Quantitative Bias Correction for Misclassification

Core Correction Formula:

For non-differential misclassification of exposure:

$$P_{\text{true}} = \frac{P_{\text{observed}} - (1 - \text{Specificity})}{\text{Sensitivity} + \text{Specificity} - 1} \tag{1}$$

Where:

- P_{true} = True prevalence of exposure
- P_{observed} = Observed prevalence of exposure
- Sensitivity = Probability of correctly classifying exposed individuals
- Specificity = Probability of correctly classifying unexposed individuals

Step-by-Step Correction Process

Parameter	Value
Sensitivity	0.80 (80%)
Specificity	0.80 (80%)
Total with disease	100
Total without disease	100

Observed Data (with misclassification):

	Disease +	Disease −	Total
Exposed	68	32	100
Unexposed	32	68	100
Total	100	100	200

Correction Calculations:

1. Calculate observed exposure prevalence in each disease group:

$$P_{\text{obs}}(\text{Exposed}|\text{Disease}+) = \frac{68}{100} = 0.68$$
$$P_{\text{obs}}(\text{Exposed}|\text{Disease}-) = \frac{32}{100} = 0.32$$

2. Apply correction formula to each group:

$$P_{\text{true}}(\text{Exposed}|\text{Disease}+) = \frac{0.68 - (1 - 0.80)}{0.80 + 0.80 - 1} = \frac{0.68 - 0.20}{0.60} = \frac{0.48}{0.60} = 0.80$$

$$P_{\text{true}}(\text{Exposed}|\text{Disease}-) = \frac{0.32 - (1 - 0.80)}{0.80 + 0.80 - 1} = \frac{0.32 - 0.20}{0.60} = \frac{0.12}{0.60} = 0.20$$

3. Calculate corrected cell counts:

$$\begin{aligned}
\text{True Exposed \& Disease}+ &= 0.80 \times 100 = 80 \\
\text{True Unexposed \& Disease}+ &= 0.20 \times 100 = 20 \\
\text{True Exposed \& Disease}- &= 0.20 \times 100 = 20 \\
\text{True Unexposed \& Disease}- &= 0.80 \times 100 = 80
\end{aligned}$$

Corrected Data:

	Disease +	Disease −	Total
Exposed	80	20	100
Unexposed	20	80	100
Total	100	100	200

Results:

Measure	Observed	Corrected
Risk Ratio	2.125	4.00
Bias Factor	0.53	1.00
% Attenuation	47%	0%

Important Assumptions:

- Non-differential misclassification (same error rates in all groups)
- Known sensitivity and specificity from validation study
- Misclassification is independent of other variables

Reference:
Fox, M. P., MacLehose, R. F., & Lash, T. L. (2021). *Applying quantitative bias analysis to epidemiologic data* (Vol. 10, pp. 978-3). New York: Springer.

Full text available at: https://link.springer.com/content/pdf/10.1007/978-3-030-82673-4.pdf