Adjusting number of cases, risk, risk difference and risk ratio estimated through an algorithm with unknown sensitivity

Giorgio Limoncella[1], **Leonardo Grilli**[1], **Emanuela Dreassi**[1], **Carla Rampichini**[1], **Robert Platt**[2], **Rosa Gini**[3]

[1] University of Florence, Florence, Italy
[2] McGill University, Montreal, Canada
[3] ARS Toscana, Florence, Italy

Validating Study Variables to Reduce Misclassification Bias: Recent Tools and Research Needs

March 27, 2025, ARS Toscana

UNIVERSITÀ DEGLI STUDI FIRENZE

UNIVERSITÀ DEGLI STUDI FIRENZE

Algorithms to retrieve variables of interest

1 Background

- Real-world evidence may be generated by making secondary use of electronic healthcare databases, whose primary purpose is different from research
- Variables of interest are measured using algorithms (**A**), which may misclassify cases leading to biased number of cases and measures of risk and association

Usual Strategy in Pharmacoepidemiology research

Define a highly specific algorithm at the expense of sensitivity (**SE**).

Consequences of low sensitivity and study objective

- Number of cases, prevalence, and incidence may be underestimated
- Risk ratio is unbiased **if sensitivity is non differential**
- Risk difference may be biased, often underestimated in absolute value

Objective

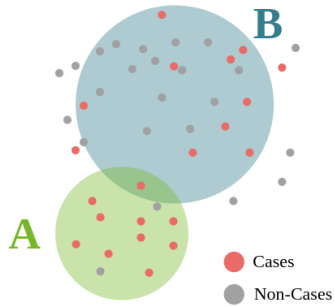
To develop a methodology to:

- reduce bias from low sensitivity
- test the assumption of non-differential sensitivity

Methods: screening algorithm

We build on a strategy suggested by Lanes and Beachler[1]:

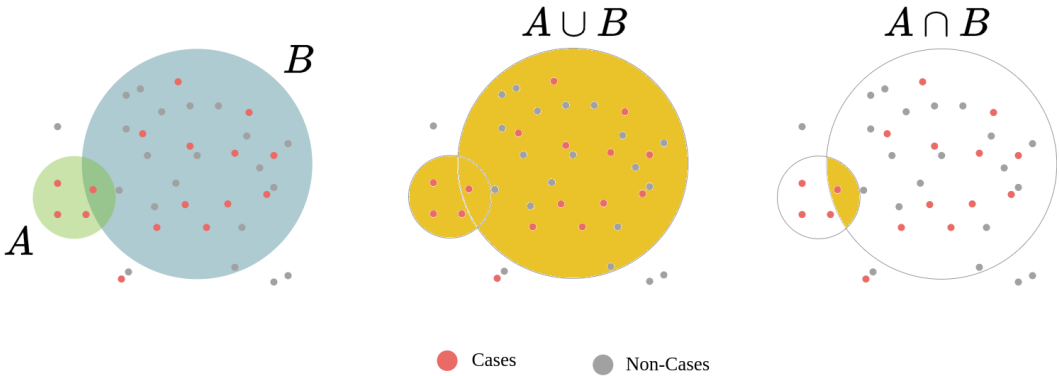
1. Develop a screening algorithm (**B**)
2. Perform a validation study on $PPVs$
3. Estimate sensitivity of A



[1] Lanes, Stephan, and Daniel C. Beachler. "Validation to correct for outcome misclassification bias." *Pharmacoepidemiology and Drug Safety* 32.6 (2023): 700-703.

Composite algorithms

2 Methods



Sensitivity of A as function of $PPVs$

2 Methods

To estimate sensitivity of A, we exploit relation between sensitivities across algorithms:

$$SE_{A\cup B} = SE_A + SE_B - SE_{A\cap B}$$

and interrelation between validity indices [2] e.g.:

$$SE = \frac{P \times PPV}{\pi}$$

then:

if $SE_{A\cup B} = 1$

$$SE_A = \frac{P_A PPV_A - P_{A\cap B} PPV_{A\cap B}}{P_A PPV_A + P_B PPV_B - P_{A\cap B} PPV_{A\cap B}}$$

if $SE_{A\cup B} \leq 1$

$$SE_A \leq \frac{P_A PPV_A - P_{A\cap B} PPV_{A\cap B}}{P_A PPV_A + P_B PPV_B - P_{A\cap B} PPV_{A\cap B}}$$

where: P is the observed prevalence and π is the true prevalence

[2] Bollaerts, Kaatje, et al. "Disease misclassification in electronic healthcare database studies: Deriving validity indices—A contribution from the ADVANCE project." PLoS One 15.4 (2020): e0231333.

Summary of application

2 Methods

Population not stratified per exposure

	No screening algorithm	**With screening algorithm**	
	PPV_A available	PPV_A, PPV_B, $PPV_{A\cap B}$ available	
	–	$SE_{A\cup B} = 1$	$SE_{A\cup B} < 1$
Sensitivity of A	Not estimable	Unbiased	Upper bound
Number of cases	Underestimated	Unbiased	Reduced underestimation
Risk	Underestimated	Unbiased	Reduced underestimation

Summary of application

2 Methods

Population not stratified per exposure

	No screening algorithm	**With screening algorithm**	
	PPV_A available	PPV_A, PPV_B, $PPV_{A\cap B}$ available	
	–	$SE_{A\cup B} = 1$	$SE_{A\cup B} < 1$
Sensitivity of A	Not estimable	Unbiased	Upper bound
Number of cases	Underestimated	Unbiased	Reduced underestimation
Risk	Underestimated	Unbiased	Reduced underestimation

Population stratified per exposure

	No screening algorithm	**With screening algorithm**		
	PPV_A available in both strata	PPV_A, PPV_B, $PPV_{A\cap B}$ available in both strata		
	–	$SE_{A\cup B} = 1$	$SE^e_{A\cup B} = SE^{\bar{e}}_{A\cup B} < 1$	$SE^e_{A\cup B} \neq SE^{\bar{e}}_{A\cup B}$
Risk Ratio	Unbiased if $SE^e_A = SE^{\bar{e}}_A$	Unbiased	Unbiased	Unbiased if $SE^e_A = SE^{\bar{e}}_A$
Risk Difference	Biased	Unbiased	Lower bound in absolute value	Biased
Differentiality of SE_A	Not testable	Testable	Testable	Not Testable

Non-Differentiality Test

2 Methods

Assuming $SE^{E}_{AUB} = SE^{\bar{E}}_{AUB}$:

$$H_0 : \frac{SE^{e}_{A}}{SE^{\bar{e}}_{A}} - 1 = 0$$

$$\Downarrow$$

$$H_0 : \frac{P^{e}_{A} PPV^{e}_{A} (P^{\bar{e}}_{A} PPV^{\bar{e}}_{A} + P^{\bar{e}}_{B} PPV^{\bar{e}}_{B} - P^{\bar{e}}_{A\cap B} PPV^{\bar{e}}_{A\cap B})}{P^{\bar{e}}_{A} PPV^{\bar{e}}_{A} (P^{e}_{A} PPV^{e}_{A} + P^{e}_{B} PPV^{e}_{B} - P^{e}_{A\cap B} PPV^{e}_{A\cap B})} - 1 = 0$$

We conducted simulation studies performing the test with the **bootstrap** method, computing the power of the test across different scenarios.[3]

[3] Limoncella, Giorgio, et al. "Addressing bias due to measurement error of an outcome with unknown sensitivity in database epidemiological studies." American Journal of Epidemiology (2024): kwae423.

Some notes on the assumptions

- $SE^{e}_{A} = SE^{\bar{e}}_{A}$ is not always realistic, especially in safety studies:

  Knowing that a patient used a new medication or one with a suspected adverse reaction, the coder may use a more specific diagnostic code

- $SE_{A \cup B} = 1$ may be realistic in particular scenarios, e.g. A = *infarction* and B = *death*

- $SE^{e}_{A \cup B} = SE^{\bar{e}}_{A \cup B} < 1$ may be likely in several scenarios, e.g. outcomes that require hospitalisation

ENTRESTO (Post-authorisation safety study)

3 Case studies

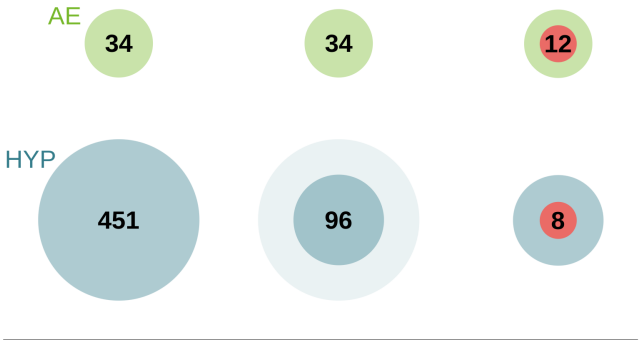
- **Outcome of interest**: Angioedema
- **Main algorithm**: 9951 ICD9CM diagnostic code (angioneurotic edema, **AE**)
- **Screening algorithm**: 9950, 99527, 7088, 7080, 7081, 7089, 37633, 37482, 47825, 4786, 47875, 5088, 7823, 9952 ICD9CM diagnostic codes (hypersensitivity, **HYP**)

$SE_{AE \cup HYP} = 1$

This would imply that all the subjects with angioedema accessed the hospital and were diagnosed with one of the codes included in the AE or HYP algorithms

ENTRESTO results

3 Case studies



$$PPV_{AE} = 35.4\%$$

$$PPV_{HYP} = 8.3\%$$

$$\Downarrow$$

$$\mathbf{SE_{AE}} \leq 24\%^{*}$$

$$\mathbf{N} \geq 49^{*}$$

* C.I. can be computed by bootstrap

Discussion

- **Limitation**:
  - Some assumptions are still required (e.g., assuming non-differentiality of $SE_{A \cup B}$ to test that of SE_A)
  - The methodology does not yet cover all scenarios (misclassification of exposure, hazard, ...)
- Estimate on validity indices can be integrated into quantitative bias analysis
- Code for non-differentiality test available on GitHub: `https://github.com/GiorgioLimoncella/NonDifferentialityTest`

Considerations for the investigator

Is it possible to conduct a validation study for the main outcome variable?

Can this be done stratified by exposure status?

And if so, could a screening algorithm be defined that, even with limited specificity, would still identify some true cases missed by the primary algorithm?

And if so, can you assume that the sensitivity of $SE_{A \cup B}$ is non-differential?

Considerations for the Researchers

How a possible delay in detecting events can be managed?

What happen if this delay is differential?

Can we envision how to estimate a possible delay and adjust the analysis?