Quantitative Bias Analysis in Practice

Anna Schultze
London School of Hygiene and Tropical Medicine

With thanks to **Ian Douglas** and **Marleen Bokern** for slides

Validate study variables to reduce misclassification bias: recent tools and research needs

HYBRID WORKSHOP

27 March 2025 - 14.30-18.30
ARS Toscana - Villa La Quiete, via Pietro Dazzi 1 - Florence, Italy

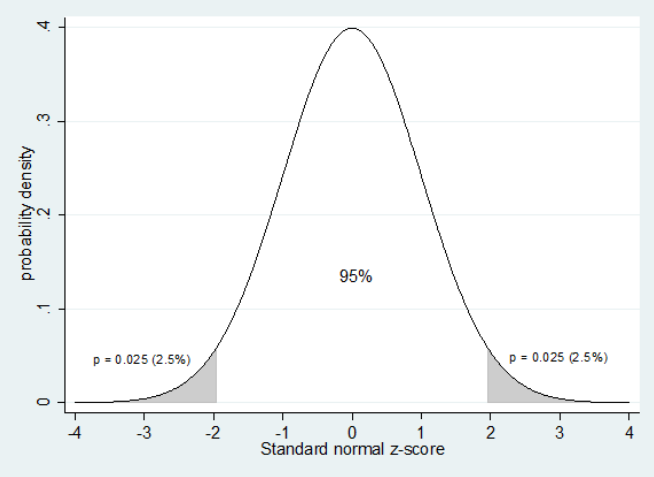
ELECTRONIC HEALTH RECORDS RESEARCH GROUP

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

How do we handle error in epidemiological studies?

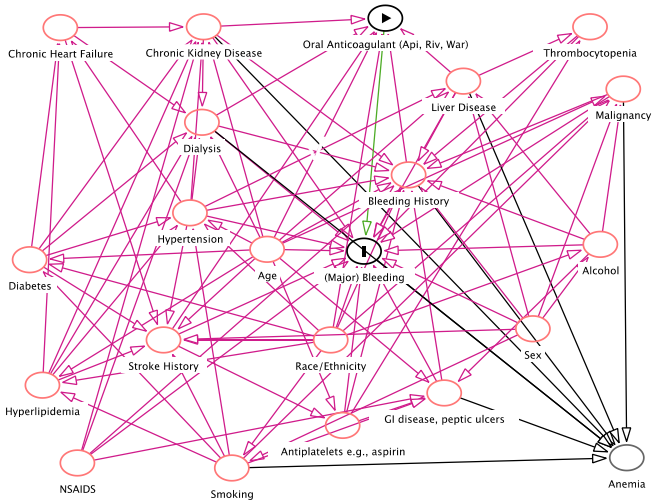
Random Error

Quantitative assessments
Confidence intervals, p-values



Confounding

Careful adjustment
Often sensitivity analyses



Misclassification

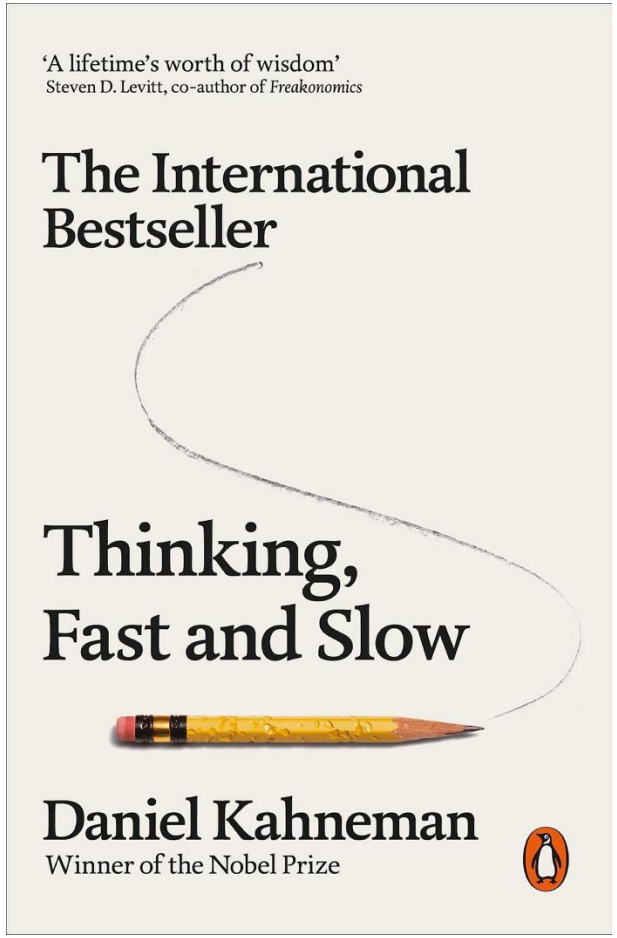
"blah blah … bias towards the null"

DAG from Yap et al, in press

The status quo needs improving

- ”Non-differential misclassification biases towards the null”, ***except…***
  - **Chance:** bias towards null in expectation, not always
  - **More than 2 categories:** may bias away from null
  - **Dependency of 2 misclassified variables:** hard to predict
  - **Outcome misclassification with 100% specificity:** no bias
  - **Misclassification of a confounder:** direction of confounding

- The extent of bias matters
  - RR = 1.5 vs. RR = 1.6 vs. RR = 55
  - We are bad at intuiting the extent of bias

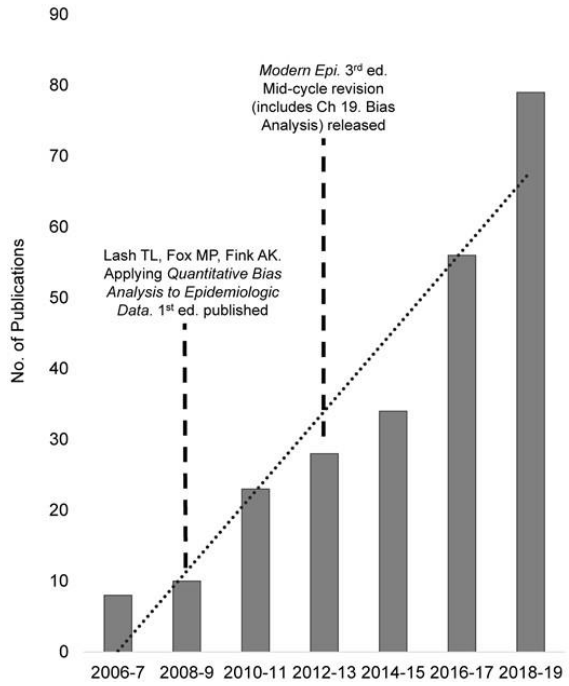


Yland et al, AJE, 2022; Lash et al, Epidemiology, 2021

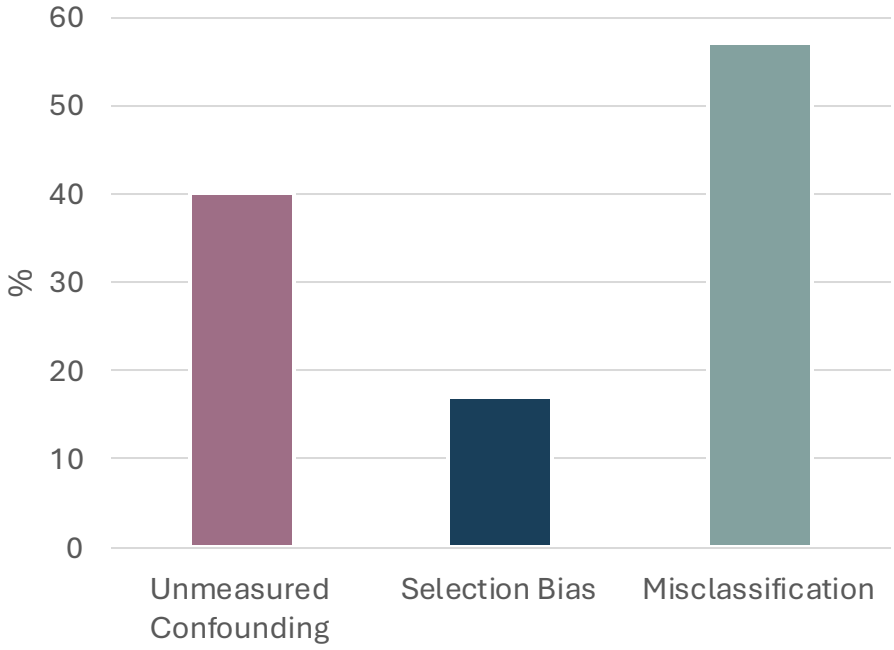
Quantitative Bias Analysis (QBA)

- Methods which quantify the **direction, magnitude** and **uncertainty** of systematic error

QBA is increasingly used



Most often for misclassification



Petersen et al, IJE, 2021

QBA for misclassification

- **Aim:** correct an effect estimate, and its uncertainty, for misclassification
- Assumptions - bias parameters : sensitivity, specificity, NPV or PPV
  - Internal validation **29%**
  - External validation (assuming transportability) **48%**
  - Educated guess **24%**
- Different methods
  - Simple / Multidimensional **36%**
  - Probabilistic **57%**

Petersen et al, IJE, 2021

Simple QBA

Expected Truth – corrected for outcome misclassification

		Exposure	
		E +	E -
Outcome	Outcome	$A = \frac{a - (a + c)(1 - Sp)}{Se - (1 - Sp)}$	$B = \frac{b - (b + d)(1 - Sp)}{Se - (1 - Sp)}$
	No outcome	$C = (a + c) - A$	$D = (b + d) - B$
		a + c = A + C	b + d = B + D

Probabilistic Bias Analysis (PBA)

Uses Monte Carlo sampling methods to generate a distribution of bias-adjusted effect estimates

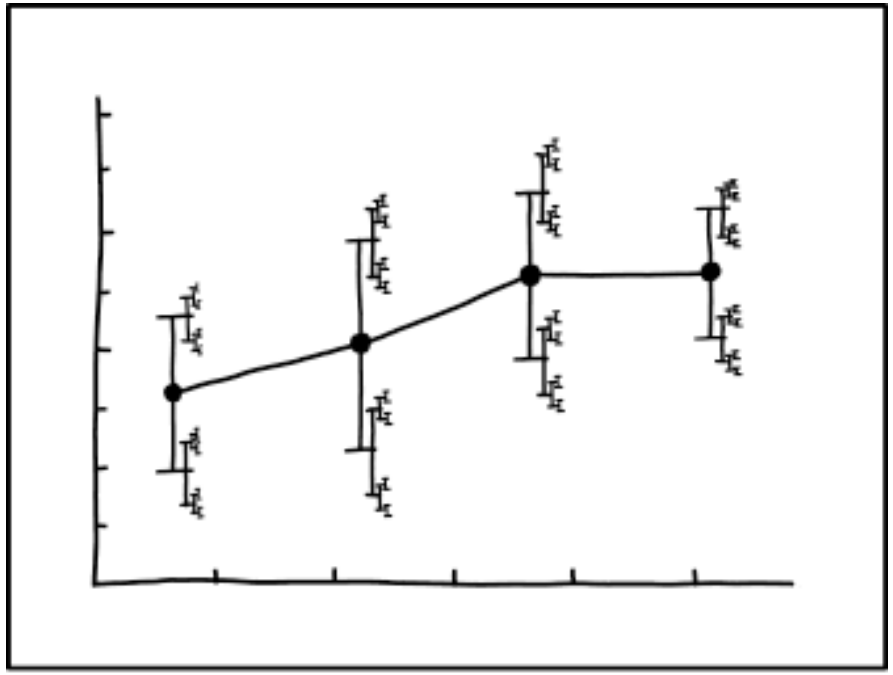
1) Assign probability distributions to bias parameters
2) Sample from this distribution
3) Use the sampled values to correct 2x2 table
    a. *Estimate prevalence of outcome in each exposure group*
    b. *Use prevalence, SE and SP to calculate PPV and NPV*
    c. *Correct 2x table using PPV and NPV*
    d. *Calculate corrected OR*
4) Repeat many times

Probabilistic Bias Analysis (PBA)

Summary Level	Record Level
• Calculated on 2x2 table • No confounder adjustment	• Simulated outcome for each record • Full analytical flexibility

Incorporating uncertainty

- Want to quantify **both** random and systematic error
- Simple bias analysis:
  - Cannot calculate 95%CI directly from the corrected cell counts --> separate formulas
- Probabilistic bias analysis:
  - Simple approximate method
  - Simulates random error around bias adjusted estimate
  - Summarize the resulting distribution as median and simulation interval (2.5th to 97.5th percentile)
  - ‘total error’ - both systematic and random

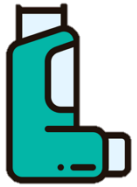


Fox, MacLehose and Lash, 2023

Example: COVID-19 drug studies

- First lockdown
- Lots of interest in how drugs impact covid risk
- Extreme time pressure
- ICS proposed as potential repurposing candidate
- Original study (Schultze et al, 2020) → harm
  - Later RCTs → null, or possible protection
- Marleen Bokern PhD:
  - explore impact of different errors
  - how can we do better next time?

Inhaled corticosteroids (ICS) and risk of COVID death



People with COPD in CPRD

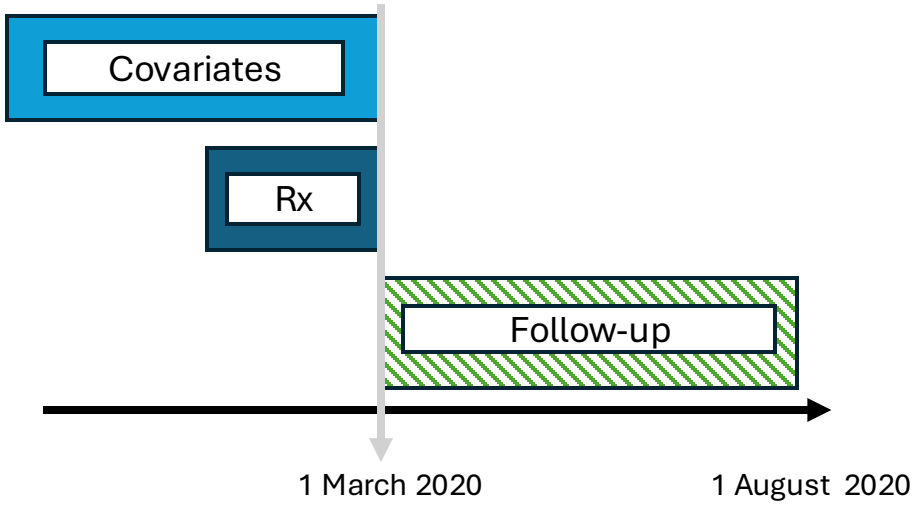
ICS compared to LABA/LAMA

National death registration (ONS)

ICD-10 codes (U07.1, U07.2)

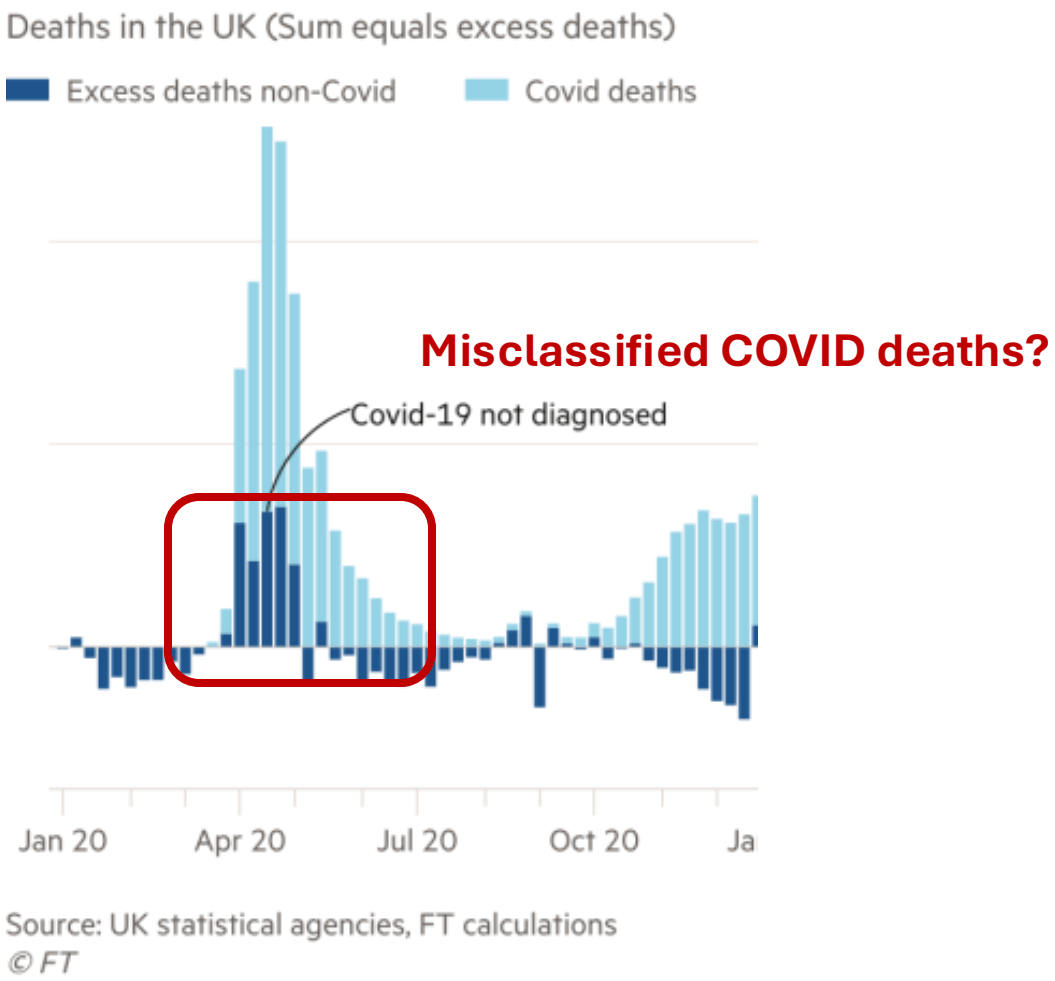
Schultze et al, 2020; Bokern et al, 2025

The basic study design



Schultze et al, 2020; Bokern et al, 2025

COVID-19 outcome misclassification

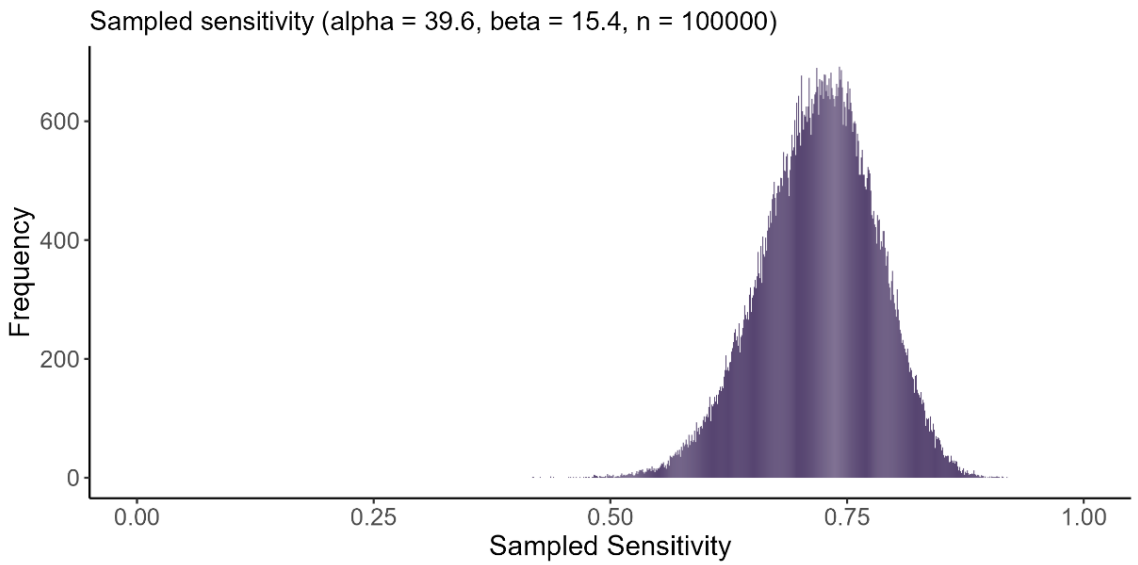


- Low sensitivity - > 20% of COVID deaths misclassified
- Reduced specificity - limited testing capacity
- Misclassification of cause
  - ***why*** – not whether – someone died
- Was this an issue?

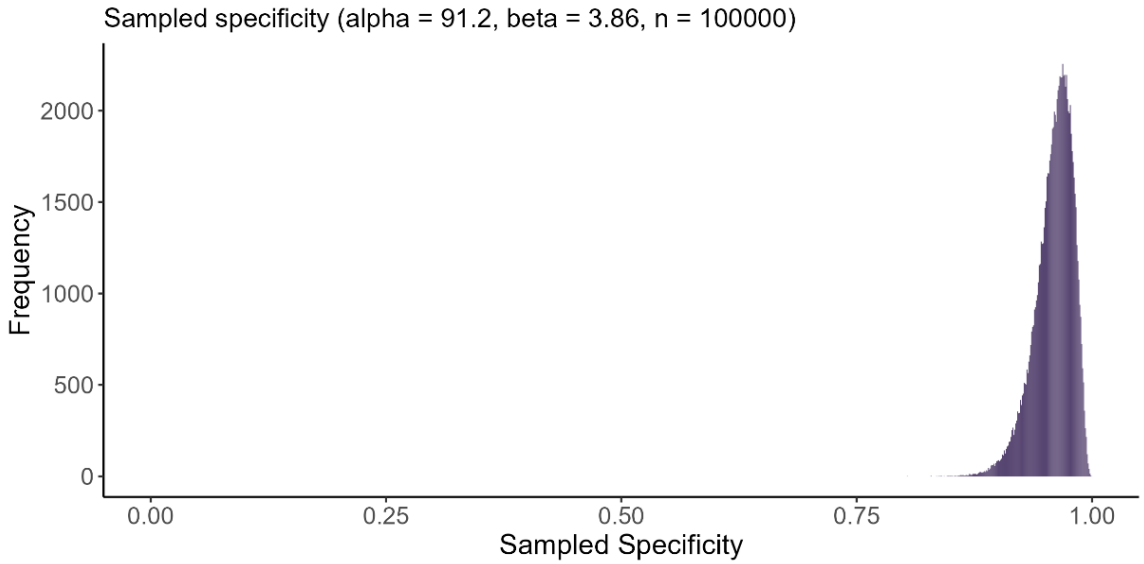
Kontopantelis, 2021

Bias parameters & assumptions

- No validation study
- Solution : use three sources of death information
  - Excess deaths estimates, ONS death registrations and UKHSA (deaths among those with + test)
  - Assume excess deaths were all due to COVID-19



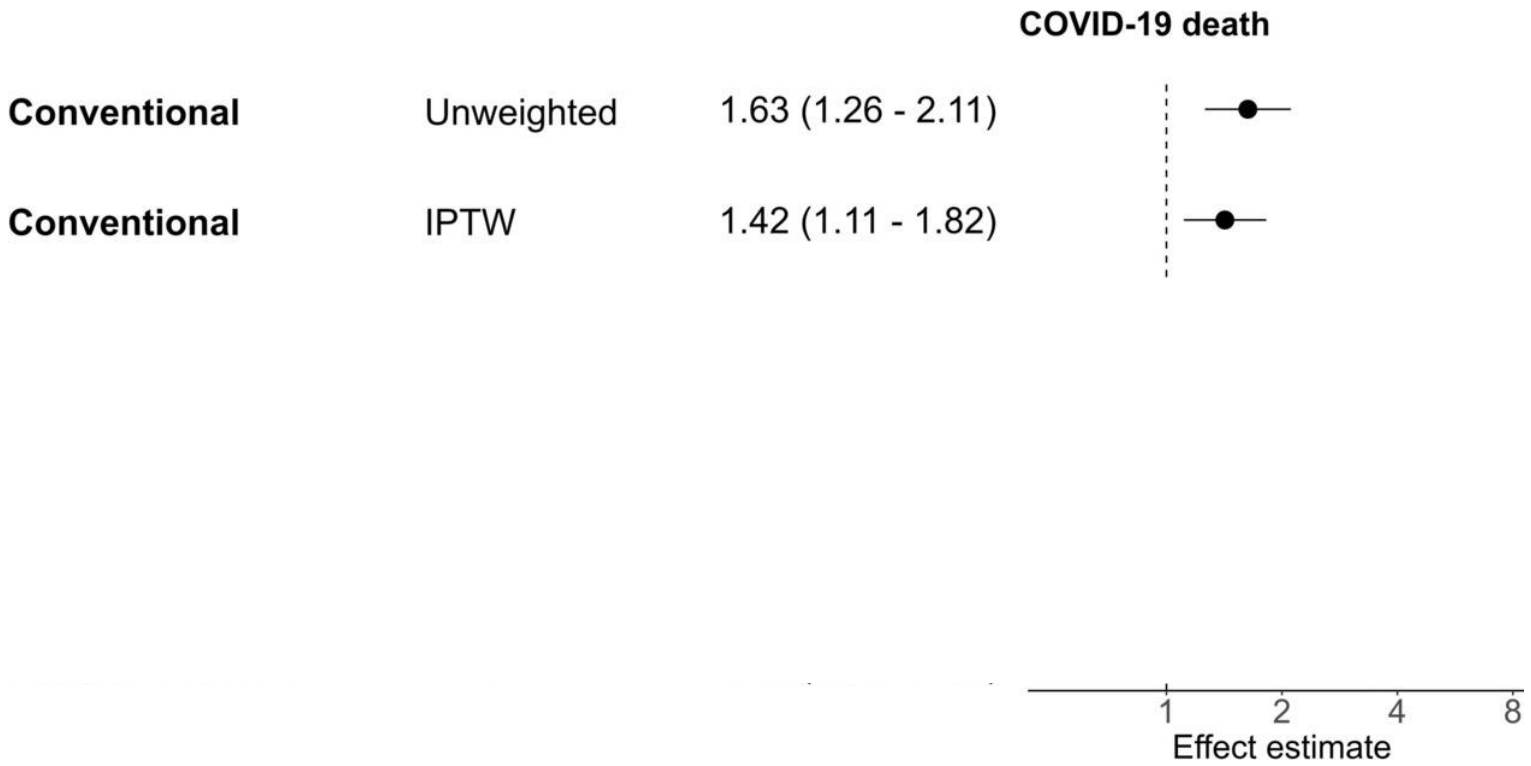
median = 0.72
2.5th%-ile = 0.60
97.5th%-ile = 0.83



median = 0.96
2.5th%-ile = 0.91
97.5th%-ile = 0.99

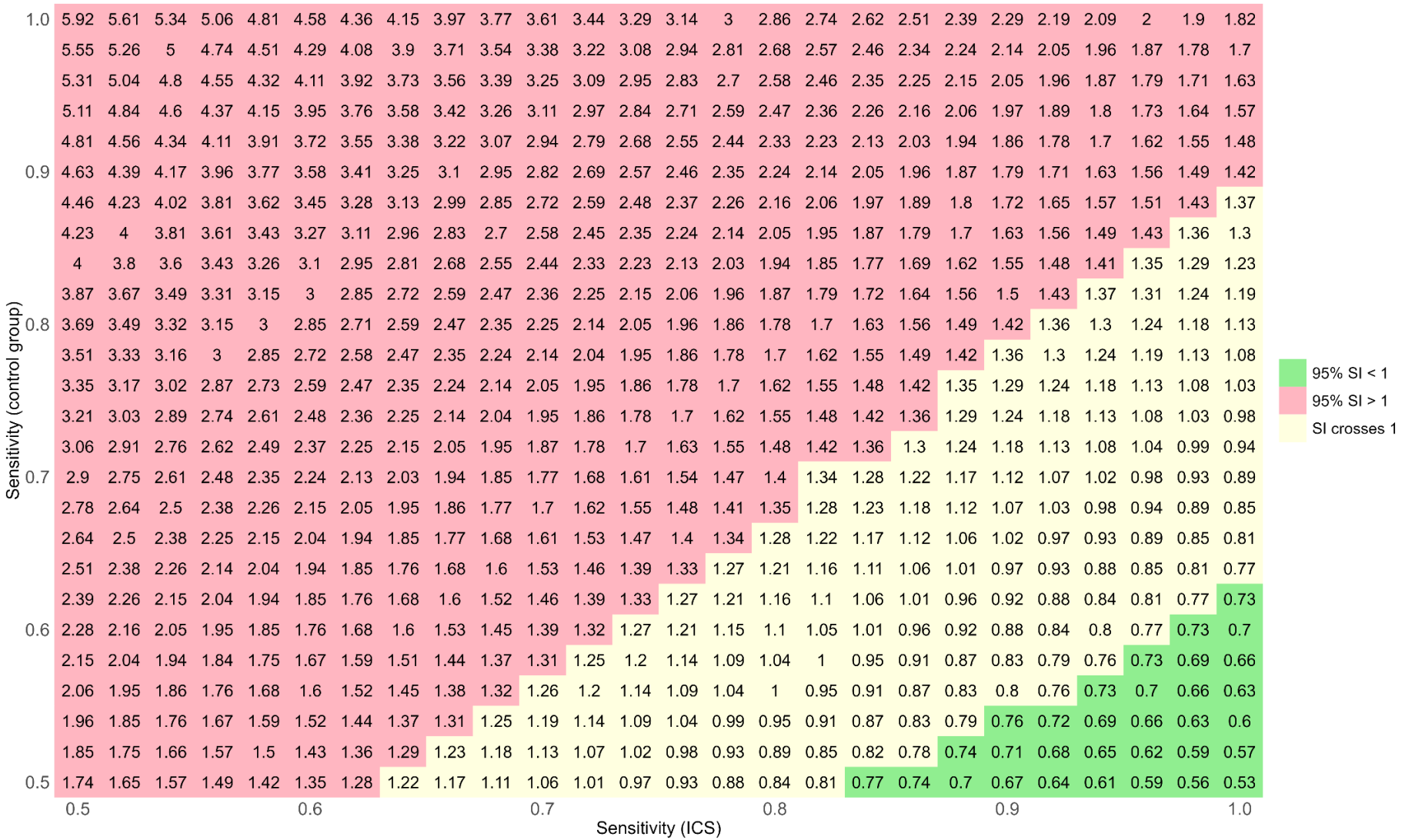
Results – non-differential sensitivity

Odds ratios and 95% confidence or simulation intervals for COVID-19 hospitalisations and deaths, comparing ICS/LABA (± LAMA) users to LABA/LAMA users



Results – differential sensitivity

Point estimates under a range of scenarios



100,000 simulations were run per combination of sensitivity values. Specificity was set at 0.97. SI = simulation interval

Interpretation

- We missed a lot of COVID-19 deaths, but in the context of high specificity non-differential reduction in sensitivity did not matter
  - Small shift away from null driven by reduced specificity
  - ***Increased uncertainty***
- The extent of differential misclassification needed for us to 'miss' a protective effect would need to be substantial
  - Unlikely in context of active comparator
- Importance of considering multiple sources of error
  - Confounding!!

Recommendations to Xabi

- If you are doing a validation study, report validity measures by exposure / outcome

- Pre-specify when you will run a QBA, and how you will define the bias parameters

    - If your main results are very imprecise, QBA may not be informative
    - Very important to ensure bias parameters are reasonable for your target population
        - If large number of negative cell counts, implies bias parameters not supported by data

- To incorporate uncertainty → probabilistic bias analysis (either summary or record-level)

- If you are doing a distributed analysis, summary-level PBA may be easier to implement
    - Could be run on summary data by a single analyst
    - Record level is more computationally intensive

- If confounding has a large impact, record-level PBA likely preferable
    - If PS-weighting, might be able to apply summary-level PBA to weighted 2x2 summary tables

Recommendations to Ersilia / Robert

- How do we best use QBA in more complex scenarios?
  - Bias parameters vary with values of one or more (unmeasured) confounders
  - Bias parameters vary over time
  - Time to event outcomes, where time is also misclassified

- When is it okay to 'transport' validity measures (SE, SP?) between populations?

- How does QBA compare to alternative methods to handle misclassification?
  - Multiple imputation

- How can interpretation of QBA be improved?
  - Tendency for QBA to *'not substantially change conclusions'* (Petersen, 2021)
  - Should interpretation be pre-specified?

- Why is QBA not more widely used by researchers and regulators?
  - Many tools / resources are available : so what are the barriers?

Thank you!

Any questions?

Contact: anna.schultze@lshtm.ac.uk

ELECTRONIC HEALTH RECORDS RESEARCH GROUP

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE