.2.

Results

Baseline Characteristics

US and Canadian Patients With Complete Data for Covariates

Table. Baseline Characteristics of US and Canadian Patients With Complete Data for Covariates

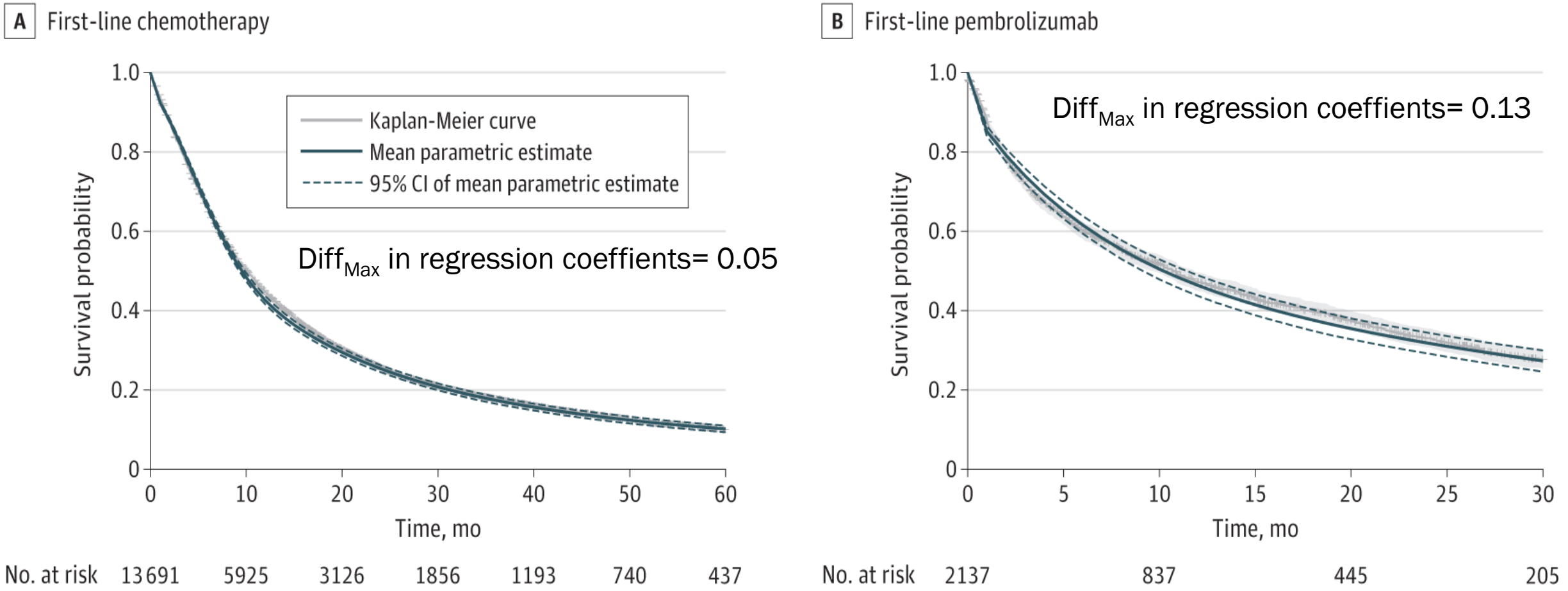
Characteristic	Patients, No./total No. (%)					
	First-line chemotherapy			First-line pembrolizumab		
	US (n = 8447)	Canada (n = 1476)	SMD	US (n = 1653)	Canada (n = 287)	SMD
Age at index date, mean (SD), y	67.34 (9.25)	65.07 (9.53)	0.242	71.64 (9.81)	69.01 (8.95)	0.280
Sex						
Female	3602/8447 (42.6)	703/1476 (47.6)	0.111	803/1653 (48.6)	149/287 (51.9)	0.066
Male	4845/8447 (57.4)	773/1476 (52.4)		850/1653 (51.4)	138/287 (48.1)	
Cancer stage at diagnosis						
IIIb or IIIc	2679/8447 (31.7)	264/1476 (17.9)	0.324	94/1653 (5.7)	27/287 (9.4)	0.140
IV	5768/8447 (68.3)	1212/1476 (82.1)		1559/1653 (94.3)	260/287 (90.6)	
ECOG performance status						
0-1	6625/8447 (78.4)	1091/1476 (73.9)	0.106	1107/1653 (67.0)	209/287 (72.8)	0.127
≥2	1822/8447 (21.6)	385/1476 (26.1)		546/1653 (33.0)	78/287 (27.2)	
Tumor histological characteristics						
Nonsquamous	5168/8447 (61.2)	1228/1476 (83.2)	0.507	1256/1653 (76.0)	244/287 (85.0)	0.229
Squamous	3279/8447 (38.8)	248/1476 (16.8)		397/1653 (24.0)	43/287 (15.0)	
Smoking history						
Ever	7808/8447 (92.4)	1343/1476 (91.0)	0.051	1521/1653 (92.0)	255/287 (88.9)	0.106
Never	639/8447 (7.6)	133/1476 (9.0)		132/1653 (8.0)	32/287 (11.1)	
Time from diagnosis to index date, median (IQR), mo	1.12 (0.72-1.63)	1.84 (1.25-2.76)	0.330	1.25 (0.89-1.81)	1.81 (1.30-2.52)	0.148
Time since January 1, 2011, median (IQR), y[a]	5.28 (3.53-7.02)	4.58 (2.50-6.44)	0.297	7.68 (6.78-8.68)	7.89 (7.28-8.52)	0.192
No. of comorbidities						
0	6188/8447 (73.3)	837/1476 (56.7)	0.362	1062/1653 (64.2)	169/287 (58.9)	0.109
≥1	2259/8447 (26.7)	639/1476 (43.3)		591/1653 (35.8)	118/287 (41.1)	
No. of sites of metastases						
0-1	7304/8447 (86.5)	877/1473 (59.5)	0.638	1367/1653 (82.7)	170/285 (59.6)	0.527
≥2	1143/8447 (13.5)	596/1473 (40.5)		286/1653 (17.3)	115/285 (40.4)	

Abbreviations: ECOG, Eastern Cooperative Oncology Group; SMD, standardized mean difference.

[a] All eligible patients were previously diagnosed de novo with advanced NSCLC on or after January 1, 2011.

Survival Curves for US Patients

Goodness of fit on total US data as a **positive control**

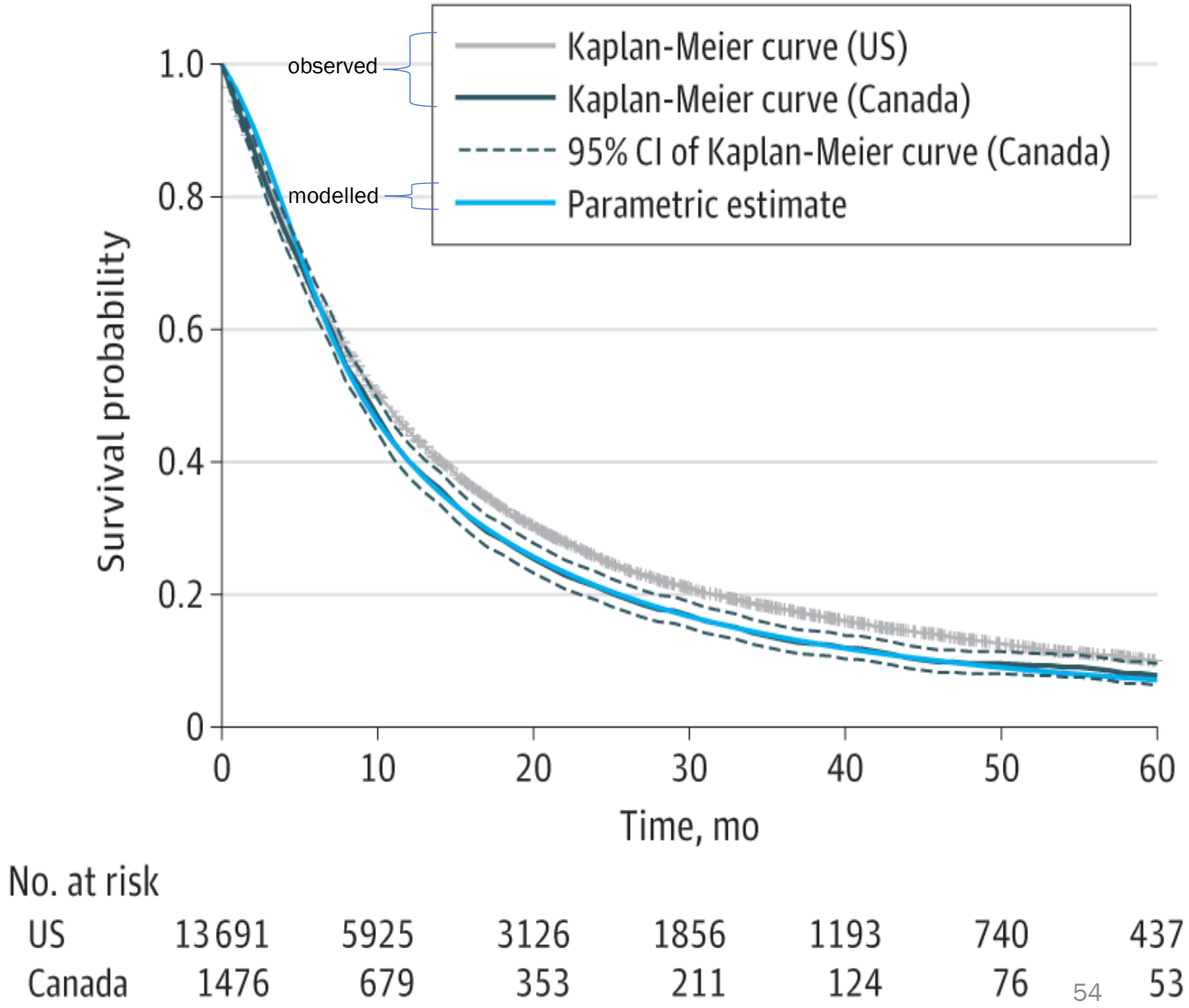


- Comparison of unadjusted Kaplan-Meier curves vs standardized parametric estimates (outcome model).
- Kaplan-Meier curves and parametric estimates for the sample population were expected to overlap by design (positive control).

Transportability results

First-line chemotherapy

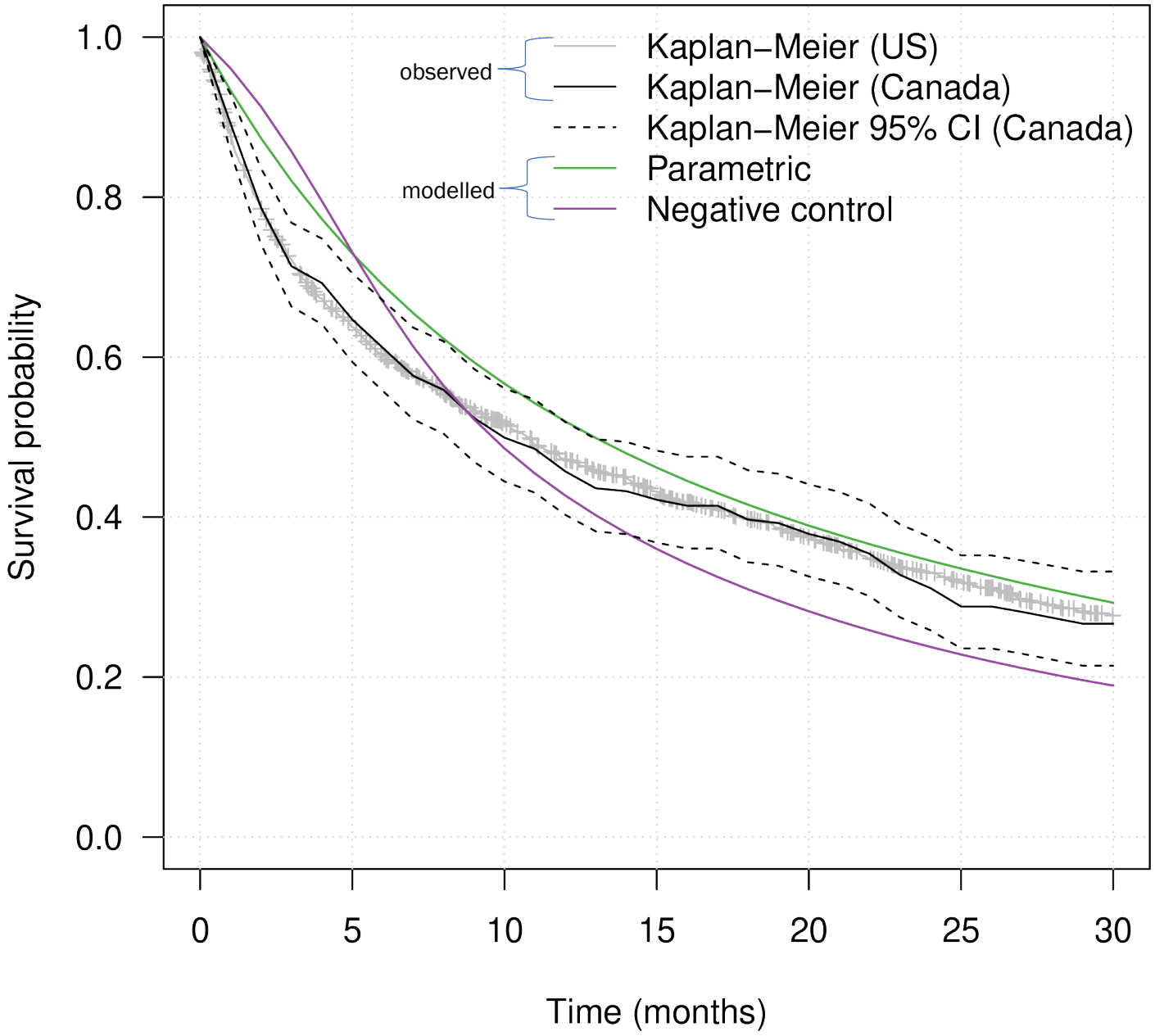
- After adjustment for baseline covariates, the transported curve $P_{FH \rightarrow 02}$ (blue) almost completely overlapped with the target KM_{02} (black)
    - Mean absolute difference was 0.56%
- Therefore, the model is transportable for the 1L chemo group



Transportability results

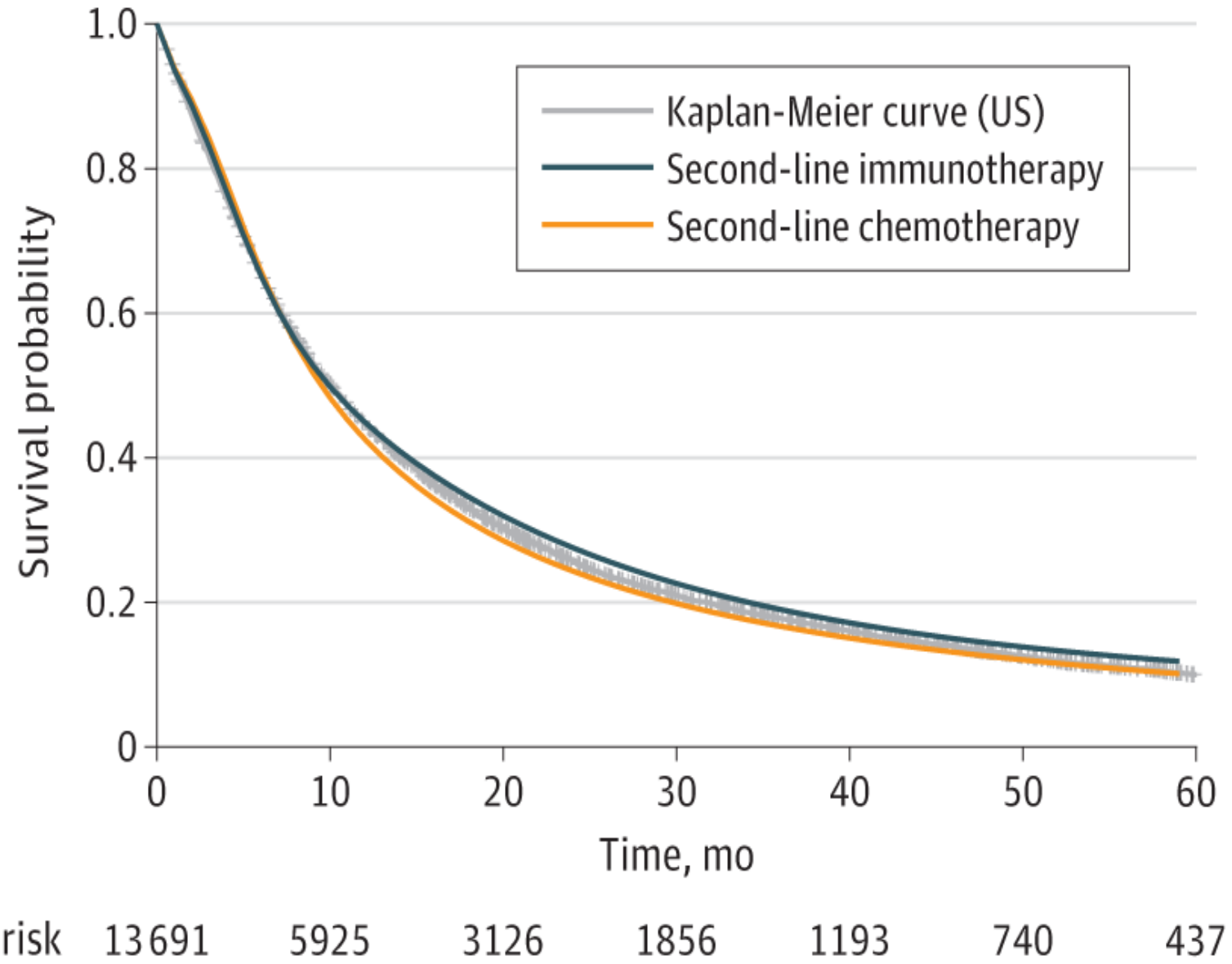
First-line pembrolizumab

- After adjustment for baseline covariates, the transported curve $P_{FH \rightarrow 02}$ (green) was similar to the target KM_{02} (black)
    - Overestimated survival initially, but progressively aligned closer
    - Mean absolute difference was 4.54%
- Before adjustment, survival curves were similar (grey and black curves)
- Negative control (purple) used a mismatched outcome model where the 1L chemotherapy model was standardized to 1L pembrolizumab covariates in Canada
    - Mean absolute difference was 6.64% and shape of curve was incompatible
- Therefore, the model is transportable for the 1L group



Bias Analysis

- Overall survival curves under hypothetical scenarios in which patients who received first-line platinum-doublet chemotherapy could **only** receive second-line immunotherapy or second-line chemotherapy, regardless of drug costs.
- The index date (time zero) corresponds to the time of initiation of first-line treatment.
- The gray Kaplan-Meier curve (US) represents observed risks.
- Numbers at risk pertain to US patients



Discussion

Transportability of overall survival for real-world patients with advanced non-small cell lung cancer from the US to Canada

Implications for regulatory and health technology assessment

Plat. doublet chemo or Pembro mono as 1L

Uncertainty around the external validity of survival outcomes derived from real-world data from US patients

- HTA decision-making **outside of the US.**

Evaluate **transportability of overall survival (OS)** for

- Transport from a large US real-world database → Canadian practices.

Analysis completed using patients with imputed data for baseline covariates

Transported OS estimates showed <5% mean absolute difference from the observed OS in the target population

- 0.56% and 4.54% respectively

Negative control analysis using a mismatched outcome model

- 6.64% discrepancy and incompatible survival curve shape.

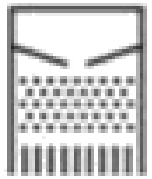
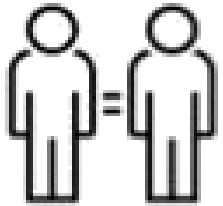
Sensitivity analysis suggests results are robust to:

- assumptions of random missingness for baseline covariates,
- unadjusted differences in baseline metastases and comorbidities
- differences in the standard of care between US and Canada*

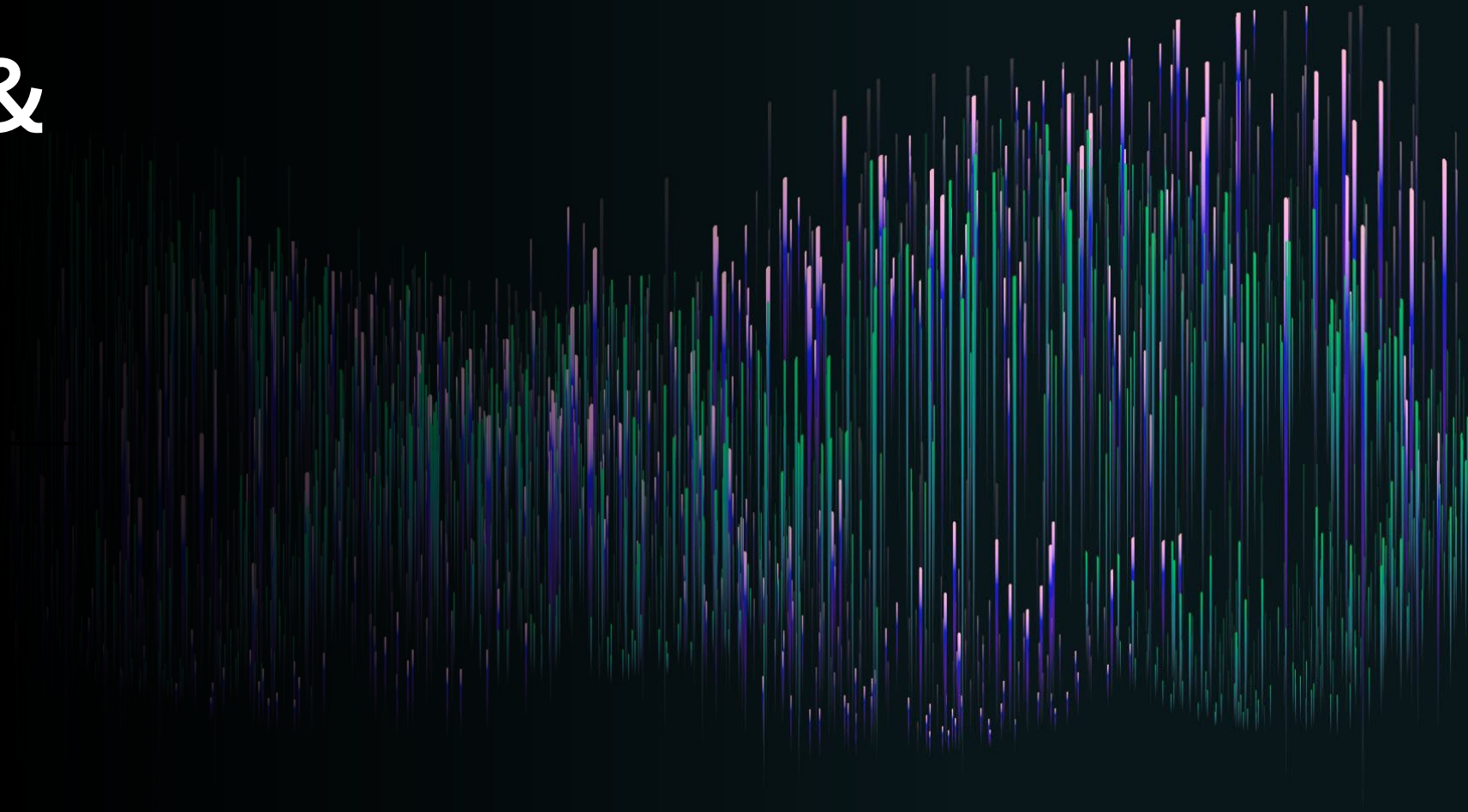
1. OS from US real-world data can be adjusted using baseline clinical characteristics to **closely approximate OS** in select groups of Canadian real-world aNSCLC patients.

2. A principled approach can be used to support regulatory decision-making and health technology assessment in target populations **outside of the US.**

Assumptions of transportability

	Exchangeability	Positivity	Consistency
	"Conditional Treatment Exchangeability"	**"Positivity of Treatment Assignment"**	**"Stable Unit Treatment Value Assumption (SUTVA) for Treatment"**
Treatment			
	Outcomes are independent of treatment assignment	Participants have a non-zero likelihood of being assigned any treatment variant	No interference from treatments received by other subjects
Selection	**"Conditional Exchangeability for Study Selection"**	**"Positivity for Study Selection"**	**"SUTVA for Selection"**
			
	All variables that might modify the treatment effect and differ between these populations are accounted for	Any subgroup has a non-zero chance of inclusion in the study sample	Outcomes remain unaffected by the participation status of other subjects

Conclusions & Future Research



Conclusions

- Demonstrates feasibility to transport OS estimates from US to Canadian patients.
- Underscores transportability analysis as a tool for confirming external validity of RWE.
- Direct implications for healthcare stakeholders in HTA decision-making.
- Ramagopalan et al's study highlights the potential of transportability in oncology.
- Sets the stage for future HTA endeavors, positioning transportability as a crucial tool in modern cancer care.

Future Areas of Research in Generalizability and Transportability Studies

- Quantitative frameworks for internal and external validity.
- Emphasis on generalizability of applied research findings.
- Addressing limitations in data availability, quality, and missing data.
- Exploration of study designs that enhance generalizability.
- Achieving consensus on “decision-grade” real-world evidence.
- Formal evaluation by regulators and HTA bodies on generalizability and/or transportability.
- Development of comprehensive frameworks and guidance on execution and interpretation of analytical methodologies.

Extensions to transportability work

Potential Criticisms

1. Impact of missing observations	• Using methods to account for missing baseline information in target population
2. Relies too much on IPD in the Target Population	• Incorporating external expert knowledge when IPD is not available for the target population
3. Specific to indication and limited treatments (narrow focus)	• Additional treatments/regimens • Expand to individuals with recurrent disease or other disease sites
4. Solidify the link between clinical and economic outcomes	• Pilot test transportability results between US and ex US via a health economic model
5. Uncertainty	• Apply various approaches to model and report uncertainty in transported estimates

Q&A

paul.arora@utoronto.ca