I^2 statistic in meta-analysis of prevalence: worthwhile or worthless?

Celina Borges Migliavaca, PhD(c)

PERSyst (Prevalence Estimates Reviews – Systematic Review Methodology) **Group**

Cochrane Training

A day with... Statistical Methods Group (SMG)

UFRGS
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

@ celinabm7@gmail.com

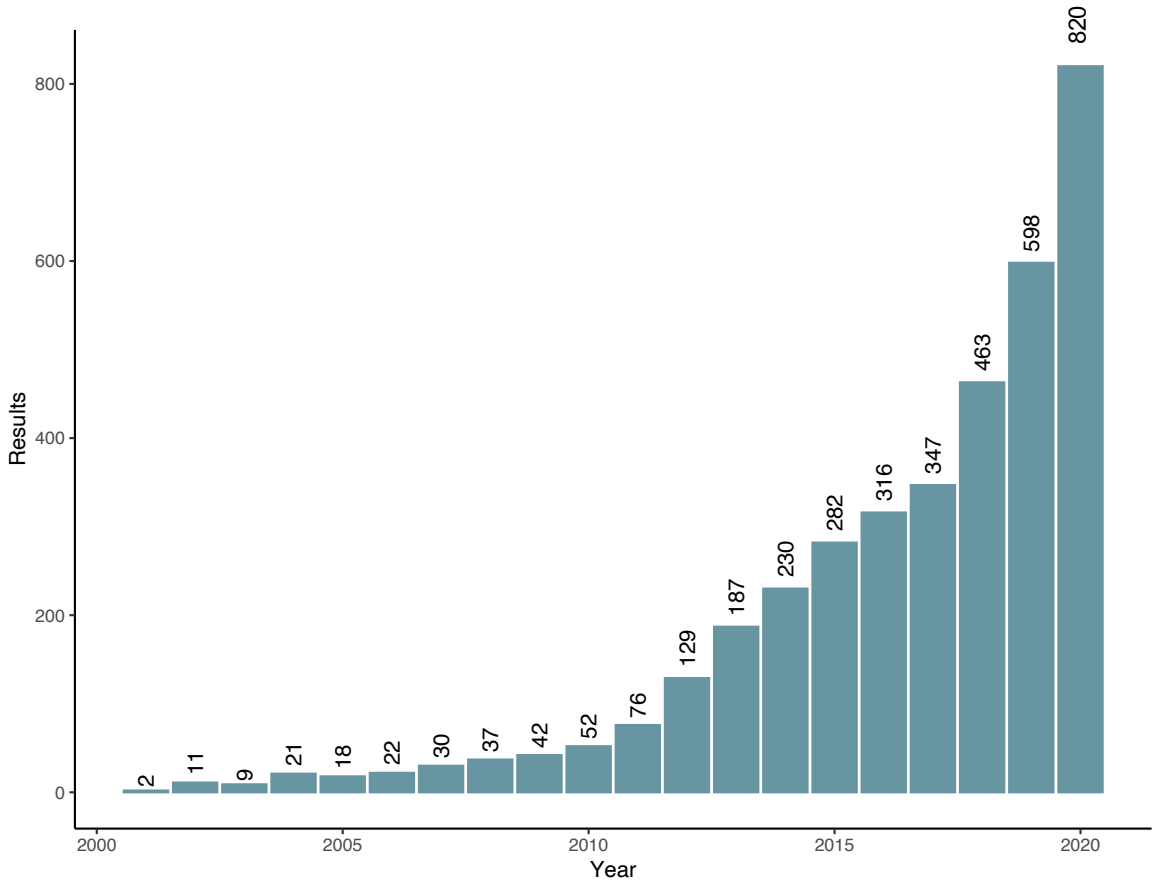
celinaborgesm

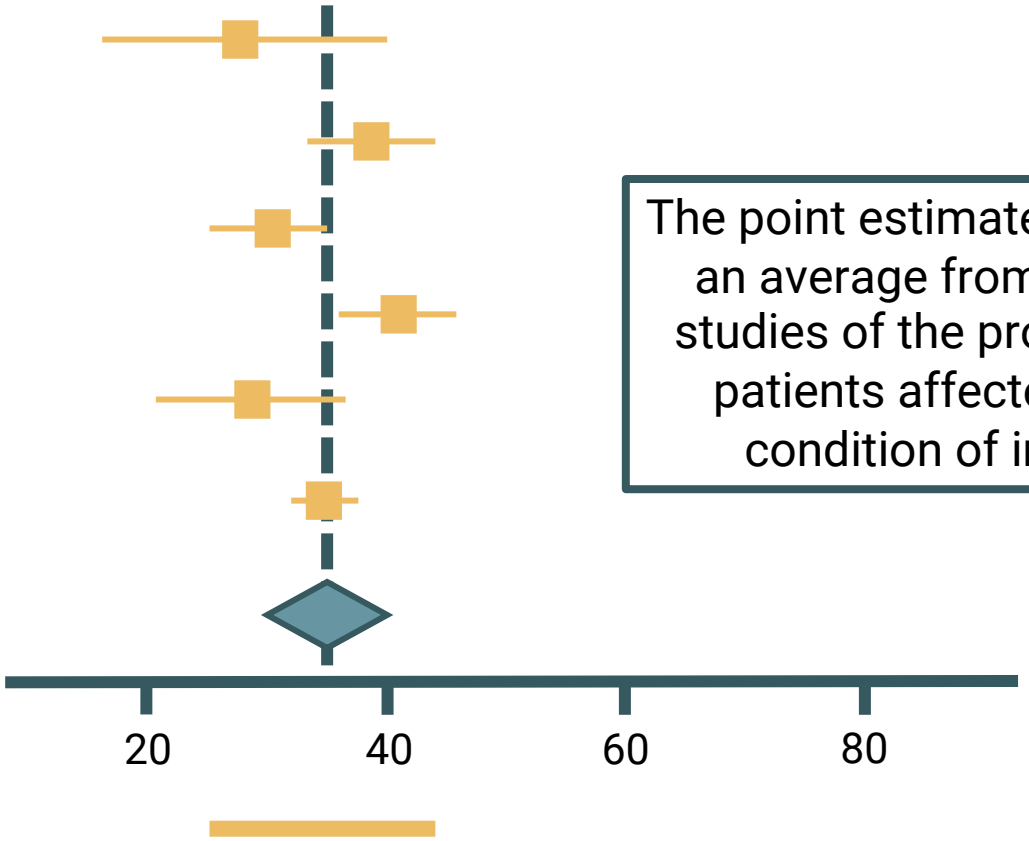
Prevalence estimates are critical for health-related decision making. They provide necessary information when estimating disease burden and are essential for issues such as priority setting and health technology assessments.

Systematic reviews and meta-analyses are the standard methods for the synthesis of available evidence to provide an answer to a specific healthcare question.

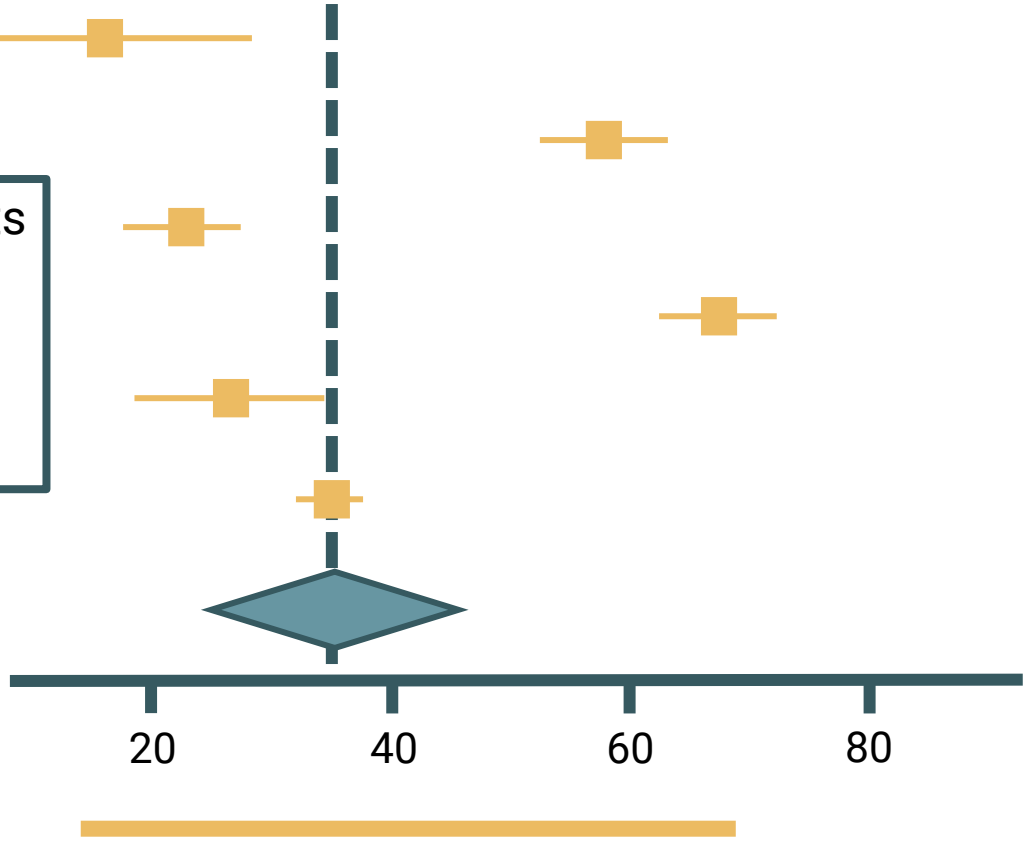
Number of systematic reviews of prevalence indexed in MEDLINE per year

Search strategy: “prevalence” AND “systematic review”, in the title.



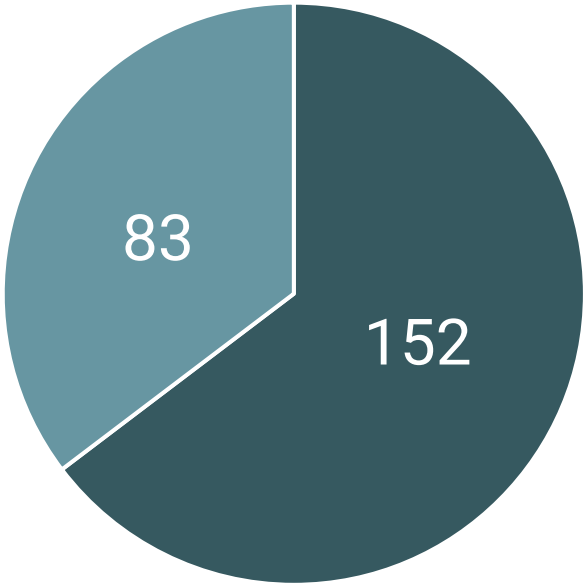


The point estimate represents an average from included studies of the proportion of patients affected by the condition of interest.



The distribution of estimates (in other words, how much they are dispersed around the average pooled estimate) is just as important as the point estimate.

235 systematic reviews of prevalence



Did not conduct meta-analysis

Conducted meta-analysis

In 144 (94.7%) reviews, authors reported that heterogeneity was assessed with I^2 statistics.

The I^2 value was reported in 134 reviews.

Median I^2: 96.9% (IQR 90.5 to 98.7)

125 (93.3%) presented $I^2 \geq 70\%$

104 (77.6%) presented $I^2 \geq 90\%$

This is not a common finding in meta-analyses for other types of data; for binary outcomes:

→ Patsopoulos et al., 2008:
- n = 1,011
- Median 21.1% (IQR 0.0 to 49.7)

→ Garcia-Alamino et al., 2017:
- n = 137
- Average 20.4% (SD 26.1)

Borges Migliavaca et al., 2020. doi: 10.1186/s12874-020-00975-3
Patsopoulos et al., 2008. doi: 10.1093/ije/dyn065
Garcia-Alamino et al., 2017. doi: 10.1136/bmjopen-2017-015888

	Number of systematic reviews	$I^2 \leq 50\%$, n (%)	$I^2 > 50\%$, n (%)	p-value
TOTAL	134	7 (5.2%)	127 (94.8%)	-
Number of studies included in the meta-analysis [a]				0.015 [c]
02 to 09	33	5 (15.2%)	28 (84.8%)	
10 to 18	31	2 (6.4%)	29 (93.6%)	
19 to 26	34	0	34 (100%)	
27 or more	36	0	36 (100%)	
Pooled estimate				0.015 [d]
Between 10% and 90%	98	2 (2.0%)	96 (98.0%)	
< 10% or > 90% (extremes)	36	5 (13.9%)	31 (86.1%)	
Transformation method				0.686 [c]
Freeman-Tukey	32	1 (3.1%)	31 (96.9%)	
Others [b]	10	1 (10%)	9 (90%)	
Not reported	92	5 (5.4%)	87 (94.6%)	

[a] Division by quartiles, approximately. [b] Other transformation methods included logit, log, arcsine and no transformation (raw). [c] Pearson's Chi-Squared Test for Homogeneity. [d] Fisher's Exact Test.

- Commonly, arbitrary thresholds were used to classify I^2 as high, moderate or low and the result of I^2 was used to justify the choice of statistical model (fixed effect or random effects).
- Subgroup analysis was conducted in 82 reviews (61.2%).
- Studies that reported high I^2 values were more likely to have conducted a sensitivity analysis.
- Only 3 (2.2%) meta-analyses estimated prediction intervals.

Discussion

- The I^2 informs the proportion of the variance in observed effect is due to variance in true effects rather than sampling error - **it does not directly inform us about the distribution of effects.**
- A high I^2 estimate it is not necessarily synonymous with important heterogeneity. In the same way, a low value of I^2 is not always an indicator of consistent and homogenous results.

Higgins et al., 2003. doi: 10.1136/bmj.327.7414.557
Borestein et al., 2017. doi: 10.1002/jrsm.1230

Discussion

- It is known that some variables may influence the estimation of I^2 - variance estimator, type of outcome (continuous, binary, proportional) and even the outcome itself.
    - In proportional data, small variance is observed even in studies with small sample size.
- In our analysis:
    - More often, meta-analyses with low value I^2 included few studies.
    - Meta-analyses with extreme pooled estimates (defined as < 10% or > 90%) more often presented a low value I^2.
        - The range of potential results is limited for conditions that are either very rare or very common.

Higgins et al., 2003. doi: 10.1136/bmj.327.7414.557
von Hippel, 2015. doi: 10.1186/s12874-015-0024-z

Discussion

- We hypothesize that authors who conduct sensitivity analysis aimed to explore the heterogeneity supposedly identified through I^2 statistic.
  - Increasing the number of analyses may lead to spurious results due to chance only.
  - Since a lower number of included studies is associated with lower I^2 values, subgroup analysis may result in subgroups with low I^2, which could be interpreted as a solution for the heterogeneity but that is not necessarily the case.
- Prediction intervals inform us the range of expected estimates - precisely the question of interest when discussing heterogeneity. However, as we observed, the estimation of prediction intervals is still underused in meta-analysis of prevalence.

Take-home messages

- Meta-analysis of prevalence commonly yields high I^2 estimates, and authors conclude their results are heterogeneous.
- However, the I^2 statistic is not an absolute index for the amount of variability observed and its estimation can be impacted by some factors such as the number of studies or the pooled result.
- When discussing heterogeneity, reviewers should focus on the description of the expected range of estimates, which can be done using prediction intervals.
- In case of substantial heterogeneity, planned sensitivity analysis can help elucidate the factors associated with the variability among estimates.

I^2 statistic in meta-analysis of prevalence: worthwhile or worthless?

Celina Borges Migliavaca, PhD(c)

PERSyst (Prevalence Estimates Reviews – Systematic Review Methodology) **Group**

Cochrane Training

A day with... Statistical Methods Group (SMG)

UFRGS
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

@ celinabm7@gmail.com

celinaborgesm