Using Probabilistic Quantitative Bias Analysis (QBA) to Account for Unmeasured Confounders When Estimating Treatment Effects in Real-World Data

Virtual ISPOR 2021

May 17, 2021

Speakers

Paul Arora

PhD

VP, Advanced Epidemiology

Cytel

Alind Gupta

PhD

Research Principal, Data Science

Cytel

Samantha Wilkinson

PhD

Data Scientist

Roche

Quantitative Bias Analysis

Introduction

Poll

1. How familiar are you with quantitative bias analysis?
2. Have you had to review a methods/results section of a paper or submission that incorporated QBA?
3. Are you using QBA in any current HEOR analyses?
4. Do you think QBA presents a satisfactory approach to handling bias commonly found in real-world data?
5. Should QBA analyses be required for ALL submissions (regulatory or payer) where RWD is incorporated?

Cytel

Challenges when using RWD

- RWD is increasingly being used in regulatory and payer settings
  - External comparator arm for single-arm trial
- Risk of bias in non-randomized studies can pose a threat to their validity
  - Statistical methods can reduce confounding due to lack of treatment randomization but other threats to validity remain
    - Unmeasured confounding
  - Regulators and payers remain reluctant to use in decision-making.
- In oncology, real-world studies often contain significant amounts of missing observations for important prognostic variables
  - Not uniformly recorded in routine clinical practice, such as Eastern Cooperative Oncology Group (ECOG) performance status
- Findings from such studies can be prone to bias if assumptions about the mechanism of missingness are incorrectly specified, or if unmeasured confounders are not accounted for

Cytel

Challenges when using RWD

Assumptions

- Perfect measurement of exposures and outcomes
- No difference in loss to follow-up between groups
- No unmeasured confounding

Bias

- Misclassification
- Loss-to-follow up
- Unmeasured confounding

Groenwold, RHH. 2021 HTAi

Cytel

Quantitative Bias Analysis (QBA)

- It is a type of sensitivity analysis
- An overarching term applied to methods that estimate quantitatively the direction, magnitude, and uncertainty associated with systematic errors that influence measures of associations. (Lash et al 2016)
- QBA methods also used to identify the impact of **missing data** and **residual confounding** on results from non-randomized studies using RWD
- QBA provides a way to assess the sensitivity of adjusted measures of association of exposures and outcomes to systematic bias
  - Computes them over a **range of assumptions** about missing observations in RWD and unmeasured confounders
  - A method to quantify impact of unmeasured confounding (not control for it) (Lash et al 2014)
- The robustness of an effect can be evaluated based on how drastic or implausible these assumptions have to be for the **observed association to be nullified.**

Lash, Timothy L., et al. "Good practices for quantitative bias analysis." *International journal of epidemiology* 43.6 (2014): 1969-1985.
Lash, Timothy L., et al. "Quantitative bias analysis in regulatory settings." *American journal of public health* 106.7 (2016): 1227-1230.

Cytel

Sensitivity Analysis

Focused

- “Simpler” scenarios
- Well characterized problem
- E-value

Broad

- Multiple scenarios and assumptions
- Probabilistic bias analysis
- Bayesian framework

Groenwold, RHH. 2021 HTAi

Cytel

A valuable approach under uncertainty

- Regulatory agencies and payers must weigh the risks and benefits of a product and strive to communicate clearly and transparently the scientific basis of their decisions. (Lash et al 2016)
  - Those decisions must sometimes be made even when the available data are imperfect.
- While random error is nearly always quantified in presentations of epidemiological research, the impact of systematic error on study findings is rarely quantified.(Lash et al 2016)
- QBA provides regulatory and payer bodies a tool to conduct reasoning under uncertainty.
- Bias analysis may not be necessary for studies that are **descriptive or exploratory** in nature, or that do not address causal links between exposures and outcomes.
  - Quantifying bias is important for studies assessing causal associations, or when policy decisions may be based on the study's results.

Lash, Timothy L., et al. "Quantitative bias analysis in regulatory settings." *American journal of public health* 106.7 (2016): 1227-1230.

Cytel

Quantitative Bias Analysis

Theory and methods

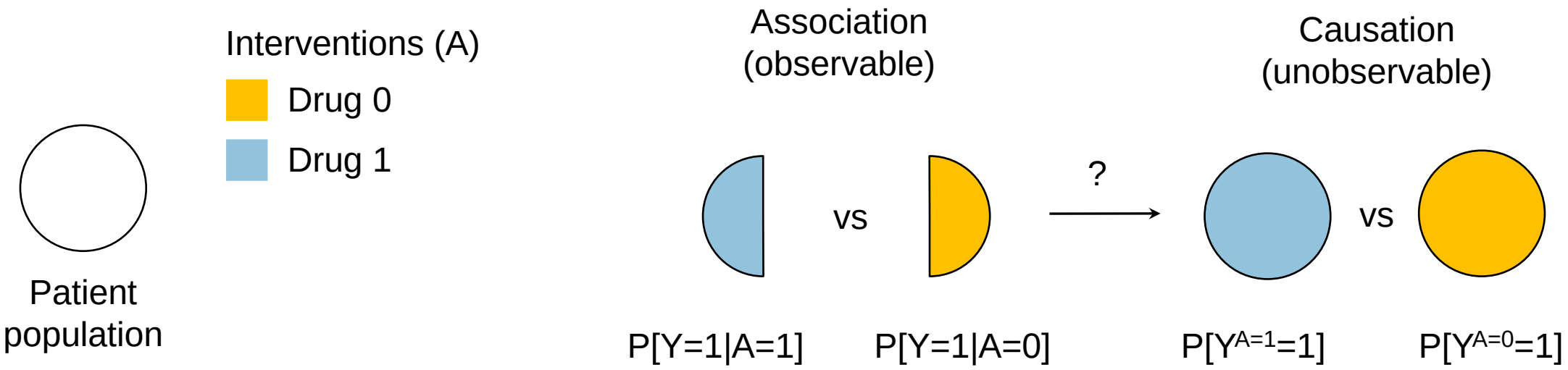
Note

- We will be using causal language
- We are ignoring collider bias, selection bias, measurement error and sample size concerns

Cytel

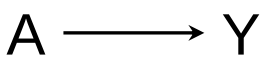
Causal framework

- Under what conditions can we infer causation from correlation?

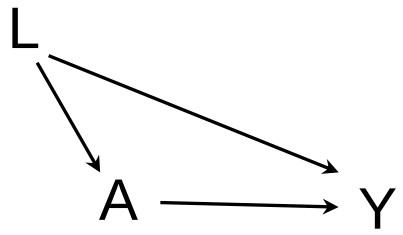


Confounding

- We are interested in the causal effect A→Y of treatment A on outcome Y
- Want observed association measure (e.g. risk ratio) to equal the causal effect
  - $P[Y^{A=1}=1] / P[Y^{A=0}=1] = P[Y=1|A=1] / P[Y=1|A=0]$
- This can not be guaranteed unless treatment A is randomized to achieve exchangeability of treatment groups (e.g. in an RCT*)
- Otherwise, prognostic factors may be imbalanced between treatment groups, and observed association is confounded and not equal to causation



Correlation(A,Y) = causation(A,Y)
E.g. RCT



Correlation(A,Y) ≠ causation(A,Y)
E.g. Non-randomized study designs

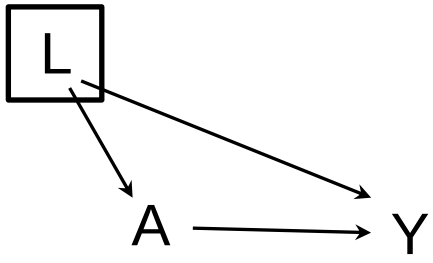
*Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, *210*, 2-21.

Cytel

Adjusting for confounders

Various methods:

- Propensity weighting/IPTW estimation, standardization
- Matching, regression adjustment
- Doubly robust estimators

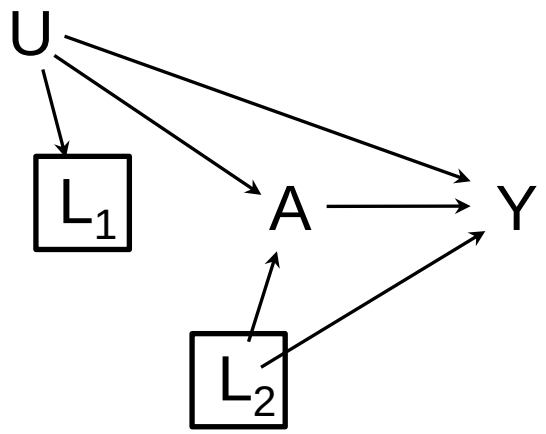


Correlation(A,Y) = causation(A,Y) after adjusting for L

Cytel

Unmeasured confounders

- To eliminate confounding, we need to block all backdoor paths
- Adjustment methods can generally reduce confounding but not necessarily eliminate it
- Researchers will usually assume that there is no unmeasured confounding
  - Difficult or impossible to verify
- Some confounders may be hard to measure or unmeasurable (e.g. socioeconomic status)

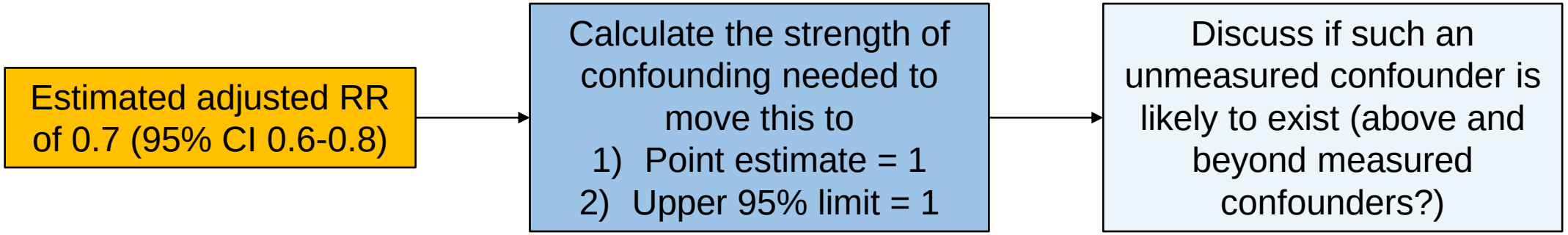


Correlation(A,Y) ≠ causation(A,Y)
after adjusting for measured confounders
L_1, L_2

Cytel

What can be done?

- We generally don't know the true value of the parameter (i.e. the causal effect) we are estimating or features of unmeasured confounders
- We can instead estimate the amount of residual confounding needed to nullify or reverse our conclusions, and then argue whether such bias is plausible

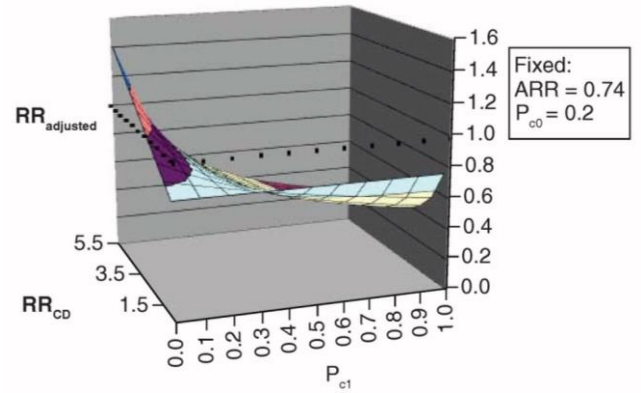


Cytel

Bias plots and E-value

- To do this, we can compute fully adjusted effect estimates (ARR) for estimated RR given a hypothetical confounder with:
  - Prevalence P in A=1
  - Prevalence P in A=0
  - Association with outcome RR_{LY}
- E-value is a scalar summary of the bias plots where $RR_{LY} = RR_{LA}$
  - Similar connotation as a p-value except for unmeasured confounding

$$ARR = RR \times \frac{P_{A=1}(RR_{LY} - 1) + 1}{P_{A=0}(RR_{LY} - 1) + 1}$$

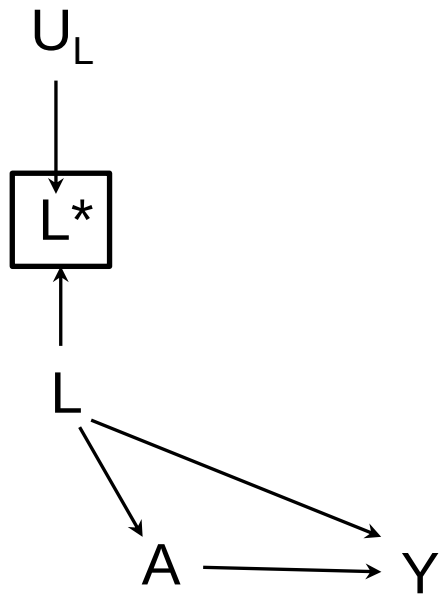


Sammon, Cormac J., et al. "Real-world evidence and nonrandomized data in health technology assessment: using existing methods to address unmeasured confounding?." (2020): 969-972.

Cytel

Partially missing data for confounders

- Real-world data often contains missing values

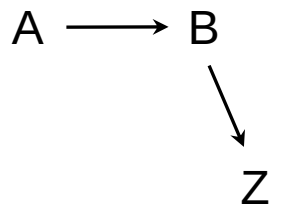


Correlation(A,Y) ≠ causation(A,Y)
after adjusting for measured L*

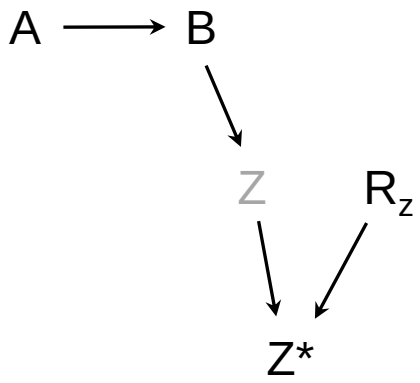
Cytel

Missing data mechanisms

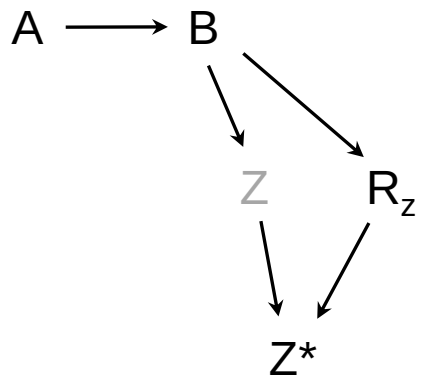
- A mechanism has to be assumed, difficult or impossible to verify
  - Due to limited sample sizes, random variability between treatment groups may be important
- We can estimate treatment effects under a variety of assumptions for missing data



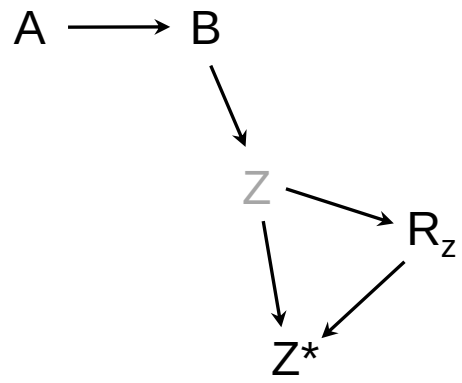
No missingness



Missing completely at random (MCAR)



Missing at random (MAR)

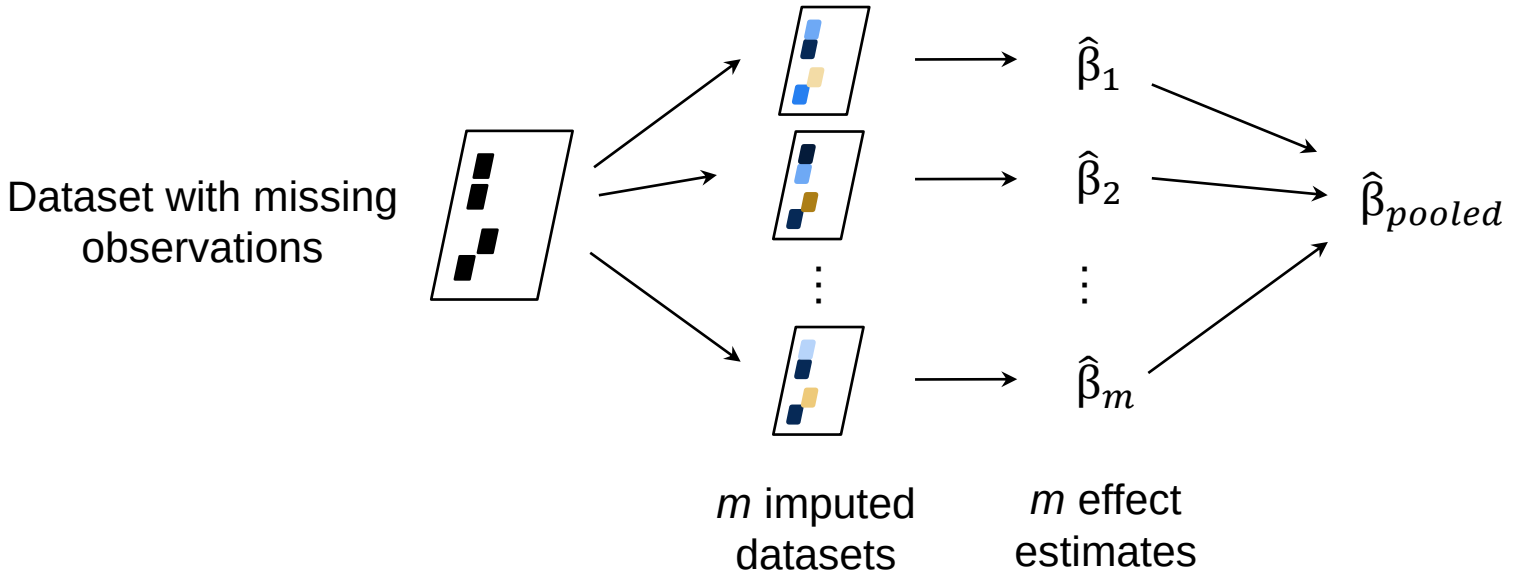


Missing not at random (MNAR)

Cytel

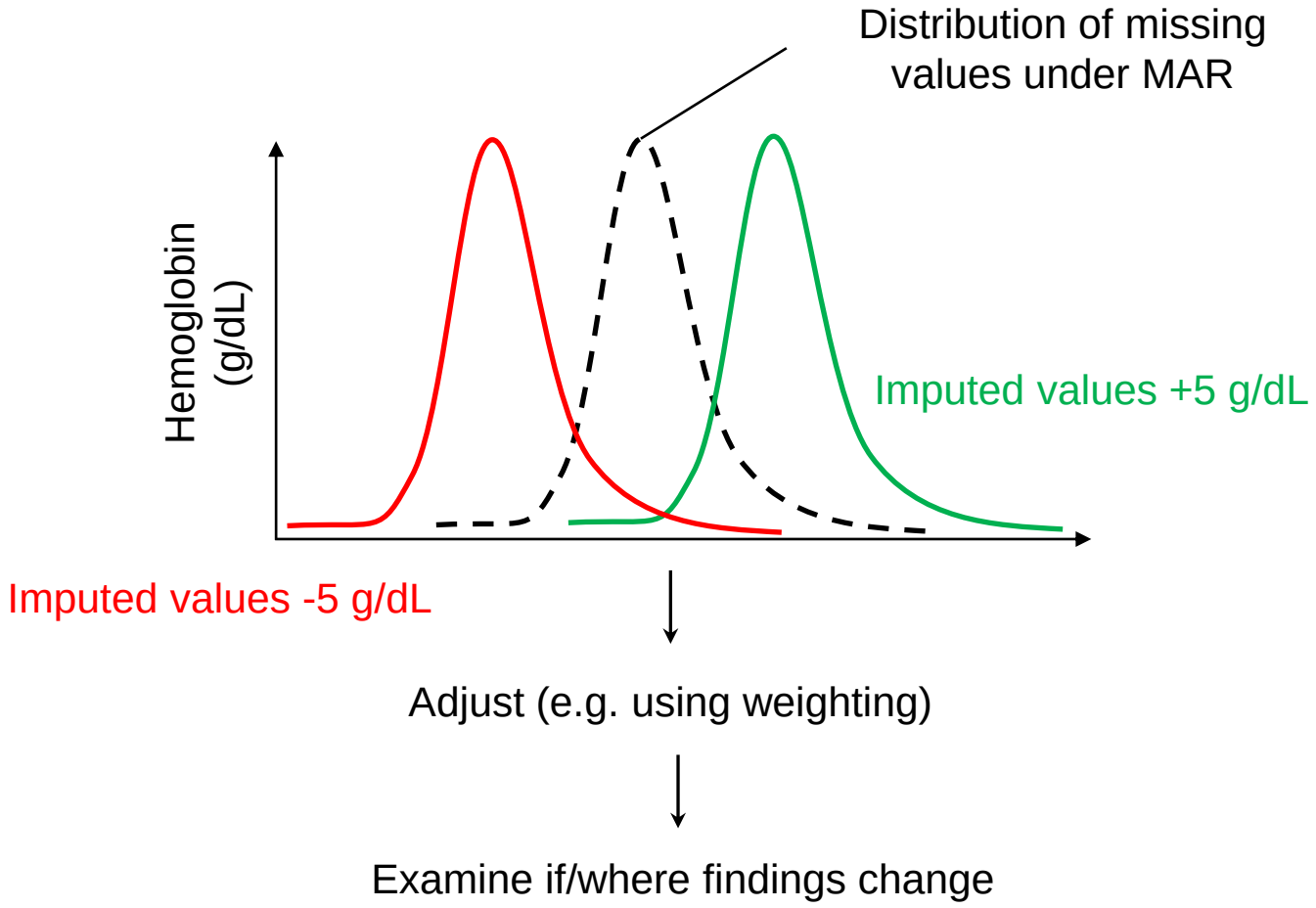
Multiple imputation

- Distribution of missing data is similar to distributed of observed data
- MICE – Fully conditional specification of the joint distribution (a Gibbs sampler)
  - E.g. $\text{logit}\{P[Z|\boldsymbol{X}]\} = \gamma_0 + \boldsymbol{\gamma}^T\boldsymbol{X}$ for binary variable Z with missing values
- Impute multiple (usually >5) datasets with stochastic sampling based on prediction error
- Run analysis on each imputed dataset separately and pool results using Rubin's rules



δ-adjustment for MNAR

- 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)

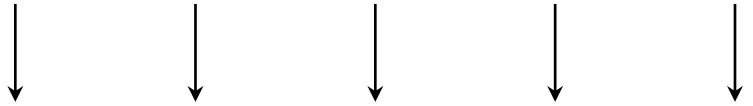
Impute Z with multiple imputation under MAR

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

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

Cytel

Conclusion

- To assess robustness to unmeasured confounding, you can
  - Calculate bias plots
  - Calculate E-value
- To assess robustness to missing data for confounders, you can
  - Compute effect estimates under MCAR, MAR
  - Identify tipping points under MNAR using δ-adjustment or NARFCS
- Then, discuss plausibility of these assumptions under which conclusions change
- If these assumptions are plausible, further sensitivity analyses may be needed using e.g. external data [Zhang, Xiang, James D. Stamey, and Maya B. Mathur. "Assessing the impact of unmeasured confounders for credible and reliable real-world evidence." *Pharmacoepidemiology and Drug Safety* 29.10 (2020): 1219-1227.]

Cytel

References

Causal inference and adjustment methods

Hernán, M. A., & Robins, J. M. (2020). Causal inference: what if. (Online pdf)

Funk, Michele Jonsson, et al. "Doubly robust estimation of causal effects." *American journal of epidemiology* 173.7 (2011): 761-767.

Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in medicine*, *34*(28), 3661-3679.

Unmeasured confounding adjustment, E-value

Schneeweiss, Sebastian. "Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics." *Pharmacoepidemiology and drug safety* 15.5 (2006): 291-303.

VanderWeele, Tyler J., and Peng Ding. "Sensitivity analysis in observational research: introducing the E-value." *Annals of internal medicine* 167.4 (2017): 268-274.

R packages: episensr, EValue

Missing data, MICE, delta adjustment

Van Buuren, Stef, Hendriek C. Boshuizen, and Dick L. Knook. "Multiple imputation of missing blood pressure covariates in survival analysis." *Statistics in medicine* 18.6 (1999): 681-694.

Van Buuren, Stef, Hendriek C. Boshuizen, and Dick L. Knook. "Multiple imputation of missing blood pressure covariates in survival analysis." *Statistics in medicine* 18.6 (1999): 681-694.

Buuren, S. van, and Karin Groothuis-Oudshoorn. "mice: Multivariate imputation by chained equations in R." *Journal of statistical software* (2010): 1-68.

VanderWeele, T. (2017). On a square-root transformation of the odds ratio for a common outcome. *Epidemiology (Cambridge, Mass.)*, *28*(6), e58.

Van Buuren, S. (2018). *Flexible imputation of missing data*. CRC press.

Cytel

Quantitative Bias Analysis

Application to comparative effectiveness research for advanced cancer

Disclosures

- Employed by and holds stock in Roche
- Roche Products Ltd, Welwyn Garden City, UK
- Opinions expressed are my own and not necessarily those of Roche

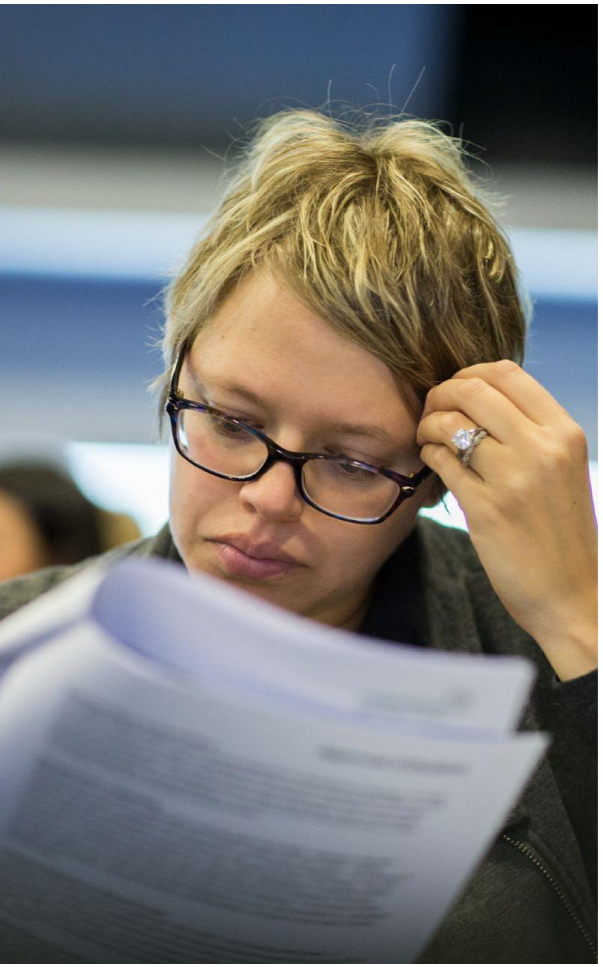
Quantitative Bias Analysis (QBA) for Comparative Effectiveness of Alectinib versus Ceritinib in Non-Small Cell Lung Cancer (NSCLC)

Samantha Wilkinson, Joshua Ray, Sreeram Ramagopalan ***Roche***
Alind Gupta, Eric Mackay, Paul Arora, Kristian Thorlund, Radek Wasiak ***Cytel***

Overview

1. **Background and context: Challenges of using RWD for reimbursement decisions**
2. **Methods, Results**
3. **Quantitative Bias Analysis (QBA) to explore uncertainty**

Comparative effectiveness of alectinib versus ceritinib in ALK+ non-small cell lung cancer (NSCLC)



Real-world data (RWD) submitted as a comparator to a single arm trial for HTA

Alectinib lower risk of death: hazard ratio: 0.65; 95% CI: 0.48-0.88

Payers cautious to assign ‘additional benefit’ to alectinib

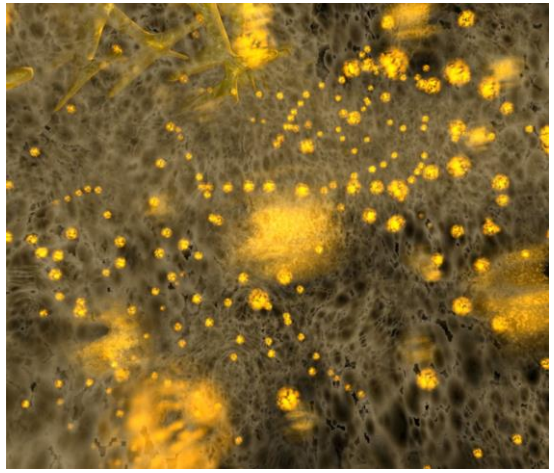
Davies, J., et al., *Comparative effectiveness from a single-arm trial and real-world data: alectinib versus ceritinib.* Journal of comparative effectiveness research, 2018. **7**(09): p. 855-865.

QBA: Two investigations to address unmeasured confounding and missing data

Unmeasured confounding: How extreme would an unmeasured confounder need to be to nullify effect estimate?

Missing data: How biased would missing values need to be to nullify effect?

Study set up



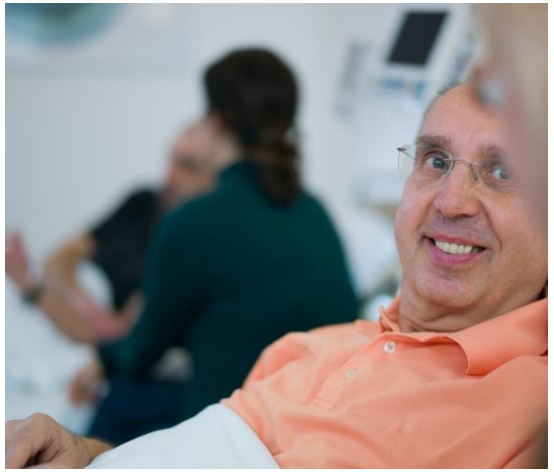
Population

Two comparisons

RCT vs. RWD
RWD vs. RWD

RCT data:
Alectinib (NCT01801111 & NCT01871805)

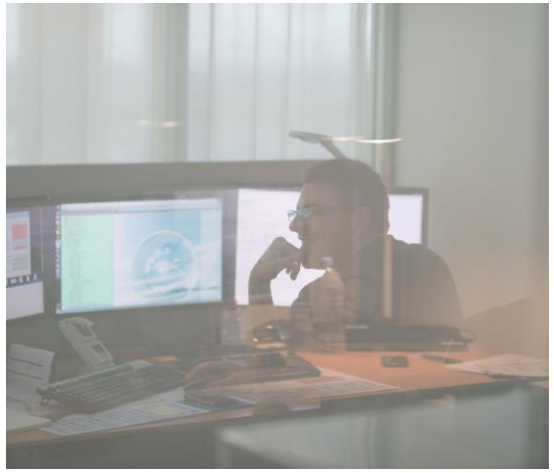
RWD:
Flatiron Health alectinib and ceritinib



Comparisons

Treatment

Alectinib vs. ceritinib after crizotinib failure in patients with ALK+ advanced non-small cell lung cancer



Outcomes

Overall survival

Adjusted using inverse probability of treatment weight (IPTW)

QBA

Sensitivity analysis

Explore tipping points for unmeasured confounding and missing data in performance score

Primary Analysis

Results

RCT vs. RWD
alectinib vs. ceritinib

	HR (95% CI)
Complete case	0.55 (0.31-0.95)
Multiple imputation	**0.59** (0.44-0.75)

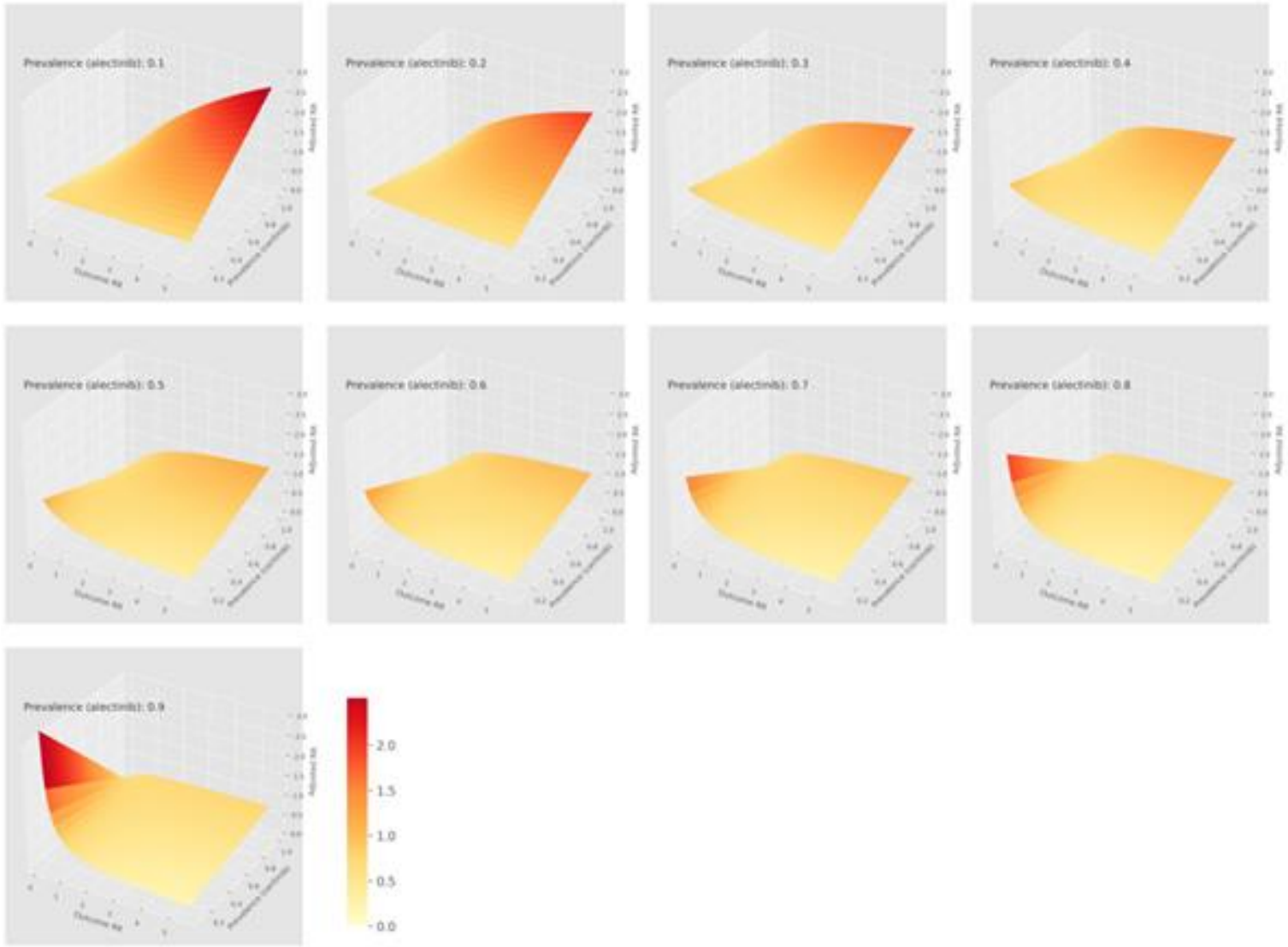
RWD vs. RWD
alectinib vs. ceritinib

	HR (95% CI)
Complete case	0.47 (0.23-0.96)
Multiple imputation	0.46 (0.29-0.63)

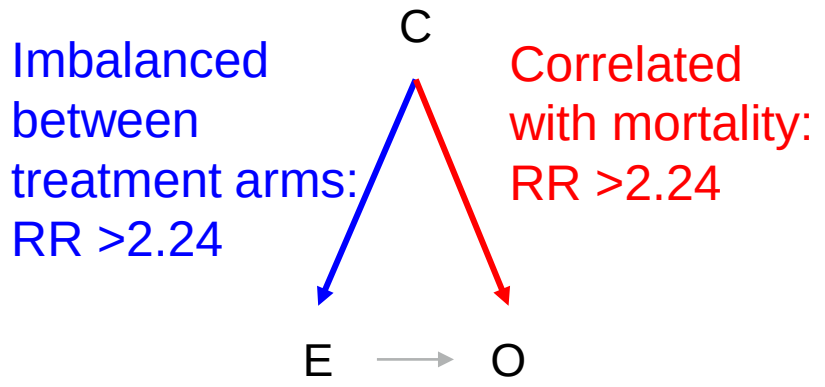
Previous study: Alectinib lower risk of death: hazard ratio: 0.65; 95% CI: 0.48-0.88

Modelling a hypothetical unmeasured confounder

Each graph shows a different prevalence of the unmeasured confounder in alectinib pts

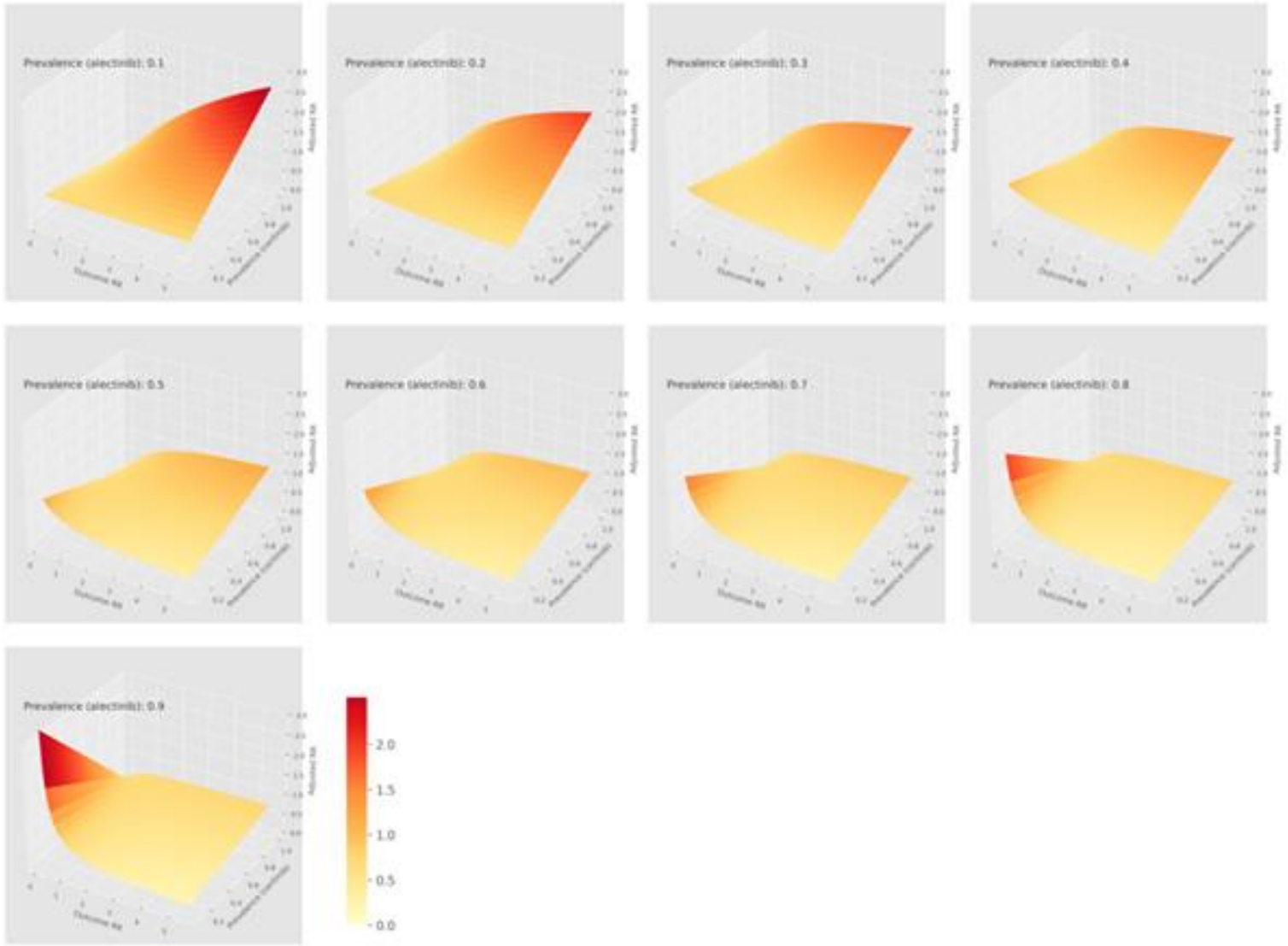


Move HR >1

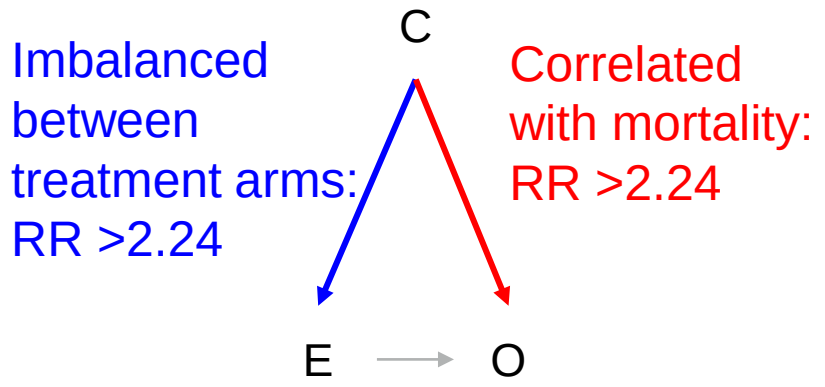


Modelling a hypothetical unmeasured confounder

Each graph shows a different prevalence of the unmeasured confounder in alectinib pts



Move HR >1



Prior chemotherapy: Association with treatment RR: 1.73, mortality RR: ~1.10.

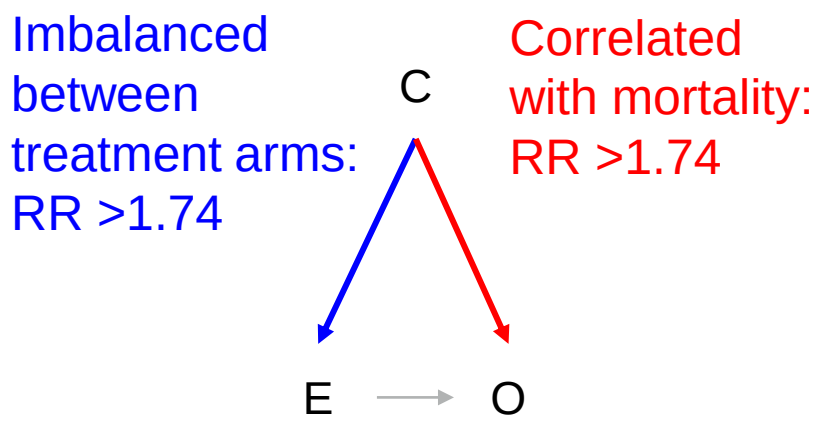
ECOG PS: Association with treatment RR: 1.19; mortality RR: <2.1.

Modelling a hypothetical unmeasured confounder

RCT vs. RWD
alectinib vs. ceritinib

Multiple imputation HR (95% CI): 0.59 (0.44-0.75)

Upper CI to cross null



Prior chemotherapy: Association with treatment RR: 1.73, mortality RR: ~1.10.

ECOG PS: Association with treatment RR: 1.19; mortality RR: <2.1.

QBA: Two investigations to address unmeasured confounding and missing data

Unmeasured confounding: How extreme would an unmeasured confounder need to be to nullify effect estimate?

Missing data: How biased would missing values need to be to nullify effect?

Balancing covariates is crucial: missing data

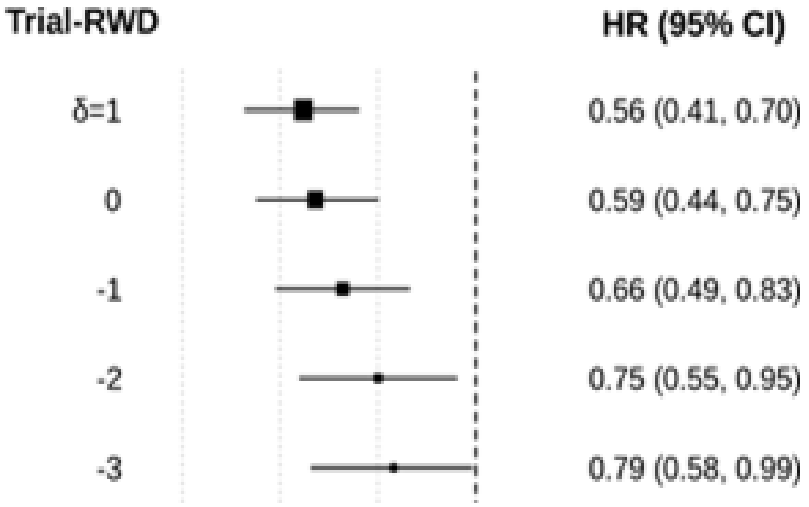
Baseline covariates

	Ceritinib RWD	Alectinib trial	Alectinib RWD
Sample size (unadjusted)	91	183	81
ECOG PS (%)			
0	18 (37.5)	64 (35.0)	21 (39.6)
1	19 (39.6)	101 (55.2)	28 (52.8)
2	11 (22.9)	18 (9.8)	4 (7.5)
Missing	**43 (47.3)**	**0 (0)**	**28 (34.6)**

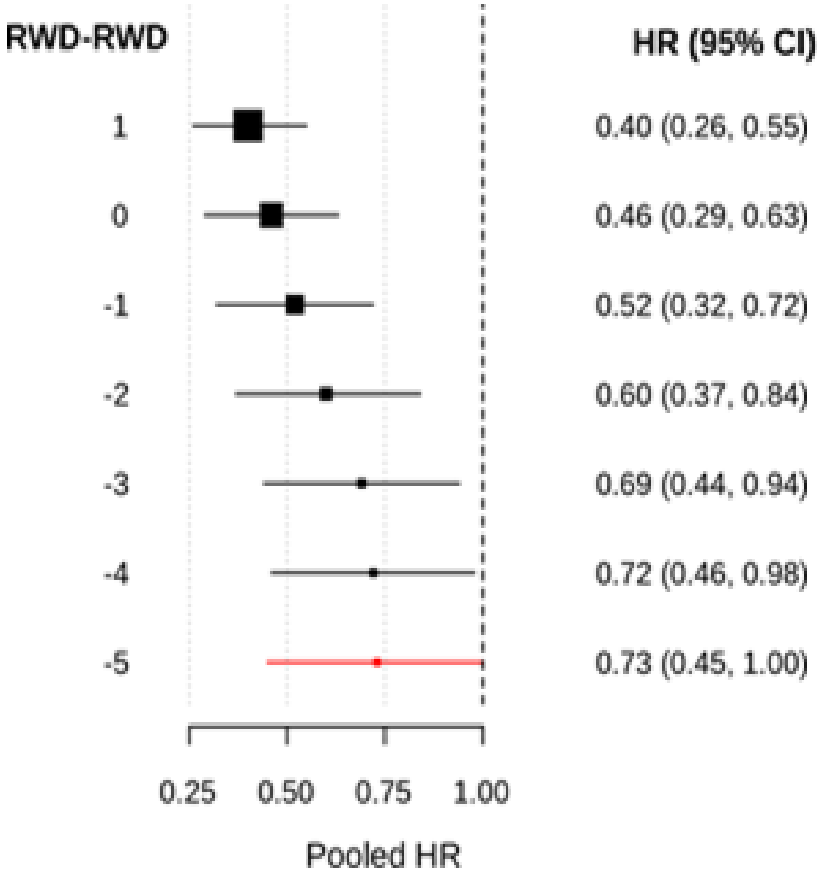
Modelling deviations from MAR: missing data

Results

RCT vs. RWD
alectinib vs. ceritinib



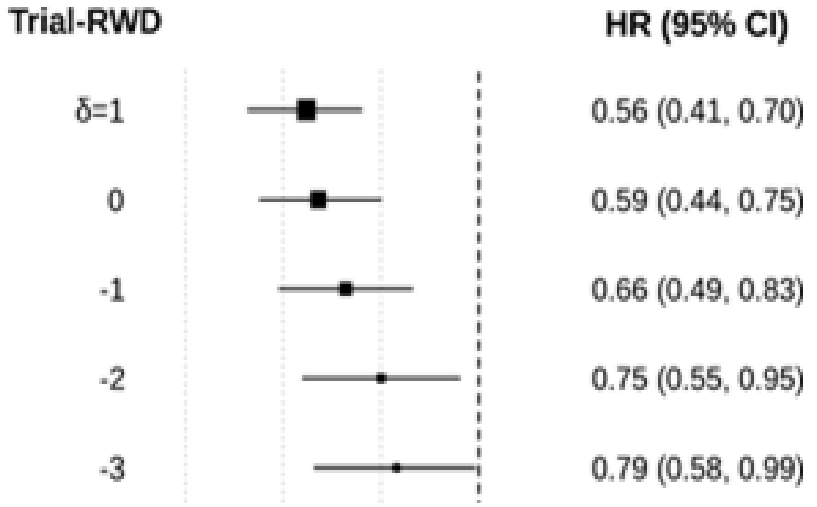
RWD vs. RWD
alectinib vs. ceritinib



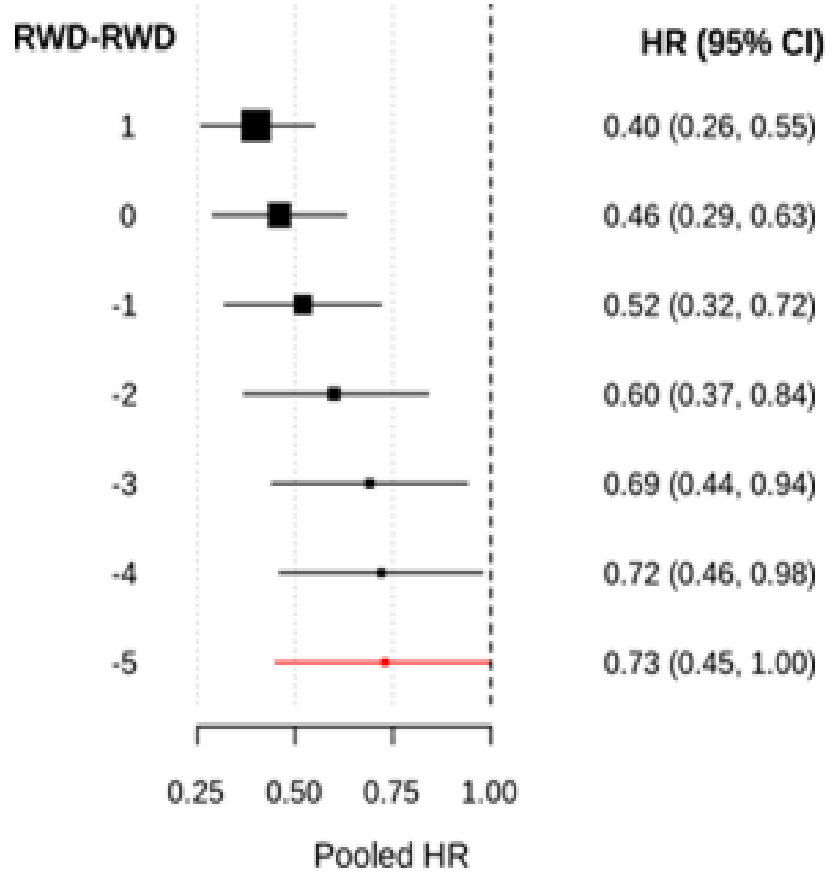
Modelling deviations from MAR: missing data

Results

RCT vs. RWD
alectinib vs. ceritinib



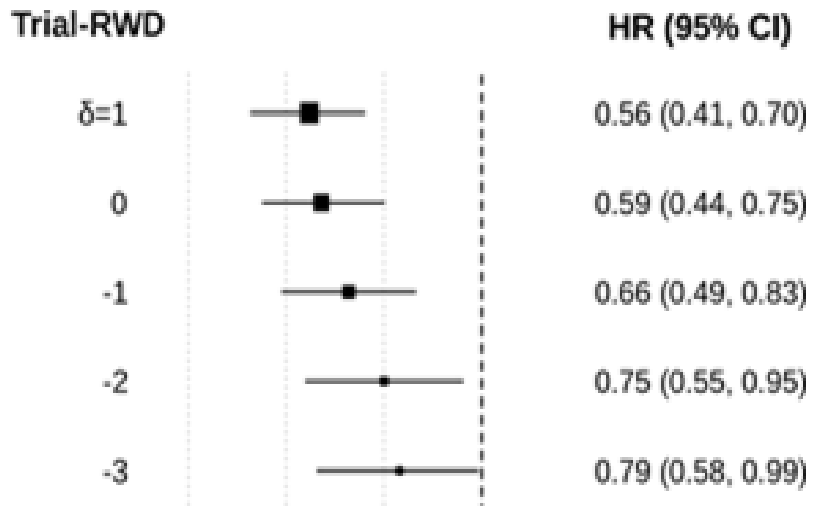
RWD vs. RWD
alectinib vs. ceritinib



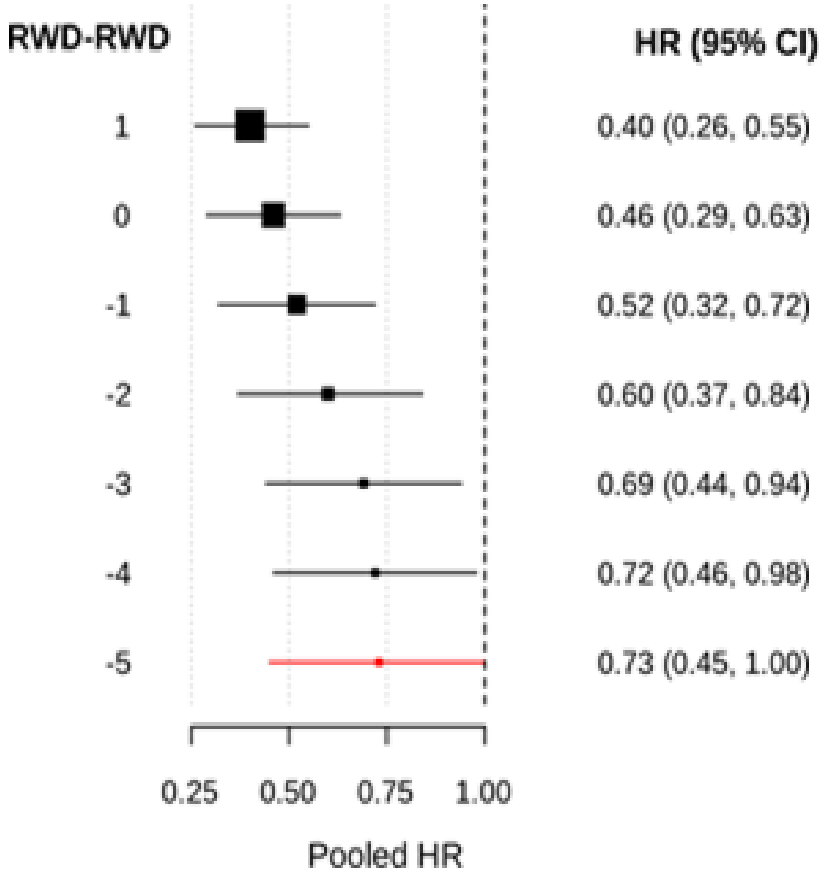
Modelling deviations from MAR: missing data

Results

RCT vs. RWD
alectinib vs. ceritinib

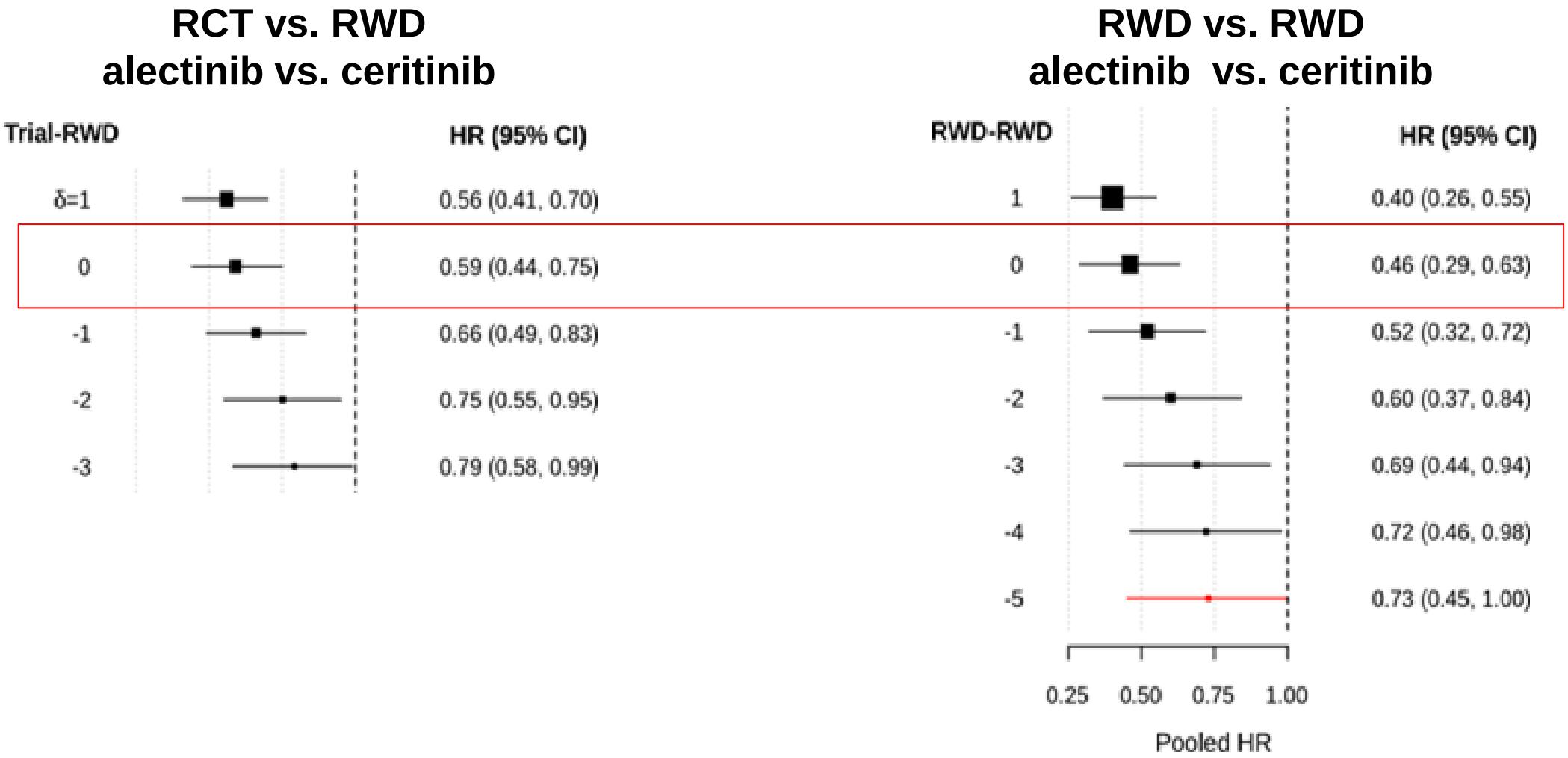


RWD vs. RWD
alectinib vs. ceritinib



Modelling deviations from MAR: missing data

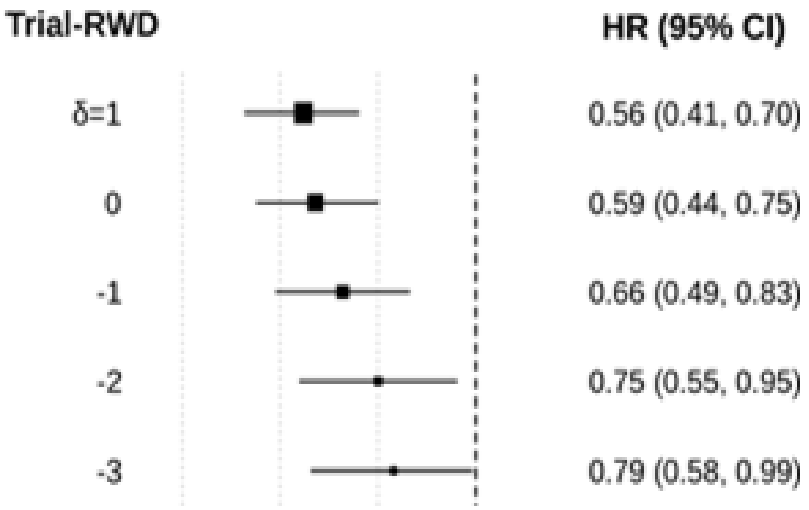
Results



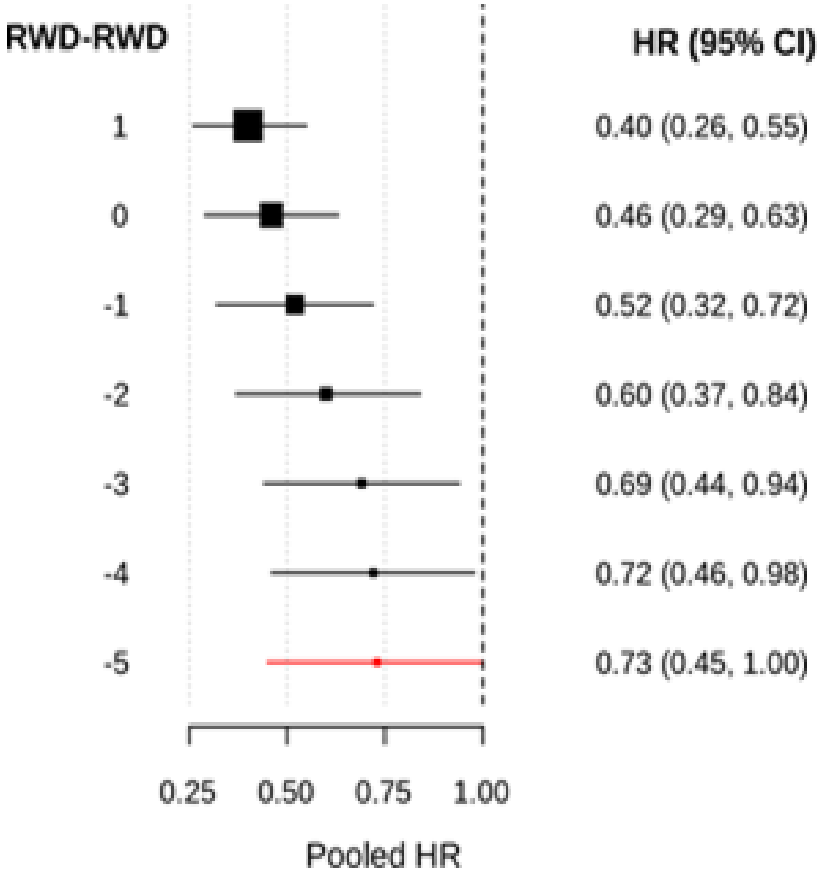
Modelling deviations from MAR: missing data

Results

RCT vs. RWD
alectinib vs. ceritinib



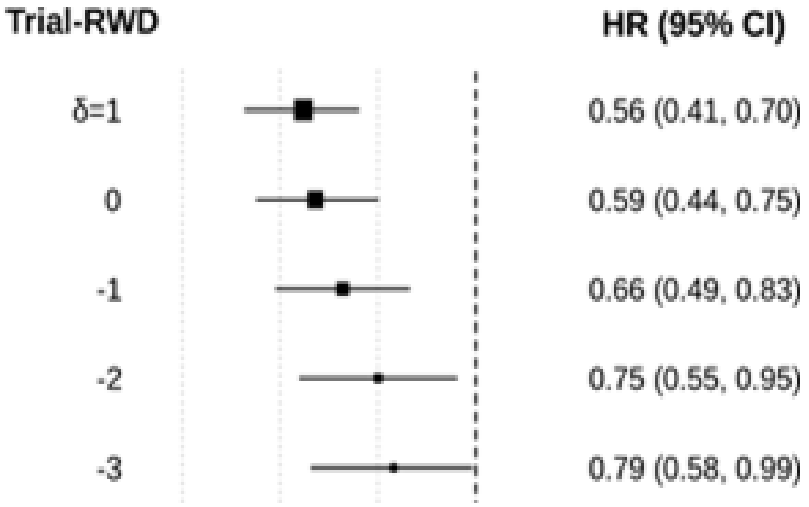
RWD vs. RWD
alectinib vs. ceritinib



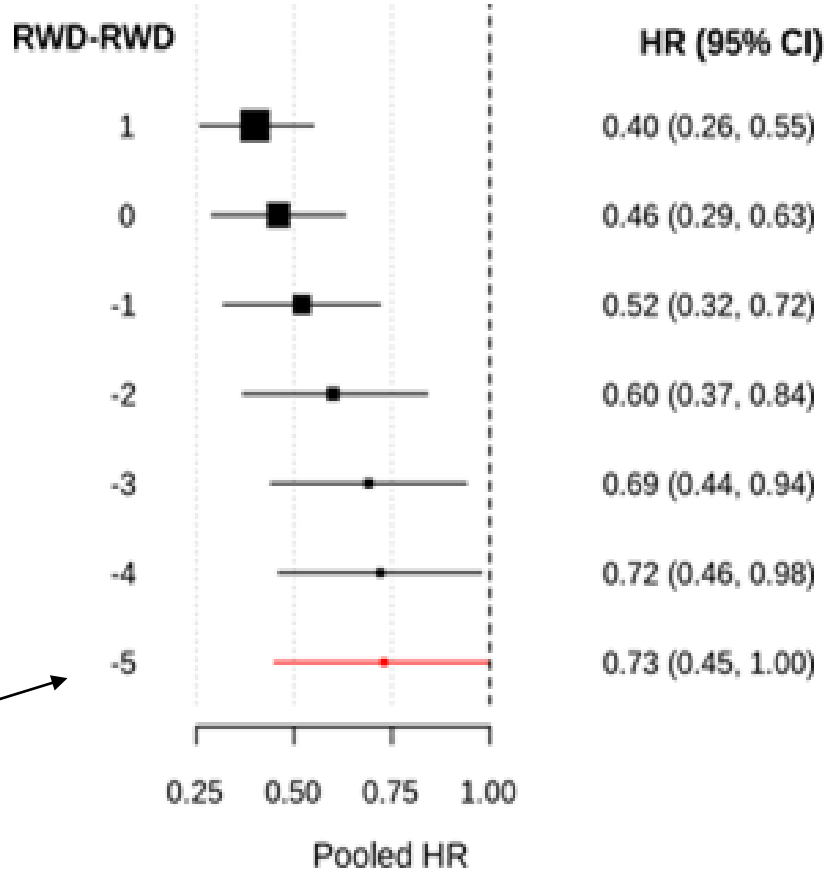
Modelling deviations from MAR: missing data

Results

RCT vs. RWD
alectinib vs. ceritinib



RWD vs. RWD
alectinib vs. ceritinib



Average shift in ECOG score of
+1.63 in ceritinib group

Limitations

Other sources of unmeasured confounding may be present

Other baseline covariates were missing: We only investigated MNAR models for ECOG

Differential measurement of outcomes and covariates in trials vs. RWD

Webster-Clark, Michael PharmD, PhD; Jonsson Funk, Michele PhD; Stürmer, Til MD, PhD Single-arm Trials With External Comparators and Confounder Misclassification, Medical Care: December 2020 - Volume 58 - Issue 12 - p 1116-1121

Summary

RWD can provide data for studies of treatment effectiveness: **Bias remains a major concern for decision-makers**

Alectinib was associated with longer OS compared to ceritinib after crizotinib failure in ALK+ patients with NSCLC

QBA: Results were robust to plausible sources of bias from unmeasured confounding and missing data

QBA should accompany submissions of non-randomized studies to aid regulatory and reimbursement decision making

References

Davies, J., et al., *Comparative effectiveness from a single-arm trial and real-world data: alectinib versus ceritinib.* Journal of comparative effectiveness research, 2018. **7**(09): p. 855-865.

VanderWeele, T.J. and O.A. Arah, *Unmeasured confounding for general outcomes, treatments, and confounders: bias formulas for sensitivity analysis.* Epidemiology (Cambridge, Mass.), 2011. **22**(1): p. 42.

Schneeweiss, S., *Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics.* Pharmacoepidemiology and drug safety, 2006. **15**(5): p. 291-303.

Little, R.J. and D.B. Rubin, *Statistical analysis with missing data*. Vol. 793. 2019: John Wiley & Sons.

Sammon, C.J., et al., *Real-world evidence and nonrandomized data in health technology assessment: using existing methods to address unmeasured confounding?* 2020, Future Medicine.

Leacy, F.P., et al., *Analyses of sensitivity to the missing-at-random assumption using multiple imputation with delta adjustment: application to a tuberculosis/HIV prevalence survey with incomplete HIV-status data.* American journal of epidemiology, 2017. **185**(4): p. 304-315.

Webster-Clark, Michael PharmD, PhD; Jonsson Funk, Michele PhD; Stürmer, Til MD, PhD Single-arm Trials With External Comparators and Confounder Misclassification, Medical Care: December 2020 - Volume 58 - Issue 12 - p 1116-1121

Doing now what patients need next

Q&A