EFSPI EUROPEAN FEDERATION OF STATISTICIANS IN THE PHARMACEUTICAL INDUSTRY
Representing Statistical Associations in Europe

Unanchored indirect treatment comparison methods and unmeasured confounding

Kate Ren, PhD
Senior Research Fellow
NIHR Advance Research Fellow
University of Sheffield

NICE Technology Appraisal Committee Member

University of Sheffield

NIHR | National Institute for Health and Care Research

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Disclaimers

- The findings and views expressed in this presentation are those of the presenter, who is responsible for its contents.
- The findings and views expressed should not be understood or quoted as being made on behalf of
  - NICE technology committee
  - National Institute for Health and Care Research (NIHR)

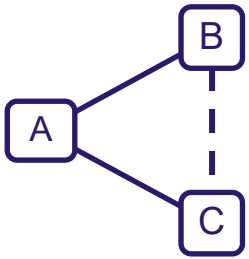
Overview

- Population-adjusted indirect comparisons (PAICs)
  - Unanchored MAIC/STC
- Unmeasured confounding
  - Quantitative bias analysis (QBA)
- Case study
  - Metastatic colorectal cancer

Population-adjusted indirect comparisons (PAICs)

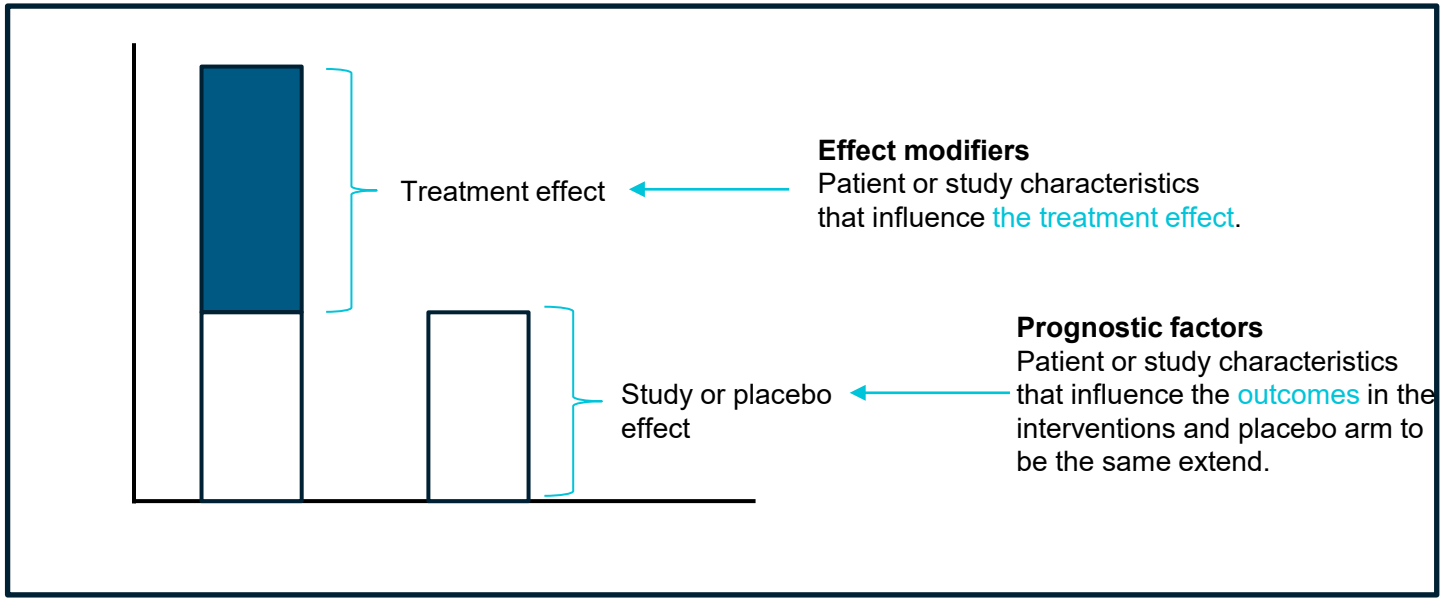
Evidence synthesis

- Evidence from multiple sources
  - **Meta-analysis**: pool evidence from independent sources
  - **Pairwise meta-analysis**: two treatments
  - **Network meta-analysis (NMA)**: more than two treatments
  - **Indirect treatment comparison (ITC)**: no head-to-head trials
    - Anchored



    - Unanchored

Effect modifiers vs. Prognostic factors

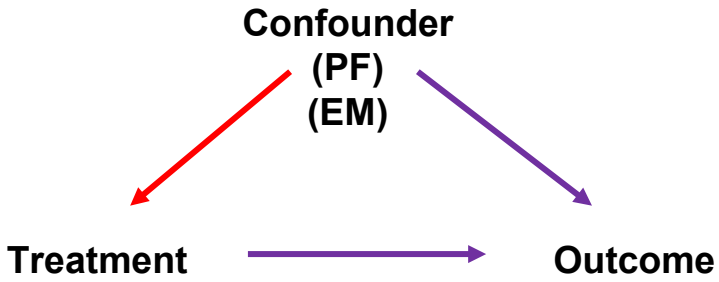


Reproduced from: Jansen JP, Trikalinos T, Cappelleri JC, Daw J, Andes S, Eldessouki R, Salanti G. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. Value Health. 2014 Mar;17(2):157-73.

Confounder and confounding issues

Confounder

- Associate with exposure and outcome
- But not an intermediate pathway



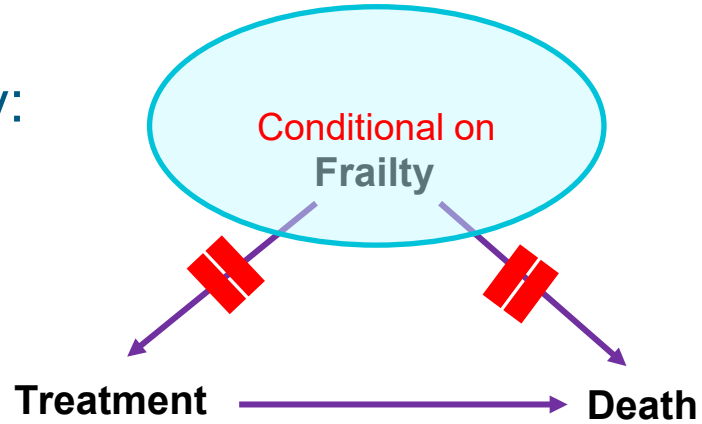
Confounding issues

- Non-RCTs
  - Treated and untreated individuals are likely to be different (in many ways)
- Estimate treatment effect in non-RCTs
  - Make fair comparisons between treated and untreated individuals
  - Measure the difference and account for them in estimating treatment effect

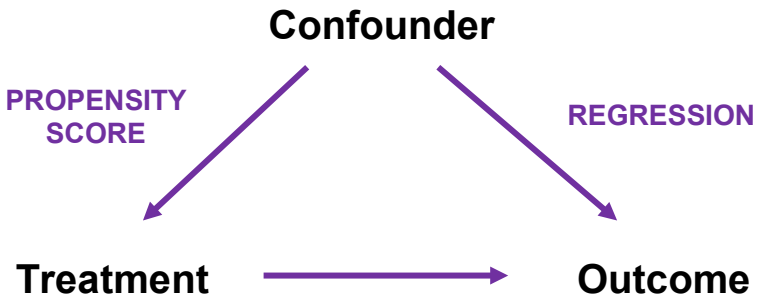
RCTs vs. Observational studies

- RCT:

- Observational study:



Propensity scores vs. Regression



- Propensity score: models the **treatment allocation mechanism**
- Regression: models the **outcome mechanism**

Poll

Have you ever been involved in a study that utilised population-adjusted indirect comparisons?

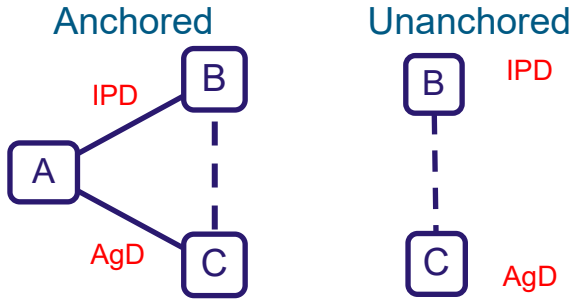
- Yes, MAIC
- Yes, STC
- Yes, ML-NMR
- Yes, other methods
- No involvement

(multiple choice)

Population-adjusted indirect comparisons (PAICs)

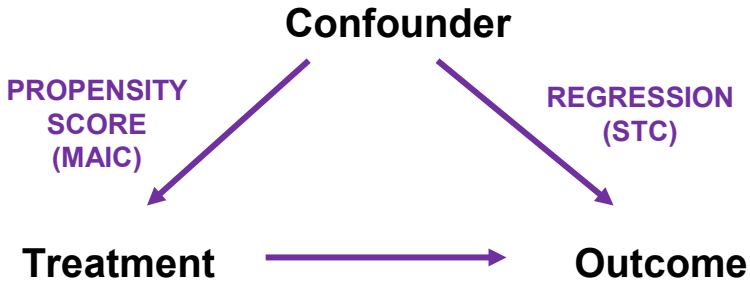
The problem

- No head-to-head trials
- Company: Individual-patient level data (IPD)
- Comparator: Aggregate data (AgD)



Adjust for between-study difference in baseline characteristics

- Matching-Adjusted Indirect Compassion (**MAIC**)
  - Population reweighting
- Simulated Treatment Comparison (**STC**)
  - Outcome regression model



Population adjustment methods assumptions

To estimate a causal effect, typically make four key assumptions:

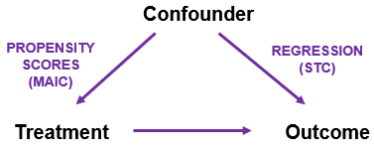
1. **Positivity** (experimental treatment assignment)
2. **Consistency** (homogeneity of effects)
3. **No interference** (Stable Unit Treatment Value Assumption [SUTVA])
4. **No unmeasured confounding** (Exchangeability, Strongly Ignorable Treatment Assignment [SITA])

PAICs assumptions

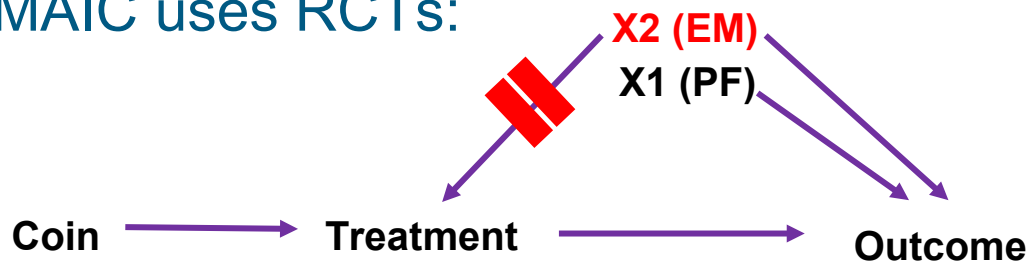
To estimate a causal effect, typically make four key assumptions:

1. **Positivity** (experimental treatment assignment)
2. **Consistency** (homogeneity of effects)
3. **No interference** (Stable Unit Treatment Value Assumption [SUTVA])
4. **No unmeasured confounding** (Exchangeability, Strongly Ignorable Treatment Assignment [SITA])
   - Conditional constancy of relative effects (**anchored** ITC)
     - Adjust for all effect modifiers
   - Conditional constancy of absolute effects (**unanchored** ITC)
     - Adjust for all effect modifiers and prognostic variables

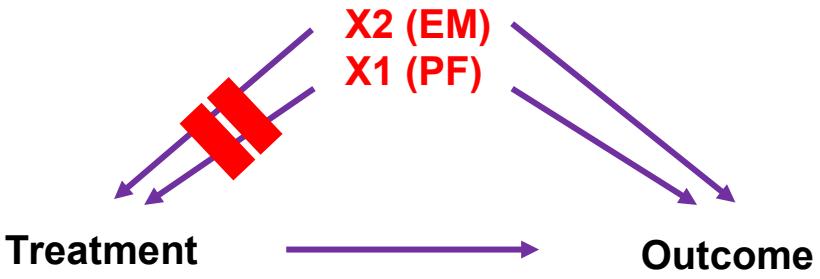
Anchored and unanchored **MAIC**



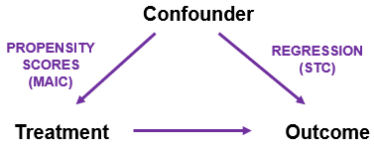
- Anchored MAIC uses RCTs:



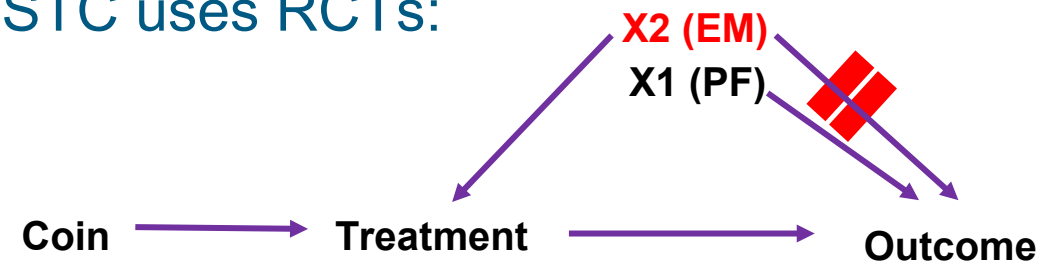
- Unanchored MAIC lacks protection from randomisation:



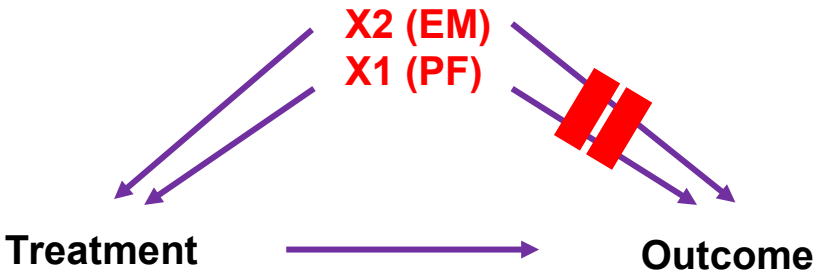
Anchored and unanchored **STC**



- Anchored STC uses RCTs:

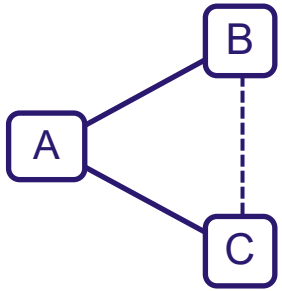


- Unanchored STC lacks protection from randomisation:



Covariates selection

- Anchored MAIC/STC



All effect modifiers
- known
- adjusted for

- Unanchored MAIC/STC (no common comparator)

All effect modifiers and prognostic variables
- known
- adjusted for

Review of PAICs

- 162 eligible records (2010-2023)*
  - Oncology: 94 (58.0%)
- Type of outcome
  - Continuous: 20 (12.4%)
  - Binary: 76 (46.9%)
  - Time-to-event: 66 (40.7%)
- Population adjustment methods
  - MAIC: 144 (88.9%)
  - STC:11 (6.8%)
  - Both MAIC and STC: 6 (3.7%)
  - ML-NMR: 1 (0.6%)
- Type of comparison
  - Anchored: 57 (35.2)
  - Unanchored: 105 (64.8%)

*Truong et al. (2023) doi:10.1002/jrsm.1653

Poll

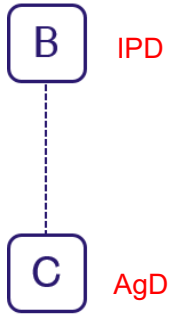
Why do you think MAIC is more frequently used than STC?

- MAIC is more intuitive to understand.
- Lack of guidance on how to perform STC.
- Everybody uses MAIC. Let's use it in our submission.

Unanchored MAIC (1)

Propensity score reweighting approach*

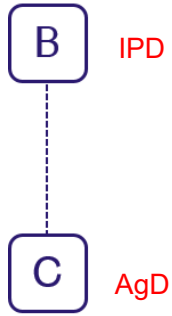
1. Create a logistic propensity score model, including all effect modifiers and prognostic factors
   - $\log(\text{weight}_i) = a_0 + a_1 x_{i1} + a_2 x_{i2} + \cdots$
     - weight_i: "trial selection" odds
   - Estimate weights using the method of moments
     - Set the weights so that the mean (potentially, higher moments: e.g. variance) of the covariates are exactly balanced across the two trial populations
       - Achieved using optimisation



* NICE DSU TSD 18

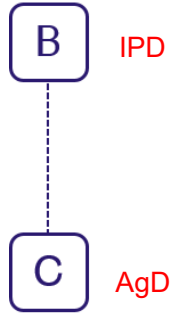
Unanchored MAIC (2)

2. Predict outcomes on treatment B in Study C population by reweighting the outcomes of the B individuals
   - $\hat{Y}_{B(C)} = \frac{\sum_{i=1}^{N_{B(B)}} Y_{i(B)} \hat{w}_i}{\sum_{i=1}^{N_{B(B)}} \hat{w}_i}$
3. Obtain the unanchored indirect comparison in Study C population
   - $\hat{d}_{BC(C)} = \hat{d}_{C(C)} - \hat{d}_{B(C)} = g(\bar{Y}_{C(C)}) - g(\hat{Y}_{B(C)})$
4. Calculate standard error
   - Robust sandwich estimator
   - Bootstrapping
   - Bayesian techniques

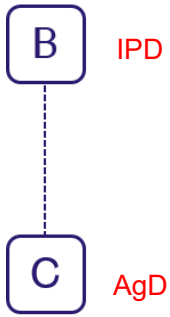


Unanchored MAIC (3)

5. Assess bias
   - NICE DSU TSD18: "*Provide evidence that absolute outcomes can be predicted with sufficient accuracy in relation to the relative treatment effects, and present an estimate of the likely range of residual systematic error. If this evidence cannot be provided or is limited, then state that the amount of bias in the indirect comparison is likely to be substantial, and could even exceed the magnitude of treatment effects which are being estimated.*"
6. Target population
   - Use the shared effect modifier assumption to transport the ITC estimate into target population if justified
   - Comment on the representativeness of Study C population
7. Present the distribution of estimated weights and ESS



Unanchored STC (1)



Outcome regression/parametric model-based approach

1. Build regression model based on the IPD from Study B, including all effect modifiers and prognostic factors

$$g\left(\theta_{i(B)}\right) = \beta_0 + \boldsymbol{\beta_1^T X_i}$$

2. Predict the treatment effect for Study C population

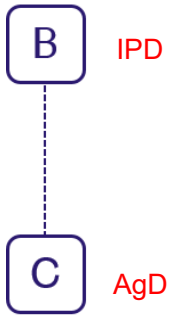
$$\hat{d}_{B(C)} = g\left(\hat{\theta}_{B(C)}\right)$$

3. Obtain the unanchored indirect comparison in Study C population, using the prediction from Step 2 and reported aggregate data for Study C

$$\hat{d}_{BC(C)} = \hat{d}_{C(C)} - \hat{d}_{B(C)} = g\left(\bar{\theta}_{C(C)}\right) - g\left(\hat{\theta}_{B(C)}\right)$$

Unanchored STC (2)

4. Calculate standard error
5. Assess bias (same as MAIC)
6. Target population (same as MAIC)
7. Present standard model fit statistics

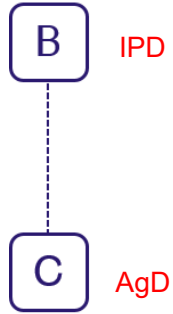


Unanchored STC (3)

- How to predict?
  - Identity link function: "Plugging-in" mean approach

$$\hat{d}_{BC(C)} = \hat{d}_{C(C)} - \hat{d}_{B(C)}$$
$$= \bar{\theta}_{C(C)} - (\hat{\beta}_0 + \hat{\boldsymbol{\beta}}_1^T \overline{\boldsymbol{X}}_{(C)})$$

  - Non-identity link function
    - "Plugging-in" mean approach: aggregation bias
    - Simulate individual-level covariates for Study C
      - NORTA/Gaussian copula
      - Adjusted absolute effect $(\hat{d}_{B(C)})$ is obtained by averaging the predictions of these individuals



NORmal To Anything (NORTA)/ Gaussian copula

To simulate a random vector $\boldsymbol{X} = (X_1, \ldots, X_k)$ with the following properties

- $X_i \sim F_{X_i}$, $i = 1, \ldots, k$ and F_{X_i} is the cumulative distribution function (CDF) for X_i; and
- $Corr(\boldsymbol{X}) = \Sigma_X$,

the NORTA algorithm proceeds as:

$$Z_i \sim MVN(0, \Sigma_Z)$$

$$\begin{pmatrix} Z_1 \\ \vdots \\ Z_i \\ \vdots \\ Z_k \end{pmatrix} \xrightarrow{X_i = F_{X_i}^{-1}(\Phi(Z_i))} \begin{pmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_k \end{pmatrix}$$

Probability integration transformation

Unanchored STC (4)

Example: binary outcome

1. Build regression model based on the IPD from Study B

   $g\left(\theta_{i(B)}\right) = \beta_0 + \boldsymbol{\beta_1^T X_i}$

   $g() = logit()$

2. Predict the treatment effect for Study C population

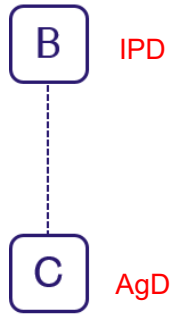
   $\hat{d}_{B(C)} = g\left(\hat{\theta}_{B(C)}\right)$

   $\hat{P}_{B(C)}(Y=1) = \frac{1}{N}\sum_{j=1}^{N} P\left(Y=1|\boldsymbol{X_j}\right)$

3. Obtain the unanchored indirect comparison in Study C population, using the prediction from Step 2 and reported aggregate data for Study C

   $\hat{d}_{BC(C)} = \hat{d}_{C(C)} - \hat{d}_{B(C)}$

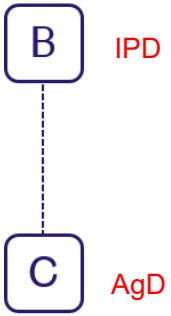
   $= \log\left(\frac{\hat{P}_{C(C)}(Y=1)}{1-\hat{P}_{C(C)}(Y=1)}\right) - \log\left(\frac{\hat{P}_{B(C)}(Y=1)}{1-\hat{P}_{B(C)}(Y=1)}\right)$



Unanchored STC (5)

- The general formula for $\hat{d}_{B(C)}$ is

$$\hat{d}_{B(C)} = g\left(\frac{1}{N}\sum_{j=1}^{N} g^{-1}\left(\hat{\beta}_0 + \widehat{\boldsymbol{\beta}}_1^{\boldsymbol{T}}\boldsymbol{X}_{\boldsymbol{j(C)}}\right)\right)$$



- Standard error of STC estimates

$$Var\left(\hat{d}_{BC(C)}\right) = Var\left(\hat{d}_{B(C)}\right) + Var\left(\hat{d}_{C(C)}\right)$$

Bootstrap

Reported summary statistics

Unmeasured confounding

Single-arm trials in HTA submissions

- Review of HTA submissions (2011-2019)*
- 433 single-arm trials
- 8 in 2011 to 102 in 2019
- 13-fold increase

*Patel et al. (2021) doi:10.1016/j.jval.2021.01.015

Confounding issue

- TA592: "**None** of the indirect comparisons provide a **reliable estimate** of relative effectiveness"
- TA567: "the results seemed **implausible**"
- TA540: "**neither** method to be **robust**"
- TA530: "… the concerns about the **robustness** of the simulated treatment comparison"
- TA478: "…**uncertainty** about the **robustness** of the results"
- TA380: "…was **not consistent** with the population in the marketing authorisation"
- …

All effect modifiers and prognostic variables
- known
- adjusted for

NICE National Institute for Health and Care Excellence

NICE DSU TSD 18

"Provide evidence that absolute outcomes can be predicted with sufficient accuracy in relation to the relative treatment effects, and present an estimate of the likely range of residual systematic error. If this evidence cannot be provided or is limited, then state that the amount of bias in the indirect comparison is likely to be substantial, and could even exceed the magnitude of treatment effects which are being estimated."

Sensitivity analysis to assess the robustness of PAIC results

A methodological systematic review of studies implementing PAICs*

Sensitivity analysis to assess the robustness of PAIC results	Statistics
No sensitivity analysis	77 (47.5%)
Adjusting for different sets of covariates	55 (34.0%)
Applying additional inclusion/exclusion criteria to the IPD study	19 (11.7%)
Using different outcome definitions	7 (4.3%)
Using different follow-up time	11 (6.8%)
Other (e.g., using different approaches for handling missing data, implementing additional anchored/unanchored comparisons)	12 (7.4%)

*Truong et al. (2023) doi:10.1002/jrsm.1653

Sensitivity analysis to assess the robustness of PAIC results

A methodological systematic review of studies implementing PAICs*

Limitations acknowledged by authors	Statistics
No acknowledgement	5 (3.1%)
Unmeasured covariates	136 (84.0%)
Important covariates not reported in one of the included studies	60 (37.0%)
Limited sample size	31 (19.1%)
Heterogeneity across studies	139 (85.8%)
Small ESS/little overlap between populations	35 (31.6%)
Lack of a common comparator	23 (14.2%)

*Truong et al. (2023) doi:10.1002/jrsm.1653

Poll

How familiar are you with the concept of quantitative bias analysis?

- Very familiar
- Somewhat familiar
- Heard of it but not familiar
- Not familiar at all

Quantitative bias analysis (QBA)

- QBA is an umbrella term for the methods used to model systematic errors which may distort the results
  - Long history in epidemiology
  - Aim: to quantitatively measure the direction, magnitude and uncertainty associated with systematic errors on study results
- The analyses can be categorised to assess the impact of violations to:
  I. no unmeasured confounders;
  II. selection, participation and missing data are random within levels of adjusted covariates;
  III. no measurement error (including misclassification)

Basic idea of QBA

- QBA requires a model (also known as a bias model)
  - For the observed data (an outcome Y, an exposure/treatment A, observed covariates O) and unmeasured covariates (U)
  - Include one or more sensitivity/bias parameters
- Values of sensitivity parameters cannot be estimated from the data alone
- Values need to be pre-specified
  - **Deterministic** QBA: fixed values for the sensitivity parameters
  - **Probabilistic** QBA: a probability distribution for the sensitivity parameters
- A tipping point analysis
  - Identify the values for the sensitivity parameters that would change the study conclusion

Sensitivity analysis for unmeasured confounding for PAICs

- Major concern of unanchored MAIC and STC approach
  - Strong assumption that both prognostic factors and effect modifiers are adjusted for
- In practice, what could be adjusted for in the analysis depends on data availability
  - Information on baseline characteristics is limited in the comparator study
- QBA for unmeasured confounding via sensitivity analysis
  - PAICs

Sensitivity analysis approach based on simulating potential confounder(s)

B

IPD

Study B: IPD

Contains n observations on an outcome Y and $J + L$ observed covariates $\boldsymbol{X} = c(\boldsymbol{O}, \boldsymbol{U})$

Note that $\boldsymbol{U}$ is observed in Study B but not measured in Study A.

$\boldsymbol{X}$

Y	O_1	...	O_J	U_1	⋯	U_L
y_1	$o_{1,1}$	...	$o_{J,1}$	$u_{1,1}$	...	$u_{L,1}$
y_2	$o_{1,2}$	...	$o_{J,2}$	$u_{1,2}$	...	$u_{L,2}$
⋮	⋮	...	⋮	⋮	...	⋮
y_n	$o_{1,n}$	...	$o_{J,n}$	$u_{1,n}$	...	$u_{L,n}$

C

AgD

Study C: aggregate data

Contains reported treatment effect in Study C population $\hat{d}_{C(C)}$, and mean of the marginal distribution for J observed covariates $\boldsymbol{O}$

$E[\boldsymbol{O}] = \overline{\boldsymbol{O}}$

$\hat{d}_{C(C)}$	$E[O_1]$	$E[O_2]$	...	$E[O_J]$

?

$E[U_1] = \overline{U}_1, E[U_2] = \overline{U}_2, \ldots, E[U_L] = \overline{U}_L$

Sensitivity parameters

Deterministic QBA for unanchored STC

Study C: aggregate data

Contains reported treatment effect for Study C population $\hat{d}_{C(C)}$, and mean of the marginal distribution for J observed covariates $\boldsymbol{O}$

Study B: IPD

Contains n observations on an outcome Y and $J + L$ observed covariates $\boldsymbol{X} = c(\boldsymbol{O}, \boldsymbol{U})$

Note that $\boldsymbol{U}$ is observed in the B study but not measured in Study C.

$$E[\boldsymbol{O}] = \overline{\boldsymbol{O}}$$

$\hat{d}_{C(C)}$ $E[O_1]$ $E[O_2]$ ... $E[O_J]$

$\boldsymbol{X}$

Y	O_1	...	O_J	U_1	...	U_L
y_1	$o_{1,1}$	...	$o_{J,1}$	$u_{1,1}$	...	$u_{L,1}$
y_2	$o_{1,2}$	...	$o_{J,2}$	$u_{1,2}$	...	$u_{L,2}$
⋮	⋮	...	⋮	⋮	...	⋮
y_n	$o_{1,n}$	...	$o_{J,n}$	$u_{1,n}$	...	$u_{L,n}$

Probabilistic QBA

Assume a distribution for $\tilde{U}$

	Deterministic QBA
Step 1: Outcome regression model	Build the regression model based on the IPD from Study B $g(\theta_{i(B)}) = \beta_0 + \boldsymbol{\beta_1^T O_i} + \boldsymbol{\beta_2^T U_i}$ $\boldsymbol{U}$ refers to the unmeasured confounders in Study C and is observed in Study B.
Step 2: Prediction	Predict the treatment effect for Study C population $\hat{d}_{B(C)} = g(\hat{\theta}_{B(C)})$
	The marginal mean of $\boldsymbol{U}$ for Study C, $\tilde{U} = E[U]$, is the sensitivity parameter.
	Assume a fixed value for $\tilde{U}$ → $\overline{O}, \Sigma^{O,U}_{IPD}$ → NORTA/Gaussian copula to sample individuals → $\boldsymbol{O_1}$ $\boldsymbol{U_1}$ ⋮ ⋮ $\boldsymbol{O_N}$ $\boldsymbol{U_N}$ → $\hat{d}_{B(C)} = g\left(\frac{1}{N}\sum_{i=1}^{N} g^{-1}(\hat{\beta}_0 + \boldsymbol{\hat{\beta}_1^T O_i} + \boldsymbol{\hat{\beta}_2^T U_i})\right)$
Step 3: Obtain relative effect	Obtain the relative treatment effect using the prediction from Step 2 and reported aggregate data for Study C $\hat{d}_{BC(C)} = \hat{d}_{C(C)} - \hat{d}_{B(C)}$

Case study

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Case study

- Re-analyse data from the PRIME study
  - A Phase III RCT of panitumumab with FOLFOX4 vs. FOLFOX4 alone in patients with previously untreated metastatic colorectal cancer
  - Obtain anonymous IPD for the PRIME study from the Project Data Sphere® platform
  - Drop the FOLFOX4 arm and treat the data in the panitumumab with FOLFOX4 arm as a single-arm trial
  - Obtain summary statistics for the FOLFOX4 arm from an external source (Cunningham et al. 2009)
  - Outcome: objective response rate

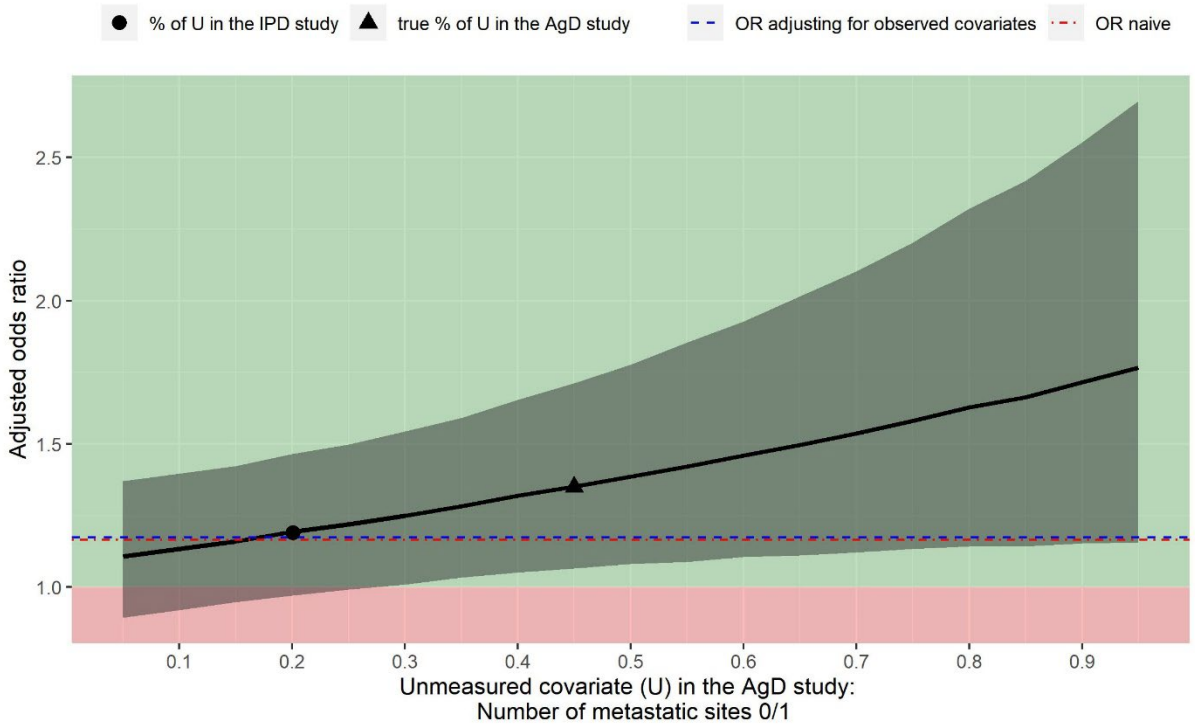
Data

	The PRIME trial		Cunningham *et al.* (2009)
Characteristic	**Panitumumab + FOLFOX4 (n=468)***	**FOLFOX4 (n=467)***	**FOLFOX4 (n=362)**
Male (%)	66	61	65
Age, years (%)			
≤65	60	62	67
65	40	38	33
ECOG performance status (%)			
0/1	95	95	93
≥2	5	5	7
Primary tumour type (%)			
Colon	67	69	56
Rectal and other	33	31	44
Number of metastatic sites (%)			
0/1	20	20	45
≥2	80	80	55
Metastatic site (%)			
Liver alone	18	16	33
Prior adjuvant chemotherapy (%)	15	12	27
Prior surgery (%)	91	91	87
Objective response rate (%)	57.9	53.3	54.1

OR from PRIME
1.20 (95% CI, 0.93 to 1.56)

Naïve ITC
1.17 (95% CI, 0.88 to 1.54)

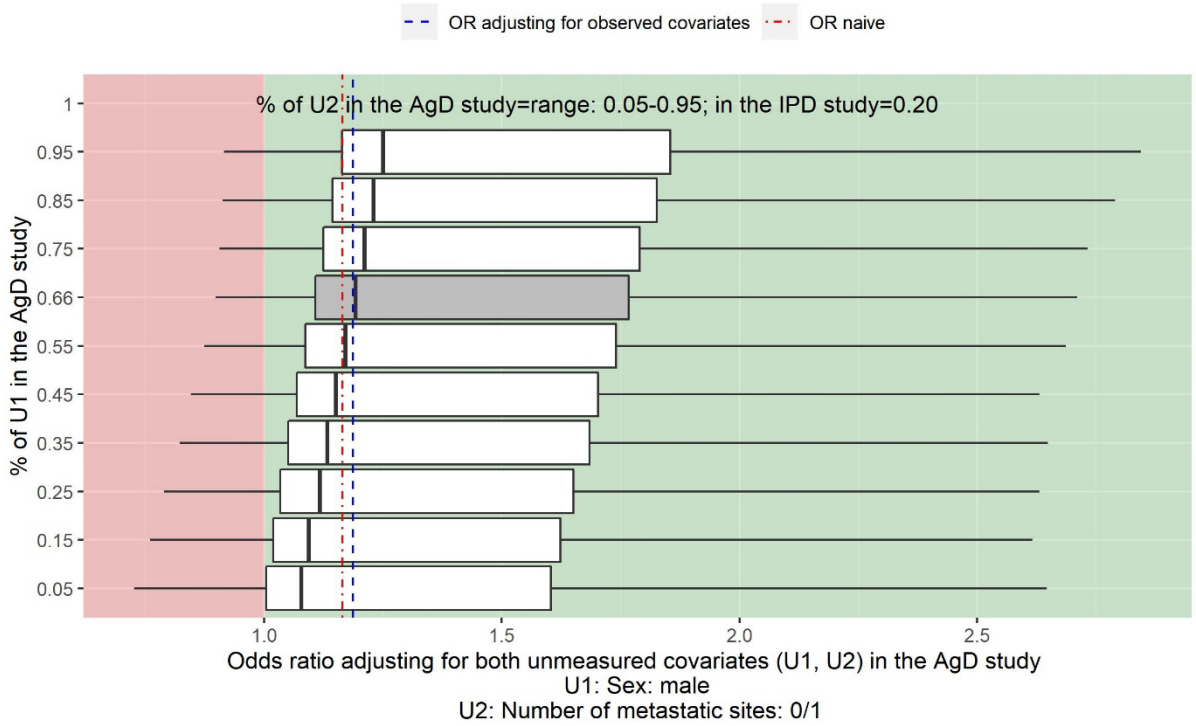
Sensitivity analysis: number of metastatic sites unmeasured



Naïve OR:
1.17 (95% CI, 0.88 to 1.54)

OR adjusted for observed **X**:
1.18 (95% CI, 0.96 to 1.44)

Sensitivity analysis: sex and number of metastatic sites unmeasured



Naïve OR:
1.17 (95% CI, 0.88 to 1.54)

OR adjusted for observed **X**:
1.19 (95% CI, 0.97 to 1.45)

Summary

- Unanchored MAIC and STC are heavily criticised for its strong assumptions
  - Robustness?
- QBA formally quantifies the bias associated with unmeasured confounding
  - Provide a quantitative assessment of the impact of this bias
  - Increase the robustness of the ITC approach for single-arm trials

Thank you!

Questions?

Reference

- Patel D, Grimson F, Mihaylova E, et al. Use of External Comparators for Health Technology Assessment Submissions Based on Single-Arm Trials. Value Health. Aug 2021;24(8):1118-1125. doi:10.1016/j.jval.2021.01.015
- Truong B, Tran LT, Le TA, Pham TT, Vo TT. Population adjusted-indirect comparisons in health technology assessment: A methodological systematic review. Res Synth Methods. Sep 2023;14(5):660-670. doi:10.1002/jrsm.1653
- Cunningham D, Sirohi B, Pluzanska A, et al. Two different first-line 5-fluorouracil regimens with or without oxaliplatin in patients with metastatic colorectal cancer. *Ann Oncol*. Feb 2009;20(2):244-50. doi:10.1093/annonc/mdn638
- Ren, S, et al. Advancing unanchored simulated treatment comparisons: A novel implementation and simulation study. Res Synth Methods. 2024 (in press)