A statistical note on Karl Pearson's 1904 meta-analysis

Harry Shannon

Department of Clinical Epidemiology & Biostatistics, McMaster University, Hamilton, ON, Canada L8S 4K1
Corresponding author: Harry Shannon. Email: shannonh@mcmaster.ca

Karl Pearson's 1904 report on *Certain enteric fever inoculation statistics* is seen as a key paper in the history of meta-analysis.[1–3] In it, Pearson raised several important methodological issues arising from his correlations between typhoid and mortality and the inoculation status of soldiers serving in various parts of the British Empire.[4]

First, he noted the 'significance' of the individual correlations. For this he used the magnitude of the correlations in relation to their 'probable errors'. Second, he pointed out the 'extreme irregularity' of the correlation values – what we would now call heterogeneity – and sought to explain why they differed. Third, he commented on the 'lowness' of the values, arguing that they were too low to convince him that the inoculation had been proven worthwhile. He felt that a better vaccine was needed.

Pearson also commented on how the data had been obtained. He was concerned that self-selection into the inoculated group by volunteers who were 'more cautious and careful' could have produced spurious estimates of effectiveness. This and his concerns about the weakness of the correlations led him to recommend that an 'experiment' be done. He did not propose a randomised controlled trial – he was writing before Fisher developed the theoretical reasons for random allocation – but Pearson clearly understood the need for comparability of groups. His solution was to call for volunteers, register them all, and only inoculate every second one.

The data available to Pearson were presented in 2×2 tables. To create a measure of effect, he computed for each table the tetrachoric correlation, which he had described a few years earlier.[5] The approach assumes the data come from a bivariate normal distribution and derives the correlation based on that distribution.

Today we would use the data in the tables to find other measures, for example, the relative odds (odds ratios). Table 1 shows Pearson's values for the correlations, along with estimates of the relative odds.

Following Pearson, the results are presented separately for the relation between inoculation and escaping typhoid (enteric) fever, and the relation between inoculation and case survival. The rank orders of the correlations and odds ratios are the same for the first set of tables and almost identical for the second. What is striking is that, even when the odds ratio reached 7.9 (the relative risk for this table was 6.9), the correlation was only 0.445, which fell in the range (0.25–0.5) that Pearson labelled 'moderate'. (Pearson used as outcomes 'escaping' disease and 'survival given disease', so that protection resulting from inoculation is reflected in positive correlations and in odds ratios greater than 1. We are used to seeing the odds ratios presented so that values below 1 show a benefit of treatment. In this case, the inverse of 7.9 is 0.13.)

A formal test of heterogeneity for the odds ratios in the first set of tables confirms Pearson's observation (Breslow-Day $X^2 = 90.6$ on 4 df, $p < 0.001$). However, this is not so for the second set, for which the test is not conventionally statistically significant: $X^2 = 6.9$ on 5 df, $p = 0.23$. Given this, it is legitimate to compute a pooled odds ratio: the Mantel-Haenszel estimate is 1.77 (95% CI 1.5–2.1).

A final point: Pearson considered the effectiveness of inoculation in two steps – whether it prevented soldiers from acquiring typhoid fever and whether it reduced mortality in those who had developed the disease. For four of the groups, it is possible to explore directly the relationship between inoculation and mortality from the disease. The odds ratios are in the range of 2.2–6.8. They are not significantly different from each other – $X^2 = 5.2$ on 3 df, $p = 0.16$. The pooled estimate is 4.5 (95% CI 3.1–6.6). At face value, it is a strong effect (the inverse is 0.22, 95% CI 0.15–0.32) by current criteria. Even so, I suspect that Pearson would still not have been convinced of the value of vaccination but would have continued to insist that further work was needed, including a proper controlled trial.

Table 1. Tetrachoric correlations calculated by Pearson and relative odds for typhoid fever data.

Dataset	Correlation	Probable error	Relative odds	95% CI
Association between 'escaping' disease and inoculation				
I	+0.373	±0.021	3.1	1.9 – 4.8
II	+0.445	±0.017	7.9	5.6 – 11.0
III	+0.191	±0.026	2.3	1.5 – 3.5
IV	+0.021	±0.033	1.1	0.8 – 1.5
V	+0.100	±0.013	1.7	1.4 – 2.2
Overall estimate[a]	+0.226		N/A	
Association between case survival and inoculation				
VI	+0.307	±0.128	2.8	0.6 – 13.6
VII	− 0.010	±0.081	0.96	0.4 – 2.1
VIII	+0.300	±0.093	2.4	1.0 – 5.7
IX	+0.119	±0.022	1.5	1.2 – 1.9
X	+0.194	±0.022	2.0	1.5 – 2.6
XI	+0.248	±0.050	2.7	1.4 – 5.1
Overall estimate[a]	+0.193		1.77	1.5 – 2.1

[a]For correlations, the overall estimate is the arithmetic mean of the correlations, as given by Pearson. For the relative odds, it is the Mantel-Haenszel pooled estimate. N/A shows that the separate estimates were heterogeneous, and hence not pooled.

References

1. Hedges LV. Commentary on pooling the results of clinical trials. *Stat Med* 1987; 6: 381–385.
2. Chalmers I, Hedges LV and Cooper H. A brief history of research synthesis. *Eval Health Prof* 2002; 25: 12–37.
3. O'Rourke K. *An historical perspective on meta-analysis: dealing quantitatively with varying study results.* The James Lind Library. See www.jameslindlibrary.org (2006, last checked 23 June 2016).
4. Pearson K. Report on certain enteric fever inoculation statistics. *BMJ* 1904; 3: 1243–1246.
5. Pearson K. Mathematical contributions to the theory of evolution. VII. On the correlation of characters not quantitatively measurable. *Philos Trans R Soc London A* 1900; 195: 1–47.