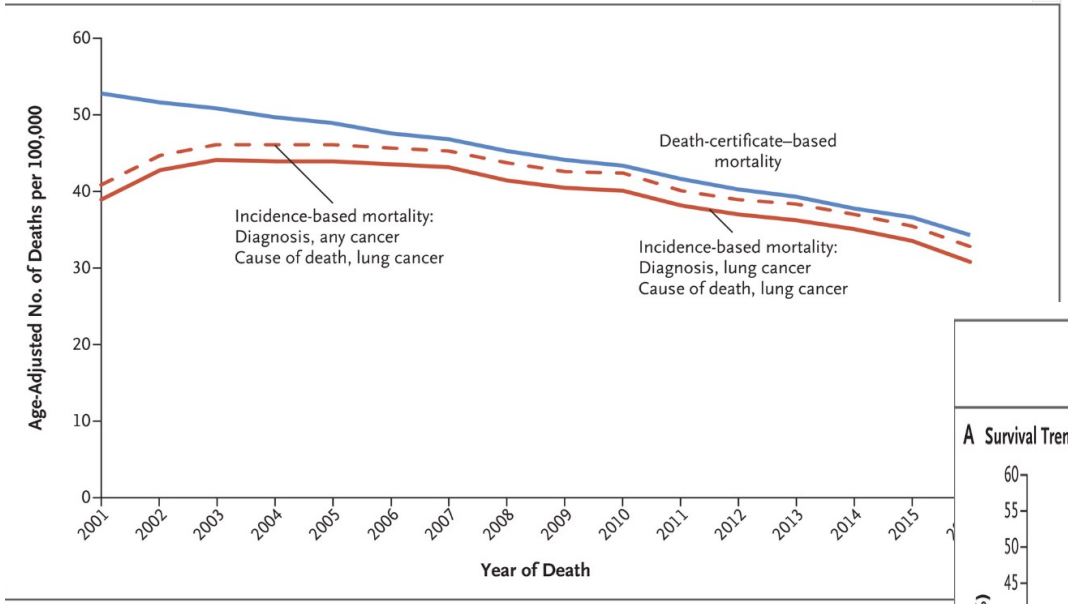
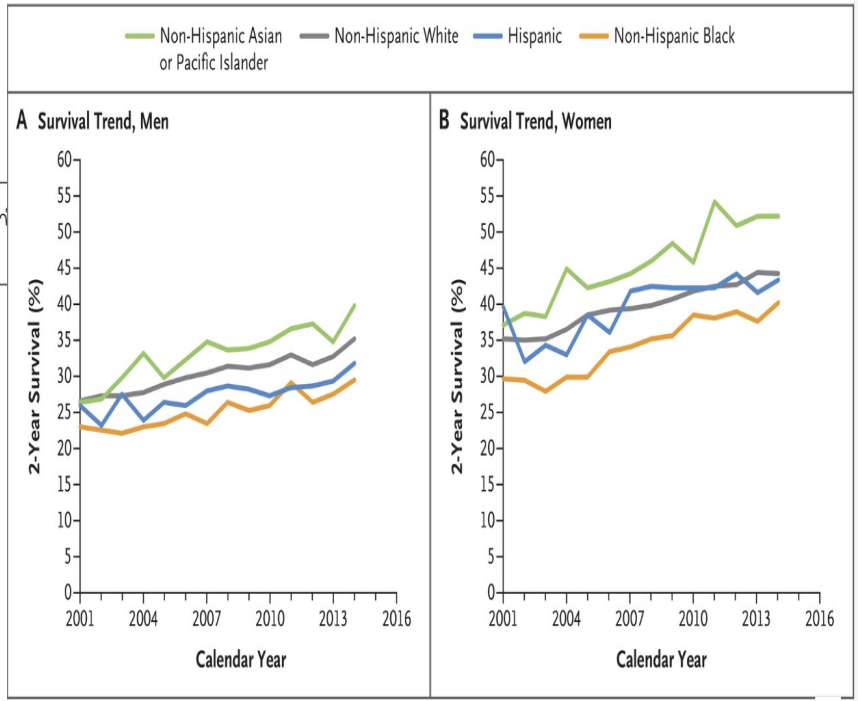
Methodologies to address concerns in using real-world evidence for measures of treatment effectiveness

Sreeram Ramagopalan, PhD

Agenda

- **Current use** of RWE in in HTA and concerns for use in decision-making
- Challenges related to **internal validity** of RWE: Quantitative bias analysis
- Challenges related to **external validity** of RWE: Transportability

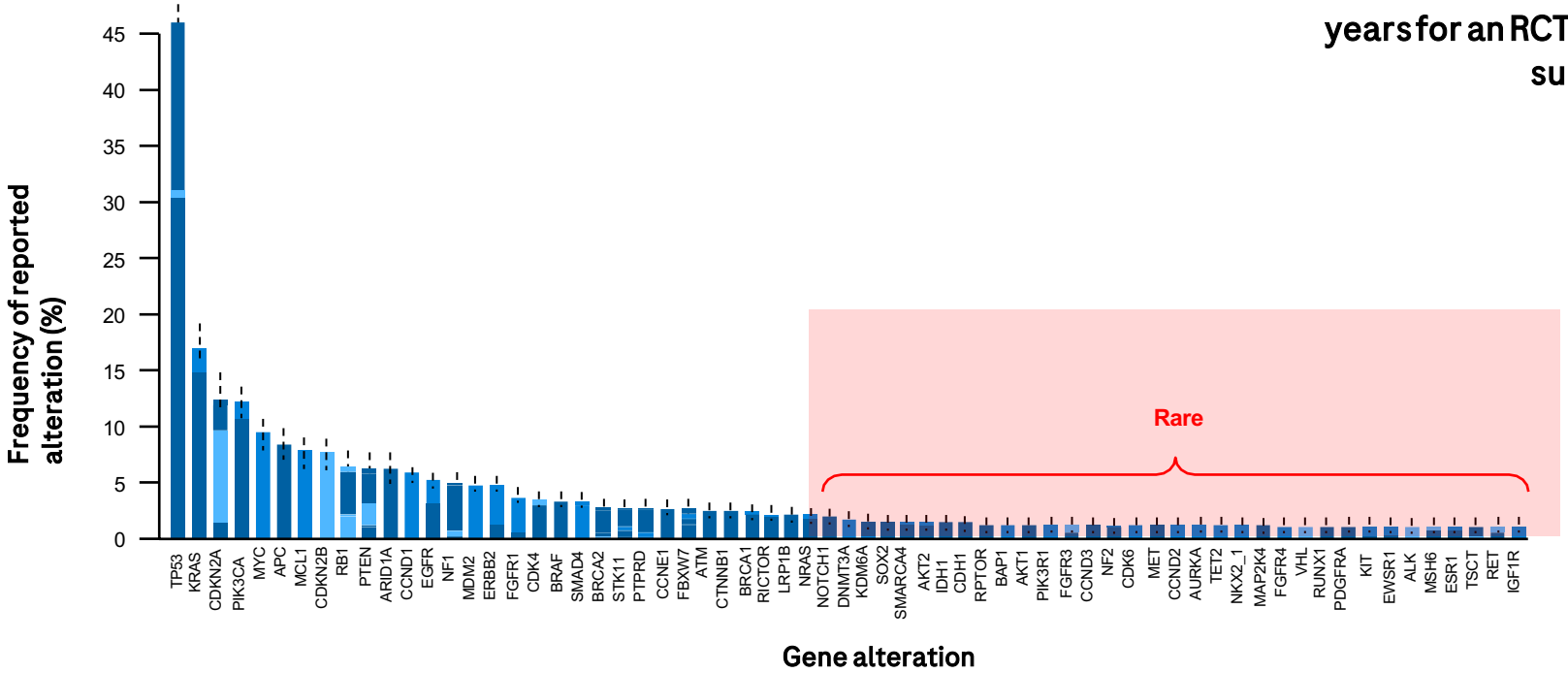

Age-Adjusted No. of Deaths per 100,000
60
50
40
30
20
10
0
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
Year of Death
Death-certificate–based mortality
Incidence-based mortality: Diagnosis, any cancer Cause of death, lung cancer
Incidence-based mortality: Diagnosis, lung cancer Cause of death, lung cancer


Non-Hispanic Asian or Pacific Islander
Non-Hispanic White
Hispanic
Non-Hispanic Black
A Survival Trend, Men
B Survival Trend, Women
2-Year Survival (%)
Calendar Year
2001
2004
2007
2010
2013
2016

Potentially targetable alterations across all cancers

17 years

Minimum time to results in years for an RCT in any NTRK sub indication (Sarcoma)



Regulators vs HTA Agencies/Payers

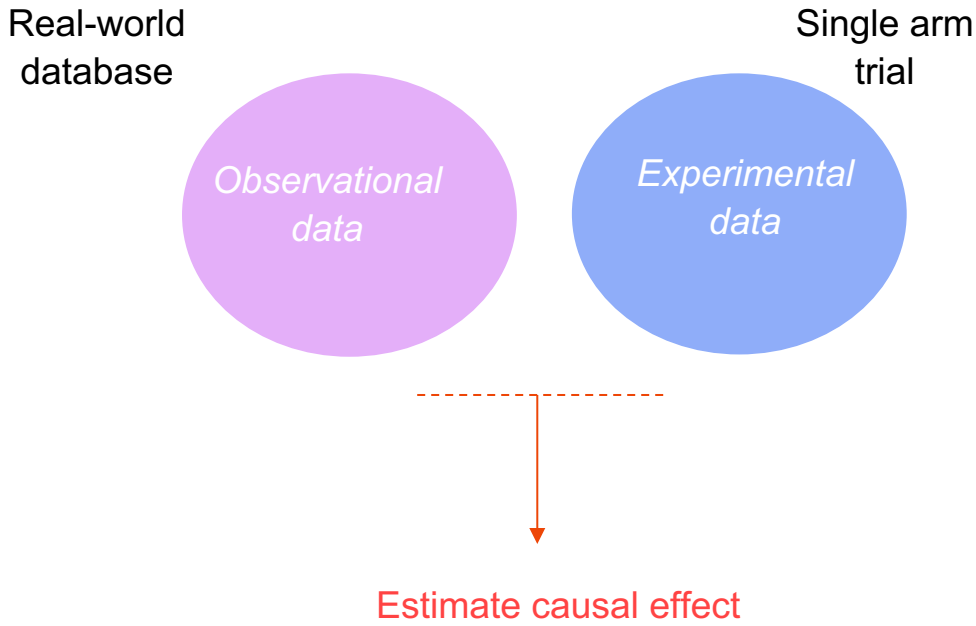
Regulators

- Positive benefit-risk of a new drug
- *Does efficacy outweigh safety?*

HTA Agencies / Payers

- Benefits over and above drugs already on the market
- *Do efficacy, safety & value outweigh those of alternative treatments?*

Synthetic Control Arm (SCA) to save the day?



Impact, acknowledgment of use of RWE by language assessment

Key

Green - language used in the HTA report interpreted as acknowledging the usefulness of RWE in the final first assessment.

Amber - RWE was acknowledged but it was unclear how or whether it was used.

Red - RWE was considered and disregarded.

Grey - RWE evidence was not found, potentially because it was not submitted or it was submitted but not mentioned in the final first report.

	Australia	Canada	England	France	Germany
Evrysdi	○	●	●	●	○
Kymriah	○	○	●	●	○
Luxturna	○	●	●	●	○
Polivy	○	●	○	○	○
Rozlytrek	○	○	○	○	●
Tecartus	●	●	●	●	○
Yescarta	●	●	●	○	○
Zolgensma	●	●	●	●	○

SCAs in HTA submissions

Criticism for bias

- HTA agencies are concerned about bias when combining experimental and observational data

- Some examples of issues related to internal validity that have been cited:
  - Unmeasured confounding in SCAs derived from RWD
  - Residual confounding by variables that are not commonly recorded in RWD, such as performance status, or unadjusted due to large amounts of missingness
    - E.g. Lack of sensitivity analysis for missingness in ECOG PS[2] score
  - Insufficient harmonization of covariates and outcomes
- Lack of data coming from "home" country
- Can quantitative bias and transportability analysis help?

[1] HTA = health technology assessment
[2] ECOG PS = Eastern Cooperative Oncology Group Performance Status

Addressing issues of internal validity in synthetic control arm (SCA) analyses

External adjustment and quantitative bias analysis (QBA)

QBASEL study

27 June 2022

QBASEL (on-going study)

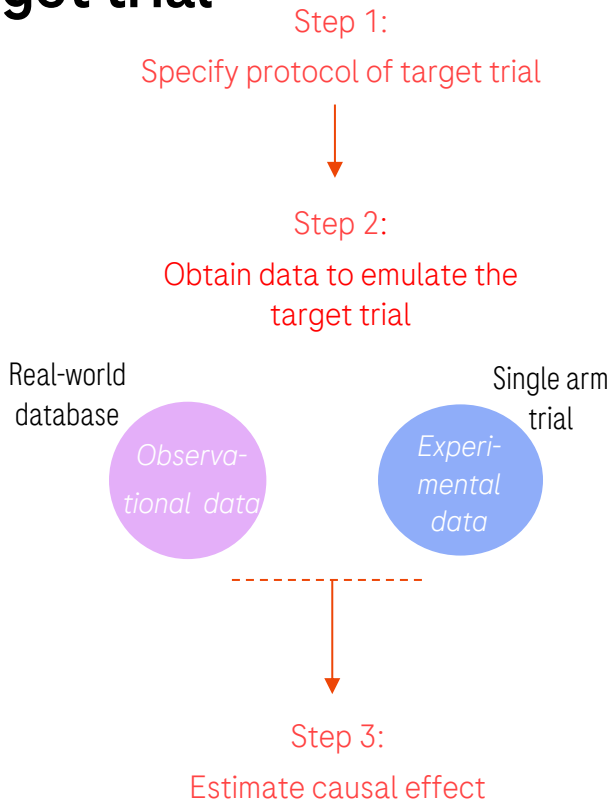
Quantitative bias analysis for comparisons between SCA from external data and lung cancer trials

- Objectives: For 14 completed randomized clinical studies in aNSCLC[1], the goals are to
  - Select **external control arms** for each study from real-world Flatiron Health database by trial emulation
  - **Adjust** for confounding and estimate hazard ratios **for overall survival (OS)** targeting an observational ITT[2] estimand
  - For each SCA analysis, summarize external information on important sources of bias and **compute bias-corrected HRs**
  - **Compare** randomized and non-randomized bias-corrected estimates
- Tipping point analyses are also planned to assess individual impact of missing data assumptions and unmeasured confounding

[1] aNSCLC = advanced non-small cell lung cancer
[2] ITT = intention-to-treat
[3] QBA = quantitative bias analysis

Overview: Emulation of control group of target trial

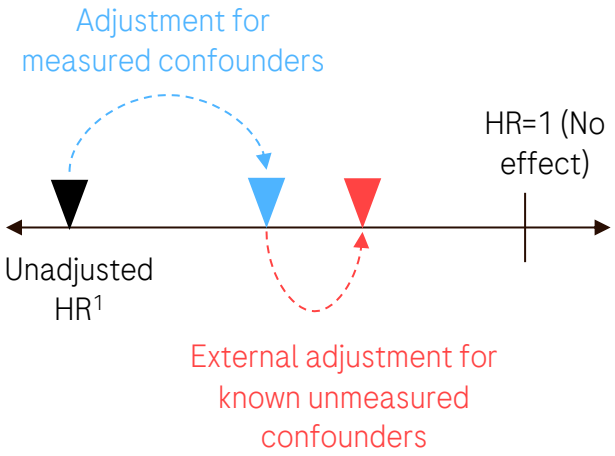
- Step 1: Specify the protocol of the target trial
  - Eligibility criteria
  - Treatment strategy (comparison is standard-of-care)
  - Outcome(s)
  - Start and end of follow-up
  - Statistical analysis
- Step 2: Obtain data to emulate the target trial
  - Recruit and follow eligible participants to treatment group of target trial (i.e., experimental data)
  - Select eligible individuals for standard-of-care group using a healthcare database (i.e., observational data)
- Step 3: Use statistical methods to adjust for differences between arms and estimate the causal effect



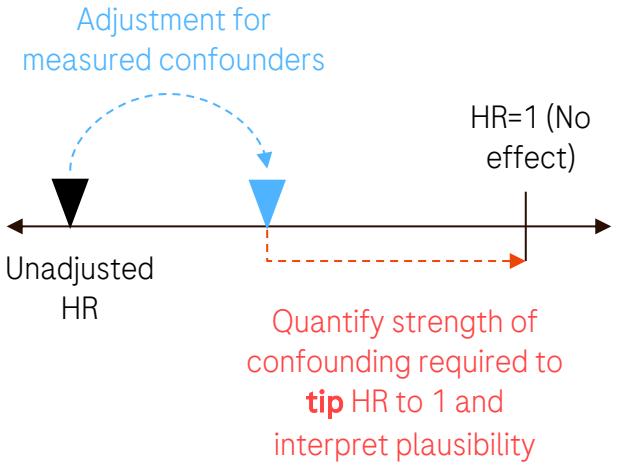
Methods for addressing bias

- Two methodologies represented below:
  - Using external information to correct for bias ("**external adjustment**")
  - **Tipping point analysis**

External adjustment



Tipping point analysis



[1]HR = hazard ratio

Collaborators

Roche

NICE National Institute for Health and Care Excellence

Cytel

HARVARD UNIVERSITY

Putnam
An Ashfield Advisory Company

Imperial College London

Universiteit Leiden
The Netherlands

Results for two trials

Some preliminary data for external bias adjustment

ALESIA

(*crizotinib* vs alectinib in ALK+[1]):

- Asian patients only
- External adjustment for a **mis**measured confounder
  - An important variable, presence of CNS[3] metastases, was not measured in-trial

GO27821

(*SOC*[2] vs SOC + onartuzumab):

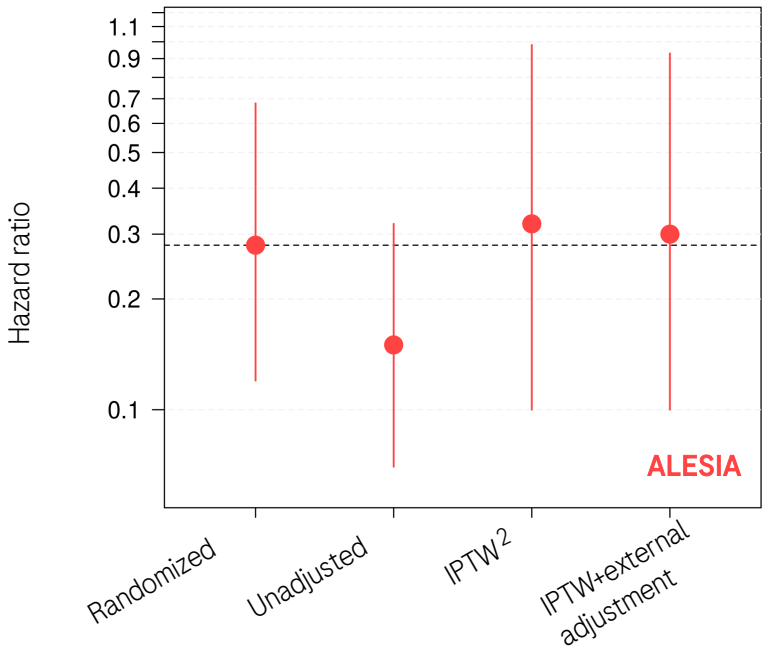
- Small sample sizes (<70 patients)
- External adjustment for an **un**measured confounder
  - A key variable, **uncontrolled** CNS[3] metastases, was measured differently in-trial than in Flatiron

[1] ALK+ = anaplastic lymphoma kinase-positive
[2] SOC = standard of care
[3] CNS = central nervous system

Results for ALESIA – Asian patients only

Some preliminary data for external bias adjustment

- After IPTW-adjustment, the HR point estimate became **notably closer** to the randomized HR
- Included **measured confounders**: age, sex, stage at diagnosis, ECOG, historical smoking status, and various labs
- For ALESIA, **presence of CNS[1] metastases** was a **mismeasured confounder**
- The trial measured various types of CNS metastases (controlled/uncontrolled), but in Flatiron "no metastasis" refers to recorded status (present/not present)
  - Thus, the distribution was **recalibrated** to that from bias parameters in literature search
- Additional external adjustment induced a small change, but brought the point **closer** to the randomized HR
- The 95% CI width has also **decreased** slightly

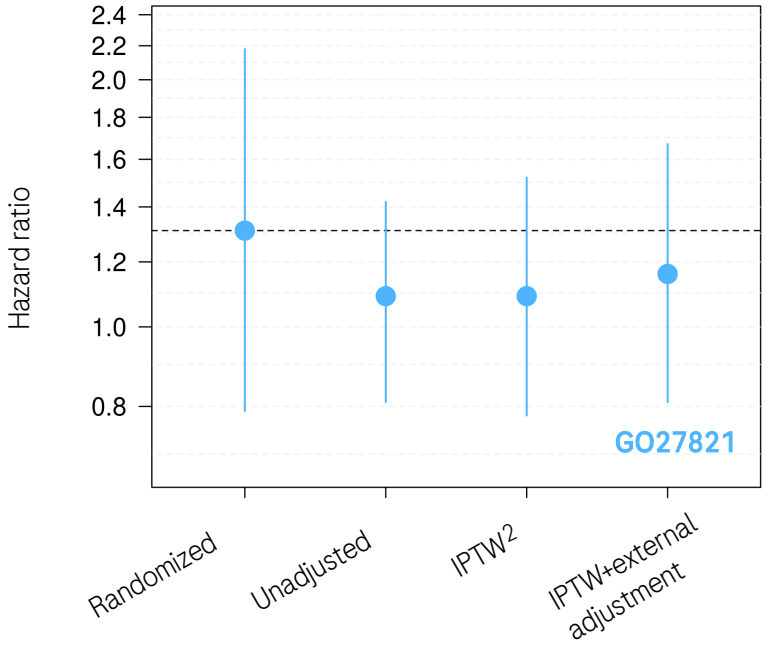


[1] CNS = central nervous system
[2] IPTW = inverse probability treatment weighting

Results for GO27821 – Small sample sizes

Some preliminary data for external bias adjustment

- After IPTW-adjustment, the HR point estimate **did not** get much closer to the randomized HR
- Included **measured confounders**: age, sex, stage at diagnosis, ECOG, historical smoking status, and various labs
- In GO27821, presence of **uncontrolled CNS[1] metastases** was an **unmeasured confounder**
  - Since CNS metastases was not measured in-trial, values were **simulated** based on **bias parameters** pertaining to:
    - prevalence,
    - association with treatment, and
    - association with the outcome extracted from the literature.
- Additional external adjustment brought the point estimate for HR marginally **closer** to the randomized HR



[1] CNS = central nervous system
[2] IPTW = inverse probability treatment weighting

Conclusions (1/2)

- Overall results accounted for a **wide range of measured confounders** in each study ranging from demographics to labs.
  - Expert opinion and literature searches substantiated the extent of what external factors and parameters were relevant to two initial studies.

- **Clinically relevant bias parameters** either addressed unmeasured CNS metastases (GO27821) or recalibrating mismeasured CNS metastases (ALESIA) and accounted for their bias in a RWD setting.
  - In each study, results showed a bias-adjusted estimate, which moved **closer** to their respective RCT HR in comparison to measured-confounder only estimates.

Conclusions (2/2)

- **Initial results suggest the utility of addressing bias in a RWD setting.**
  - Two trials show that accounting for a clinically relevant unmeasured or recalibrating a mismeasured confounder can help address RWD-concerns of bias when evaluating treatment estimates relative to a RCT gold-standard

- Additional evaluation of **robustness** of results is needed to assess under what conditions the external bias adjustment holds relative to not accounting for it when assessing its relative shift to a RCT estimate

- Ongoing research toward QBA helps **answer pragmatic questions**:
  - Will external bias adjustment be acceptable in HTA?

Addressing issues of external validity in synthetic control arm (SCA) analyses

Transportability analysis of survival patterns between real-world data sources

27 June 2022

External validity of real-world evidence (RWE)

How valid is RWE beyond the study sample used for analysis?

- RWE is increasingly being used to support randomized evidence in health technology assessments (HTAs)
- Issues of **internal validity** such as confounding and bias receive much attention
- Decision-makers are also concerned about **generalizability to the target population**
  - E.g., Impact of differences in patient characteristics, treatment guidelines, clinical practice and patient preferences across settings or countries
- Can RWE potentially be used to support decision-making in a target population where **direct estimation of treatment effects is infeasible**?

Concepts in external validity

- ***Generalizability***: study sample is a subset of target population
  - E.g., study sample = US-based EHR
    target population = entire US population

- ***Transportability***: study sample is partially or completely external to the target population
  - E.g., study sample = US patients
    target population = Canadian patients

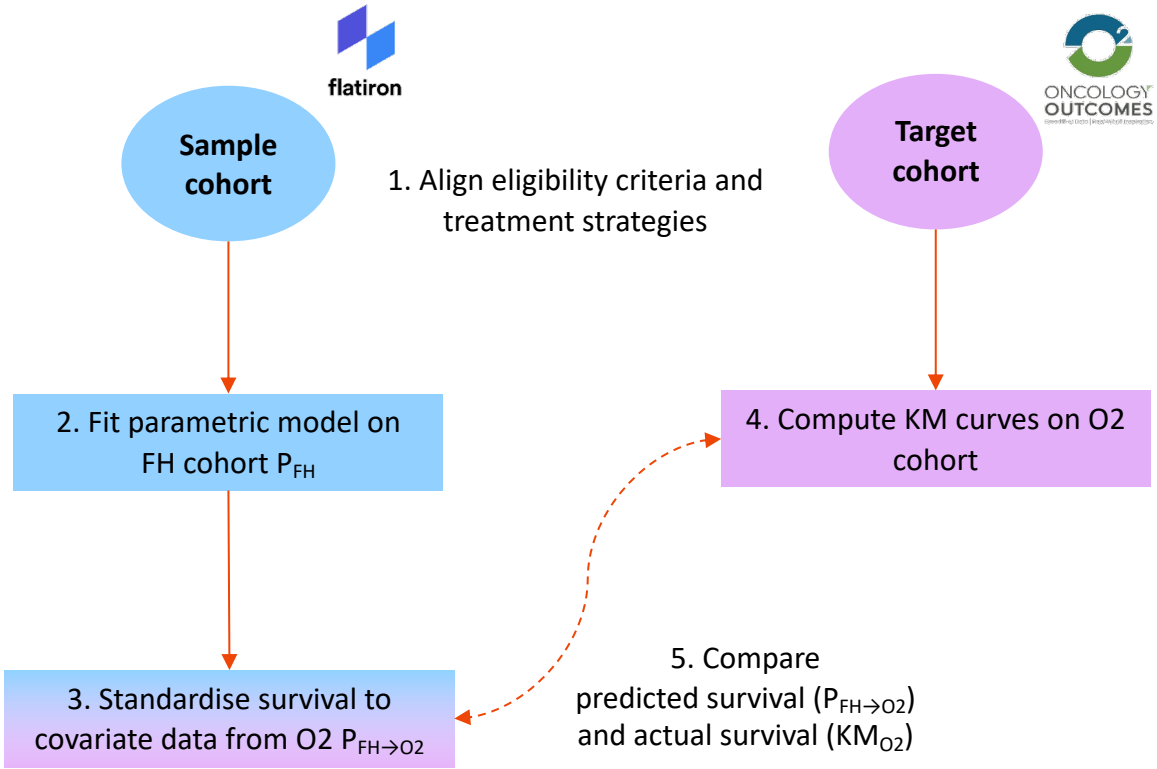
Objectives

Can RWE from US-based oncology EHR be potentially used to support HTA decisions in Canada?

- We tested if adjustment for patient characteristics for US patients with aNSCLC was sufficient to approximate OS in Canadian patients
  - ***Study sample***: US patients from Flatiron Health (FH) database
  - ***Target sample***: Canadian patients from the province of Alberta, Oncology Outcomes Research Initiative (O2)

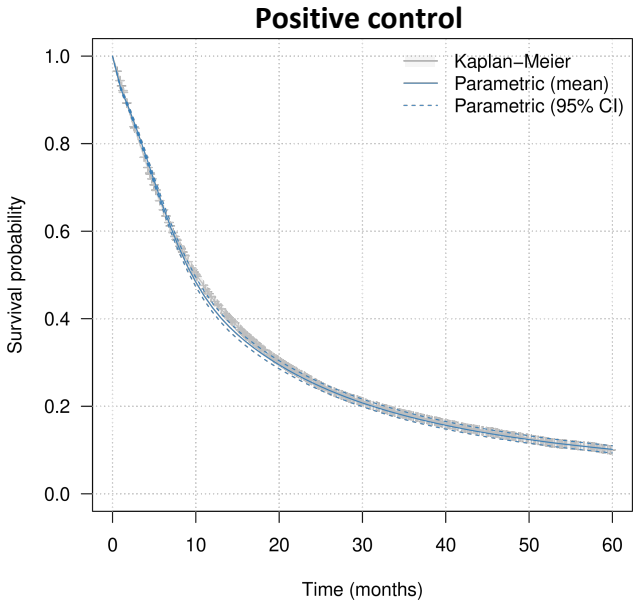
- **Treatment group**: Initiators of 1L platinum doublet chemotherapy ("1L chemotherapy/chemo")

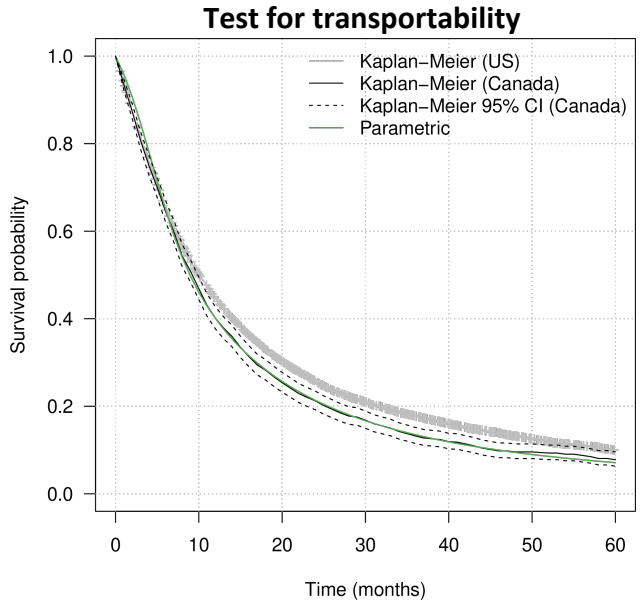
Study overview



Results

- **The parametric model P_{FH} (blue curve)** fitted on complete-case US data from FH fits OS well on the total sample patient data KM_{FH} (grey curve)
- After standardisation to baseline covariates in the target cohort, **the transported curve $P_{FH \rightarrow O2}$ (green)** almost completely overlapped with the **target Kaplan-Meier (black)**
  - Mean absolute difference was **0.56%**





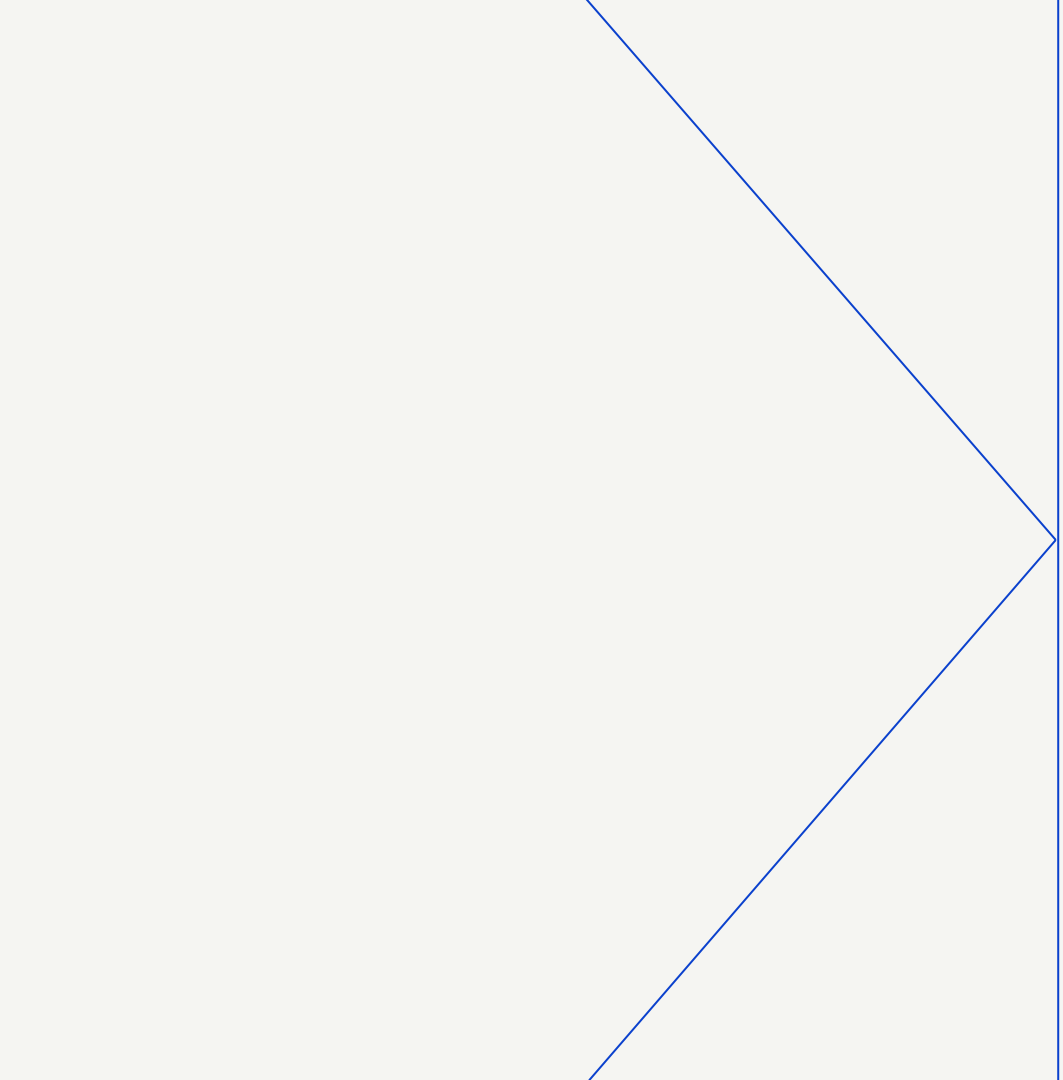
Conclusions

- Adjustment for baseline patient characteristics was able to estimate OS that was **comparable** to the target population in Canada for this treatment group
  - This result holds in general for other scenarios tested

- **Limitations**:
  - Differences in healthcare delivery, patient adherence, staff experience and standards regarding treatment dosage and/or frequency between US and Canada were not examined

- Given that the absolute risks (i.e., OS curves) were transportable, we expect that relative risks (such as hazard ratios) also transport for overall populations
  - Although we tested only select treatment groups and should be interpreted in this context, the consistency of our results suggest that for aNSCLC, **if** baseline patient characteristics are similar, **then** real-world OS may be transportable in general from the US to Canada

- **Can this help with HTA concerns when RWD does not come from the "home" population?**



Appendix B

ALESIA

- Comparison
  - **Experimental arm**: Alectinib
  - **Control arm**: Crizotinib
- Relaxed control arm start date one year to increase patient set Age 18 years or older at index date
- **Eligibility criteria**
  - ECOG of 0 through 2
  - Adequate hematological functions defined using labs measured within 7 days of index date
  - History of comorbidities – CNS metastases, hepatitis, GI disorders, autoimmune disorders, HIV
  - Patients in RWD must have initiated control treatment within the minimum and maximum date of randomization in the RCT (relaxed start date by 1 year to increase patient set (August 3, 2015 instead of August 3, 2016)

Appendix C

ALESIA (adjusted)

- Included measured confounders for weighting and adjustment of Alectinib effect:
- Age at treatment start date, ECOG, sex, stage at diagnosis, historical/present smoking status, time from treatment start date to January 1, 2011, and labs (alanine transaminase, albumin, creatinine, and white blood cell count)
- Considerations for bias in the recalibration of mismeasured central nervous system (CNS) metastases

Relationship	**Value**
Prevalence of CNS metastases	35%
Association of CNS metastases and overall survival	Normal(Log(1.4), .15)*
Association of CNS metastases and treatment	Normal(Log(1.3), .1)*
*Simulated from normal distribution and values represent mean (natural log of the odds ratio) and standard deviation, respectively.	

Appendix D

GO27821

- Comparison
  - **Control arm**: bevacizumab + platinum + paclitaxel (+ placebo in RCT, placebo not emulated in RWD)
  - **Experimental arm**: MetMAb + bevacizumab + platinum + paclitaxel
- **Eligibility criteria**
  - Age 18 years or older at index date
  - ECOG of 0 or 1
  - Non-squamous histology
  - Adequate hematological functions defined using labs measured within 7 days of index date
  - History of comorbidities – CNS metastases, diabetes with complications, uncontrolled hypertension, cardiac arrythmia and other cardiovascular complications, HIV
  - Patients in RWD must have initiated control treatment within the minimum and maximum date of randomization in the RCT (April 4, 2012 to June 28, 2013)

Appendix E

GO27821

- **Measured confounders**: age, sex, ECOG PS, cancer stage at diagnosis, smoking status, race, labs (albumin, creatinine, ALT)
- **Unmeasured confounders**: CNS metastases not excluded (including unknown or stable and/or treated disease)
  - Prevalence: 20%
  - Conditional association with OS (log HR [std deviation]): 0.7 [0.1]
  - Conditional association with treatment (log OR [std deviation]): 0.5 [0.2]

- No large weights observed (all weights were <7)

References

1. Bareinboim, Elias and Pearl, Judea. "A General Algorithm for Deciding Transportability of Experimental Results" Journal of Causal Inference, vol. 1, no. 1, 2013, pp. 107-134. https://doi.org/10.1515/jci-2012-0004
2. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: the ACTG 320 trial. Am J Epidemiol. 2010;172(1):107–115
3. Hernán MA, Vanderweele TJ. Compound treatments and transportability of causal inference. Epidemiology. 2011;22(3):368–377.
4. Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR. Generalizing Study Results: A Potential Outcomes Perspective [published correction appears in Epidemiology. 2018 Mar;29(2):e16]. Epidemiology. 2017;28(4):553-561. doi:10.1097/EDE.0000000000000664
5. Snowden JM, Rose S, Mortimer KM. Implementation of G-Computation on a Simulated Data Set: Demonstration of a Causal Inference Technique. Am J Epidemiol. 2011;173(7):731-738. doi:10.1093/aje/kwq472
6. Westreich D, Edwards JK, Lesko CR, Stuart E, Cole SR. Transportability of Trial Results Using Inverse Odds of Sampling Weights. Am J Epidemiol. 2017 Oct 15;186(8):1010-1014. doi: 10.1093/aje/kwx164. PMID: 28535275; PMCID: PMC5860052.

Baseline demographics: Platinum doublet-first line chemotherapy

Category		US	Canada	SMD
Sample size		8447	1476	
Age at index date (mean (SD))		67.34 (9.25)	65.07 (9.53)	0.242
Sex (%)	Female	3602 (42.6)	703 (42.8)	0.004
	Male	4845 (57.4)	773 (52.4)	
Cancer stage at diagnosis (%)	IIIb/IIIc	2679 (31.7)	264 (17.9)	0.324
	IV	5768 (68.3)	1212 (82.1)	
ECOG PS (%)	0-1	6625 (78.4)	1091 (73.9)	0.106
	2+	1822 (21.6)	385 (26.1)	
Tumor histology (%)	Non-squamous cell carcinoma	5168 (61.2)	1228 (83.2)	0.507
	Squamous cell carcinoma	3279 (38.8)	248 (16.8)	
Smoking history (%)	Ever	7808 (92.4)	1343 (91.0)	0.051
	Never	639 (7.6)	133 (9.0)	
Time from diagnosis to index date (months) (median [IQR])		1.12 [0.72, 1.63]	1.84 [1.25, 2.76]	0.33
Time since 2011-01-01 (years) (median [IQR])		5.28 [3.53, 7.02]	4.58 [2.50, 6.44]	0.297
Comorbidities (%)	0	6188 (73.3)	837 (56.3)	0.362
	1+	2259 (26.7)	639 (43.3)	
Number of sites of metastases	0-1	730 (86.5)	877 (59.5)	0.638
	2+	1143 (13.5)	596 (40.5)	

* Note that the values for O2 represent summaries after single imputation of missing baseline ECOG PS and smoking history. Without imputation, due to missingness in ECOG PS and smoking history, sample sizes would be much lower in O2.

Baseline demographics: Pembrolizumab first line monotherapy

Category		US	Canada	SMD
Sample size		1653	287	
Age at index date (mean (SD))		71.64 (9.81)	69.01 (8.95)	0.28
Sex (%)	Female	803 (48.6)	149 (51.9)	0.066
	Male	850 (51.4)	138 (48.1)	
Cancer stage at diagnosis (%)	IIIb/IIIc	94 (5.7)	27 (9.4)	0.14
	IV	1559 (94.3)	260 (90.6)	
ECOG PS (%)	0-1	1107 (67.0)	209 (72.8)	0.127
	2+	546 (33.0)	78 (27.2)	
Tumor histology (%)	Non-squamous cell carcinoma	1256 (76.0)	244 (85.0)	0.229
	Squamous cell carcinoma	397 (24.0)	43 (15.0)	
Smoking history (%)	Ever	1521 (92.0)	255 (88.9)	0.106
	Never	132 (8.0)	32 (11.1)	
Time from diagnosis to index date (months) (median [IQR])		1.25 [0.89, 1.81]	1.81 [1.30, 2.52]	0.148
Time since 2011-01-01 (years) (median [IQR])		7.68 [6.78, 8.68]	7.89 [7.28, 8.52]	0.192
Comorbidities (%)	0	1062 (64.2)	169 (58.9)	0.109
	1+	591 (35.8)	118 (41.1)	
Number of sites of metastases	0-1	1367 (82.7)	170 (59.6)	0.527
	2+	286 (17.3)	115 (40.4)	

* Note that the values for O2 represent summaries after single imputation of missing baseline ECOG PS and smoking history. Without imputation, due to missingness in ECOG PS and smoking history, sample sizes would be much lower in O2.

Baseline Characteristics

Category		US	Canada	SMD
Sample size		301	138	
Age at index date (mean (SD))		65.00 (9.52)	63.53 (8.67)	0.161
Sex (%)	Female	130 (43.2)	68 (49.3)	0.123
	Male	171 (56.8)	70 (50.7)	
Cancer stage at diagnosis (%)	IIIb/IIIc	30 (10.0)	23 (16.7)	0.198
	IV	271 (90.0)	115 (83.3)	
ECOG PS (%)	0-1	223 (74.1)	N/A	N/A
	2+	78 (25.9)	N/A	
Tumor histology (%)	Non-squamous cell carcinoma	227 (75.4)	92 (66.7)	0.193
	Squamous cell carcinoma	74 (24.6)	46 (33.3)	
Smoking history (%)	Ever	282 (93.7)	126 (91.3)	0.091
	Never	19 (6.3)	12 (8.7)	
Time from diagnosis to index date (months) (median [IQR])		6.67 [4.41, 9.37]	10.13 [7.17, 15.58]	0.702
Time since 2011-01-01 (years) (median [IQR])		3.50 [2.65, 4.24]	5.33 [3.86, 8.03]	0.783
Comorbidities (%)	0	231 (76.7)	N/A	N/A
	1+	70 (23.3)	N/A	
Number of sites of metastases	0-1	247 (82.1)	N/A	N/A
	2+	54 (17.9)	N/A	

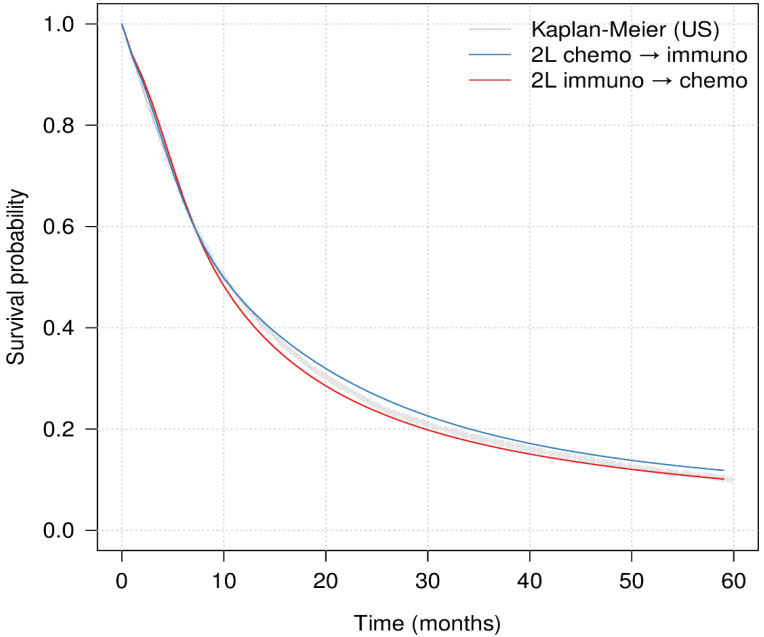
- Note that the values for O2 represent summaries after single imputation of missing baseline ECOG PS and smoking history. Without imputation, due to missingness in ECOG PS and smoking history, sample sizes would be much lower in O2.
- ECOG PS, comorbidities and metastases were not reliably available beyond the 1L setting in the Canadian database, therefore only unadjusted outcomes were compared

Bias analyses

- **Tipping point analyses** for impact of comorbidities and number of metastatic sites
  - Resulted in implausible values required to tip results to >5% mean absolute difference
  - Results

- Comparing 2L immuno vs chemo for these patients using G-computation, the maximum risk difference of 3.50% [95% CI 1.96-4.97%] was observed at month 20 after 1L initiation, and the mean absolute difference over 60 months was 2.13%
  - Unlikely to tip our results under plausible scenarios of differences in 2L for initiators of 1L chemotherapy between US and Canada

OS for 2L immuno vs 2L chemo amongst initiators of 1L chemo



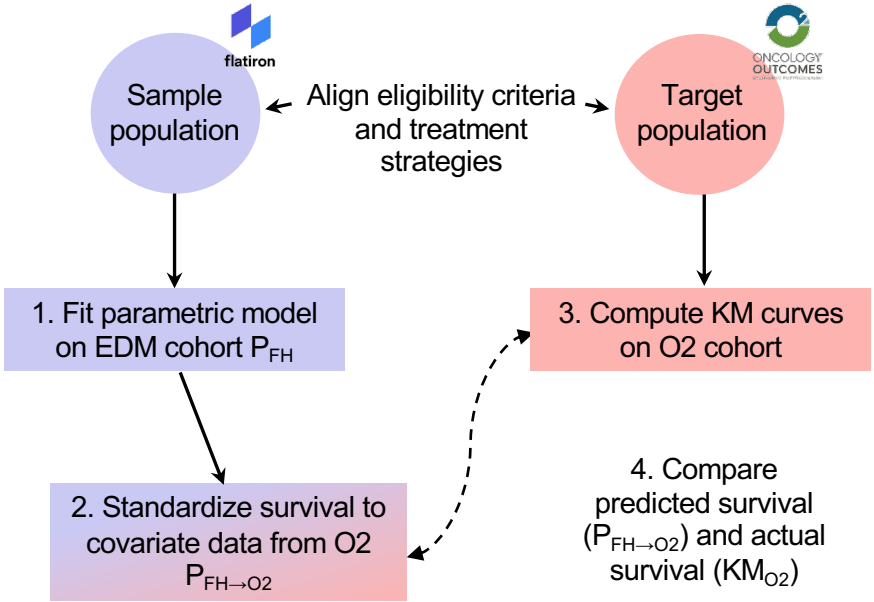
Cytel

Transportability US to CAN

2022 March 28 – v2.0

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_{O2} 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 $P_{FH \rightarrow O2}$ represents the model adjusted for individual-level baseline covariates from the target group O2
  - $P_{FH \rightarrow O2} \approx KM_{O2}$ if transportability "holds". A threshold of <5% mean absolute difference between $P_{FH \rightarrow O2}$ and KM_{O2} implied sufficient similarity for this study.



1L Pt doublet chemotherapy

Final results

Eligibility criteria

- Initiated Pt doublet chemotherapy as first-line systemic therapy after diagnosis of aNSCLC
  - Defined as cisplatin or carboplatin with one of paclitaxel, pemetrexed, gemcitabine, vinorelbine or etoposide starting January 1, 2011
  - Any dose and regimen was permitted at the discretion of the treating physician

- At index date (1L initiation)
  - ≥18 years of age
  - Incident diagnosis of advanced NSCLC
  - Any cancer histological type except NOS

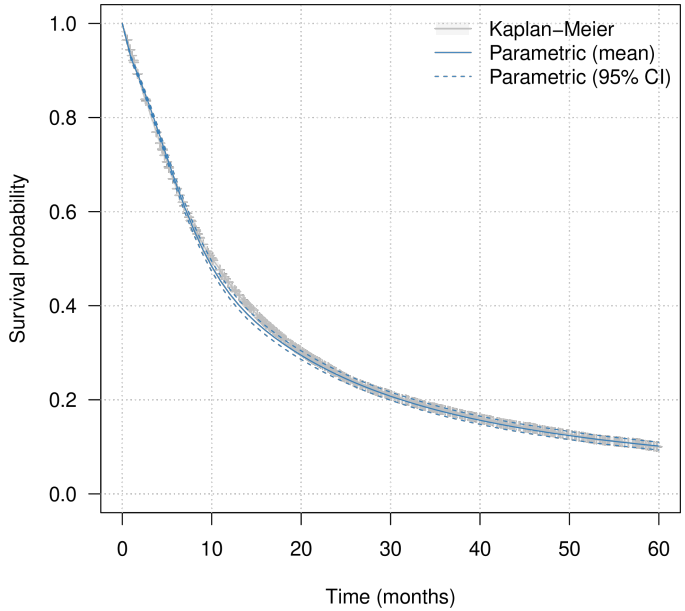
Baseline characteristics

	Category	US	Canada	SMD
Sample size		8447	1476	
Age at index date (mean (SD))		67.34 (9.25)	65.07 (9.53)	0.242
Sex (%)	Female	3602 (42.6)	703 (42.8)	0.004
	Male	4845 (57.4)	773 (52.4)	
Cancer stage at diagnosis (%)	IIIb/IIIc	2679 (31.7)	264 (17.9)	0.324
	IV	5768 (68.3)	1212 (82.1)	
ECOG PS (%)	0-1	6625 (78.4)	1091 (73.9)	0.106
	2+	1822 (21.6)	385 (26.1)	
Tumor histology (%)	Non-squamous cell carcinoma	5168 (61.2)	1228 (83.2)	0.507
	Squamous cell carcinoma	3279 (38.8)	248 (16.8)	
Smoking history (%)	Ever	7808 (92.4)	1343 (91.0)	0.051
	Never	639 (7.6)	133 (9.0)	
Time from diagnosis to index date (months) (median [IQR])		1.12 [0.72, 1.63]	1.84 [1.25, 2.76]	0.33
Time since 2011-01-01 (years) (median [IQR])		5.28 [3.53, 7.02]	4.58 [2.50, 6.44]	0.297
Comorbidities (%)	0	6188 (73.3)	837 (56.3)	0.362
	1+	2259 (26.7)	639 (43.3)	
Number of sites of metastases	0-1	730 (86.5)	877 (59.5)	0.638
	2+	1143 (13.5)	596 (40.5)	

* Note that the values for O2 represent summaries after single imputation of missing baseline ECOG PS and smoking history. Without imputation, due to missingness in ECOG PS and smoking history, sample sizes would be much lower in O2.

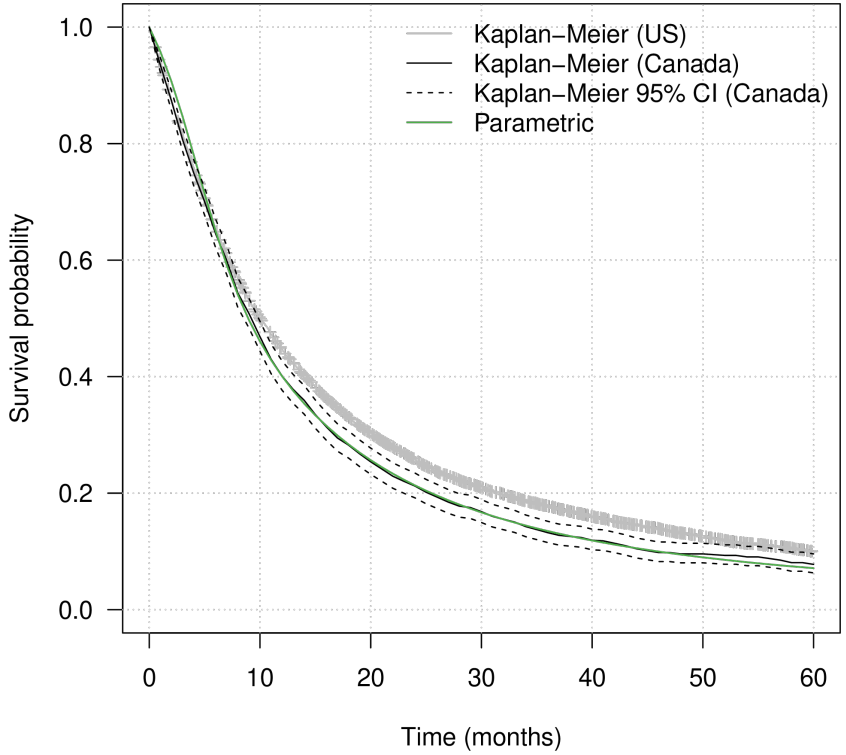
Goodness-of-fit (positive control)

- The parametric model P_{FH} (blue curve) fitted on complete-case US data fits OS well on the total sample patient data KM_{FH} (grey curve)
    - 60 months of follow-up was modelled for 1L chemotherapy



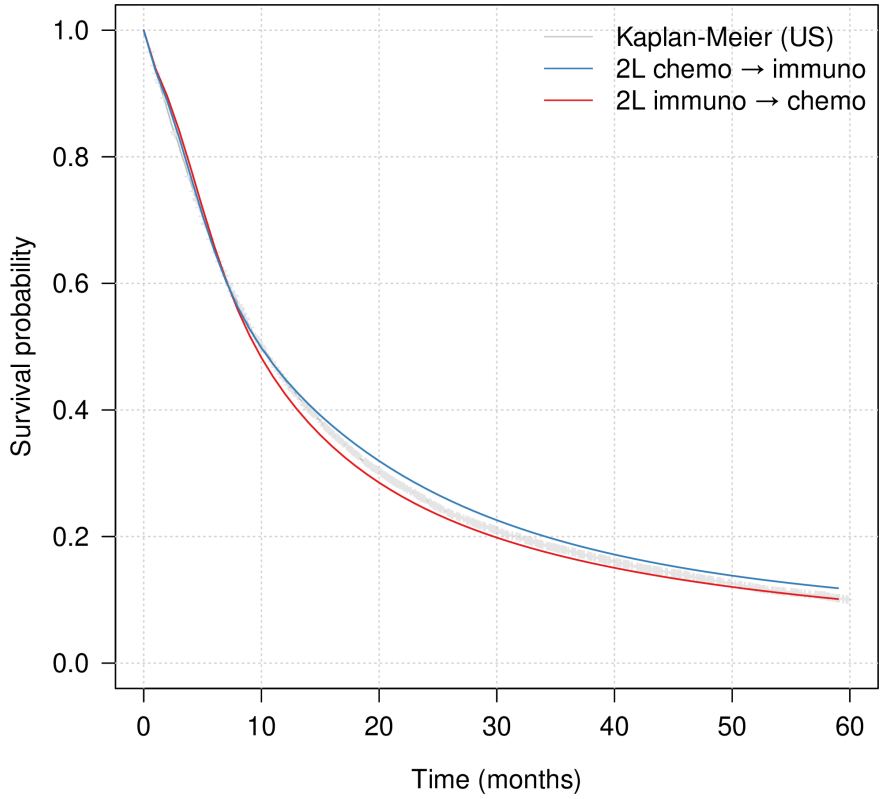
Transportability results

- After adjustment for baseline covariates, the transported curve $P_{FH \rightarrow O2}$ (green) almost completely overlapped with the target KM_{O2} (black)
  - Mean absolute difference was 0.56%
- Therefore, the model is transportable for the 1L group



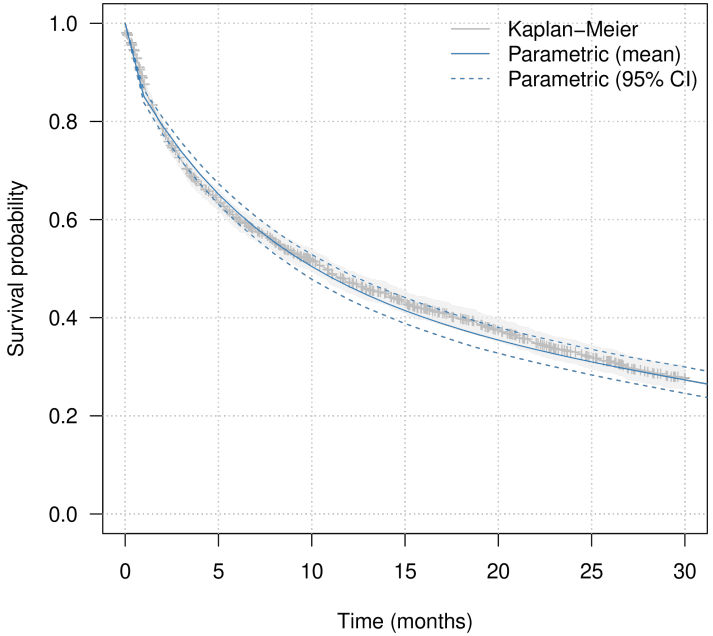
QBA summary

- Tipping points were not identified for presence of metastases; for comorbidities, tipping points were identified at >80% prevalence in the US data
    - This is a highly implausible prevalence
- (Ignoring drug costs) As shown in the figure, under worst-case hypothetical scenarios comparing 2L chemotherapy versus 2L PD-(L)1 immunotherapy amongst the 1L chemotherapy group, the maximum discrepancy was 3.50% [95% CI 1.96-4.97%]
    - Maximum absolute difference was 2.13%
- According to results from multi-country analysis (not shown here), 39% of patients in FH EDM receive 2L immunotherapy vs 18% in O2
- Therefore, under plausible scenarios, our results are robust for 1L chemotherapy group



Goodness-of-fit (positive control)

- The parametric model P_{FH} (blue curve) fitted on complete-case US data fits OS well on the total sample patient data KM_{FH} (grey curve)
    - 30 months of follow-up was modelled for 1L chemotherapy



Transportability results

- After adjustment for baseline covariates, the transported curve $P_{FH \rightarrow O2}$ (green) was similar to the target KM_{O2} (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