The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*, *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
1-year subscription	$ 79	1-year subscription	$115
2-year subscription	$155	2-year subscription	$225
3-year subscription	$225	3-year subscription	$329
3-year subscription (electronic only)	$210	3-year subscription (electronic only)	$210
1-year student subscription	$ 48	1-year student subscription	$ 79
1-year university library subscription	$ 99	1-year university library subscription	$135
2-year university library subscription	$195	2-year university library subscription	$265
3-year university library subscription	$289	3-year university library subscription	$395
1-year institutional subscription	$225	1-year institutional subscription	$259
2-year institutional subscription	$445	2-year institutional subscription	$510
3-year institutional subscription	$650	3-year institutional subscription	$750

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The *Stata Journal*, electronic version (ISSN 1536-8734) is a publication of Stata Press. Stata, STATA, Stata Press, Mata, MATA, and NetCourse are registered trademarks of StataCorp LP.

The Stata Journal (2012)
12, Number 3, pp. 505–514

Exact and mid-p confidence intervals for the odds ratio

Morten W. Fagerland
Unit of Biostatistics and Epidemiology
Oslo University Hospital
Oslo, Norway
morten.fagerland@medisin.uio.no

Abstract. The odds ratio is a frequently used effect measure for two independent binomial proportions. Unfortunately, the confidence intervals that are available for it in Stata and other standard software packages are generally wider than necessary, particularly for small-sample and exact estimation. The performance of the Cornfield exact interval—the only widely available exact interval for the odds ratio—may be improved by incorporating a small modification attributed to Baptista and Pike (1977, *Journal of the Royal Statistical Society, Series C* 26: 214–220). A further improvement is achieved when the Baptista–Pike method is combined with the mid-p approach. In this article, I present the command `merci` (mid-p and exact odds-ratio confidence intervals) and its immediate version `mercii`, which calculate the Cornfield exact, Cornfield mid-p, Baptista–Pike exact, and Baptista–Pike mid-p confidence intervals for the odds ratio. I compare these intervals with three well-known logit intervals. I strongly recommend the Baptista–Pike mid-p interval.

Keywords: st0271, merci, mercii, confidence intervals, odds ratio, mid-p, exact, quasi-exact, Cornfield, Baptista–Pike

1 Introduction

The odds ratio is an important effect measure for two independent binomial proportions. It is routinely used in case–control studies, frequently used in cohort studies, and also used in clinical trials and as a summary measure in meta-analyses. Confidence intervals for the odds ratio are typically based on the approximate normal distribution of the logarithm of the estimate of the odds ratio. Such intervals are denoted logit intervals and sometimes include an adjustment to the observed counts of events and nonevents. Another frequently used interval is the Cornfield exact interval. A common disadvantage of these intervals is conservatism; they are generally wider than necessary, particularly for small sample sizes and when proportions are close to 0 or 1 (Fagerland, Lydersen, and Laake forthcoming).

An example is given in figure 1, where the coverage probabilities of the Cornfield exact and Woolf logit intervals are plotted for a sample size of 20 in each group and a fixed odds ratio of 3.0. The coverage probability is the probability that the confidence interval contains the true odds ratio. The preference is for it to be close to the nominal coverage probability, here 95%. A conservative interval is an interval with a too large coverage probability. Exact intervals are required always to be conservative, whereas approximate intervals satisfy no such criterion. As is plain from figure 1, the conservatism of the Cornfield exact interval can be quite severe.

The coverage probability is an important index of performance; an interval with large coverage probability is likely to be wider than an interval with lower coverage probability. Thus for the sample size and odds-ratio values in figure 1, the Woolf logit interval outperforms the Cornfield exact interval.

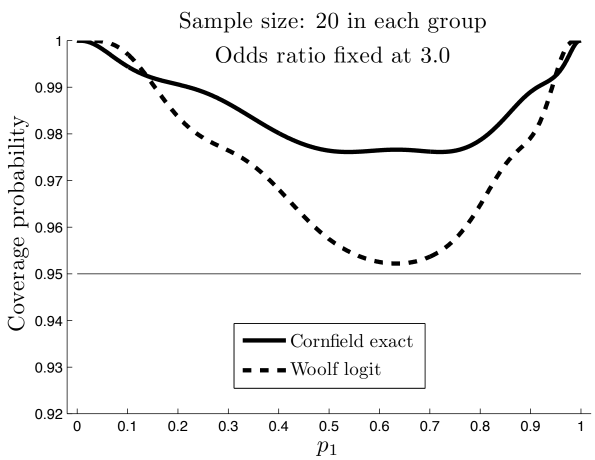

Figure 1. Coverage probabilities of the Cornfield exact and Woolf logit intervals; p_1 is the probability of event in group 1

In Stata, the main commands for producing confidence intervals for the odds ratio are `cc` and `cci`, which calculate the Cornfield exact (default), Woolf logit (option `woolf`), approximate Cornfield (option `cornfield`), and test-based (option `tb`) intervals. The approximate Cornfield interval does not always produce an upper confidence limit, even for straightforward values. For example, the command `cci 7 27 1 33, cornfield`, which corresponds to the results of the clinical trial by Perondi et al. (2004), returns the interval `(1.261, .)`. In fact, for 10,000 randomly selected counts of events and nonevents with random group sizes in the range 5–50 and excluding situations with 0 events or nonevents, 11% of the calculated approximate Cornfield intervals had a missing upper limit. The approximate Cornfield interval is thus unsuitable for general use. Test-based confidence intervals should only be used for pedagogical reasons, never for research work (see [ST] **epitab**).

Two more confidence intervals are available with the package `sbe30`: the Gart adjusted logit and the Agresti independence-smoothed logit intervals. These two intervals perform similarly to the Woolf logit interval, with Gart being slightly less conservative than the other two intervals (Fagerland, Lydersen, and Laake forthcoming).

The purpose of this article is to present three alternative confidence intervals for the odds ratio—the Baptista–Pike exact, Cornfield mid-p, and Baptista–Pike mid-p intervals—and their implementation in Stata through the command `merci` (mid-p and exact odds-ratio confidence intervals) and its immediate version `mercii`. I illustrate the performance of these intervals and show how they can improve interval estimation for the odds ratio.

2 Confidence intervals

Table 1 sets our notation for the observed counts of a 2×2 table. We denote the odds ratio by θ and use the sample proportions to estimate it:

$$\widehat{\theta} = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

Let α denote the nominal significance level, and let $z_{\alpha/2}$ denote the upper $\alpha/2$ percentile of the standard normal distribution.

Table 1. The observed counts of a 2×2 table

	Event	Nonevent	Sum
Group 1	n_{11}	n_{12}	n_{1+}
Group 2	n_{21}	n_{22}	n_{2+}
Sum	n_{+1}	n_{+2}	N

2.1 Logit intervals

Logit intervals are based on the approximate normal distribution of the logarithm of the estimate of the odds ratio, as first proposed by Woolf (1955). A confidence interval for θ is obtained by exponentiating the endpoints of

$$\log \widehat{\theta} \pm z_{\alpha/2} \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$$

To calculate the Gart adjusted logit interval (Gart 1966), we start by adding 0.5 to each cell count:

$$\widetilde{n}_{ij} = n_{ij} + 0.5, \quad i,j = 1,2 \quad \Rightarrow \quad \widetilde{\theta} = \frac{\widetilde{n}_{11}\widetilde{n}_{22}}{\widetilde{n}_{12}\widetilde{n}_{21}}$$

We then get a confidence interval for θ by exponentiating the endpoints of

$$\log \widetilde{\theta} \pm z_{\alpha/2} \sqrt{\frac{1}{\widetilde{n}_{11}} + \frac{1}{\widetilde{n}_{12}} + \frac{1}{\widetilde{n}_{21}} + \frac{1}{\widetilde{n}_{22}}}$$

The independence-smoothed logit interval by Agresti (1999) is obtained in a similar manner. Instead of adding 0.5 to all cell counts, we add the values

$$c_{ij} = 2n_{i+}n_{+j}/N^2, \quad i,j = 1,2$$

to cells n_{ij} and proceed as above.

2.2 Cornfield exact interval

Suppose that we condition on the number of events (n_{+1}) and the number of nonevents (n_{+2}) such that all marginal totals in table 1 are fixed. Any one table is then completely characterized by the count of one cell. Let x_{11} denote the number of events in group 1 for any table that might be observed given the fixed row and column sums. The probability of observing a table with x_{11} events follows the noncentral hypergeometric distribution (Cornfield 1956)

$$f(x_{11}|\theta) = \frac{\binom{n_{1+}}{x_{11}}\binom{n_{2+}}{n_{+1}-x_{11}}\theta^{x_{11}}}{\sum_{i=n_0}^{n_1}\binom{n_{1+}}{i}\binom{n_{2+}}{n_{+1}-i}\theta^{i}}$$

where $n_0 = \max(0, n_{+1} - n_{2+})$ and $n_1 = \min(n_{1+}, n_{+1})$. By inverting two one-sided Fisher exact tests, we obtain the Cornfield exact interval (L_C, U_C) by solving the following equations iteratively

$$\sum_{x_{11}=n_{11}}^{n_1} f(x_{11}|L_C) = \alpha/2 \tag{1}$$

and

$$\sum_{x_{11}=n_0}^{n_{11}} f(x_{11}|U_C) = \alpha/2 \tag{2}$$

The Cornfield exact interval is guaranteed to have coverage probabilities at least to the nominal level.

2.3 Baptista–Pike exact interval

Instead of inverting two one-sided tests, as in the Cornfield exact interval, Baptista and Pike (1977) invert one two-sided test using an acceptance region formed by ordered null probabilities. We make the following adjustments to (1) and (2)

$$\sum_{x_{11}=n_0}^{n_1} f(x_{11}|L_{\rm BP}) \times I\left\{f(x_{11}|L_{\rm BP}) \le f(n_{11}|L_{\rm BP})\right\} = \alpha \tag{3}$$

and

$$\sum_{x_{11}=n_0}^{n_1} f(x_{11}|U_{\rm BP}) \times I\left\{f(x_{11}|U_{\rm BP}) \le f(n_{11}|U_{\rm BP})\right\} = \alpha \tag{4}$$

where I is an indicator function, and n_0, n_1, and f are as defined in section 2.2. When (3) and (4) are solved iteratively such that $L_{\text{BP}} < U_{\text{BP}}$, the Baptista–Pike exact interval is given by $(L_{\text{BP}}, U_{\text{BP}})$. The Baptista–Pike exact interval is guaranteed to have coverage probabilities at least to the nominal level.

2.4 Cornfield mid-p interval

A mid-p value is calculated by subtracting half the point probability of the observed table from the ordinary p-value. The resulting mid-p test is no longer exact, and the corresponding mid-p interval can no longer guarantee coverage probabilities at least to the nominal level. To obtain the Cornfield mid-p interval, we substitute (1) and (2) with

$$\sum_{x_{11}=n_{11}}^{n_1} f(x_{11}|L_{C_m}) - \frac{1}{2} f(n_{11}|L_{C_m}) = \alpha/2$$

and

$$\sum_{x_{11}=n_0}^{n_{11}} f(x_{11}|U_{C_m}) - \frac{1}{2} f(n_{11}|U_{C_m}) = \alpha/2$$

The Cornfield mid-p interval is given by (L_{C_m}, U_{C_m}).

2.5 Baptista–Pike mid-p interval

The Baptista–Pike mid-p interval is obtained by adjusting (3) and (4) in the following manner

$$\sum_{x_{11}=n_0}^{n_1} f(x_{11}|L_{\text{BP}_m}) \times I\left\{f(x_{11}|L_{\text{BP}_m}) \leq f(n_{11}|L_{\text{BP}_m})\right\} - \frac{1}{2} f(n_{11}|L_{\text{BP}_m}) = \alpha$$

and

$$\sum_{x_{11}=n_0}^{n_1} f(x_{11}|U_{\text{BP}_m}) \times I\left\{f(x_{11}|U_{\text{BP}_m}) \leq f(n_{11}|U_{\text{BP}_m})\right\} - \frac{1}{2} f(n_{11}|U_{\text{BP}_m}) = \alpha$$

The Baptista–Pike mid-p interval is given by $(L_{\text{BP}_m}, U_{\text{BP}_m})$.

3 Comparisons of intervals

The Cornfield exact interval is the default confidence interval for the odds ratio when using Stata's `cc` or `cci` commands. As illustrated in figure 1, the Cornfield exact interval can be very conservative. We further demonstrate this in figure 2, where the coverage probabilities of the Cornfield exact and Baptista–Pike exact intervals are plotted against p_1, the probability of event in group 1, for fixed values of the odds ratio. Although both intervals are rather conservative, the Baptista–Pike exact interval improves upon the Cornfield exact interval.

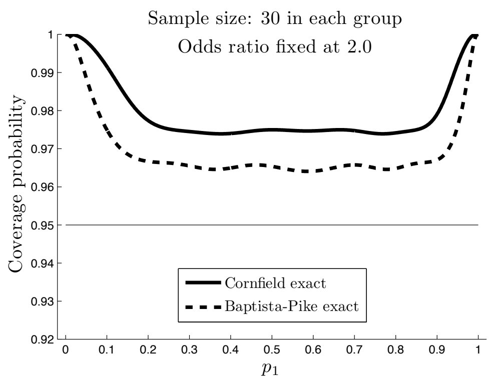

Figure 2. Coverage probabilities of two exact intervals

The coverage probabilities of the logit intervals defined in section 2.1 are shown in figure 3 for two combinations of sample sizes and fixed odds-ratio values. The three intervals perform similarly, and they have coverage probabilities considerably closer to the nominal level than the two exact intervals. The Gart adjusted logit interval is slightly less conservative (Fagerland, Lydersen, and Laake forthcoming) and slightly shorter (Agresti 1999) than the other two intervals.

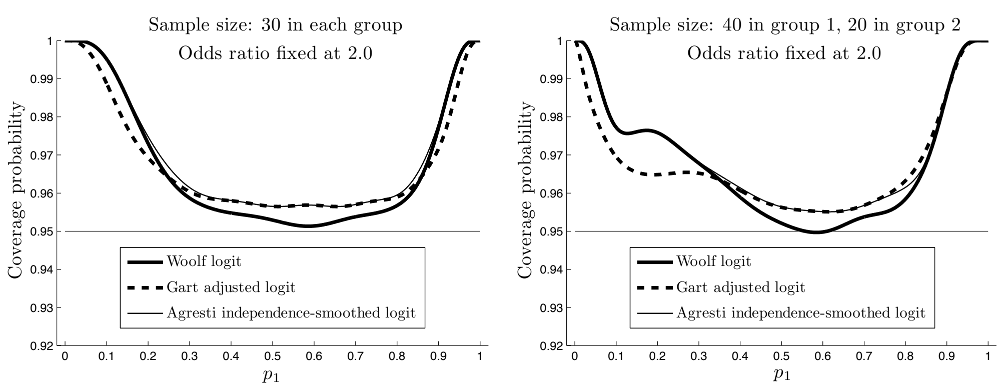

Figure 3. Coverage probabilities of three logit intervals

Figure 4 illustrates the performance of the Cornfield mid-p and Baptista–Pike mid-p intervals compared with the Gart adjusted logit interval. The Baptista–Pike mid-p interval is clearly superior to the other two intervals, particularly when proportions are close to 0 or 1. The coverage probability of the Baptista–Pike mid-p interval is sometimes below the nominal level, but the infringement is small and, for most practical purposes, inconsequential.

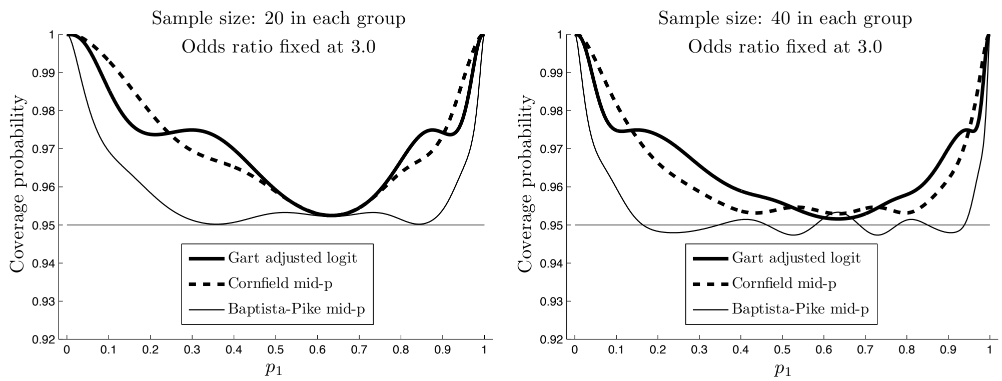

Figure 4. Coverage probabilities of the best-performing intervals

A more thorough comparison of the intervals defined in section 2 can be found in Fagerland, Lydersen, and Laake (forthcoming).

4 The merci and mercii commands

4.1 Syntax

`merci` *var_group var_event* [*if*] [*in*] [`, bp cornfield exact midp notable noccheaders level(`#`)`]

`mercii` #n_{11} #n_{12} #n_{21} #n_{22} [`, bp cornfield exact midp notable noccheaders level(`#`)`]

Dialog boxes for `merci` and `mercii` can be launched by typing `db merci` and `db mercii` at Stata's command line.

4.2 Options

`bp` requests that the Baptista–Pike method be used to calculate the confidence interval for the odds ratio. `bp` is the default option, and it will be used if the `cornfield` option is not specified.

`cornfield` requests that the Cornfield method be used to calculate the confidence interval for the odds ratio.

`exact` requests that the Baptista–Pike exact interval or the Cornfield exact interval be calculated, depending on whether the `cornfield` option is specified.

`midp` requests that the Baptista–Pike mid-p interval or the Cornfield mid-p interval be calculated, depending on whether the `cornfield` option is specified. `midp` is the default option, and it will be used if the `exact` option is not specified.

`notable` suppresses the 2×2 table from output.

`noccheaders` requests that group 2/group 1 and event/nonevent be used for the 2×2 table headers instead of the case–control specific headers cases/controls and exposed/unexposed.

`level(#)` specifies the confidence level, as a percentage, for the confidence interval. The default is `level(95)` or as set by `set level`.

4.3 Saved results

`merci` and `mercii` save the following in `r()`:

Scalars

`r(or)`	odds ratio
`r(lb_or)`	lower bound of confidence interval
`r(ub_or)`	upper bound of confidence interval

Macros

`r(method)`	confidence interval method

5 Examples

In section 1, I briefly mentioned the randomized clinical trial by Perondi et al. (2004) and that the approximate Cornfield interval fails to produce an upper confidence limit for it. In that trial, 68 children with cardiac arrest were randomized to standard ($n = 34$) or high dose ($n = 34$) epinephrine. The primary outcome measure was survival after 24 hours, and the results are summarized in table 2. The estimate of the odds ratio is $\widehat{\theta} = 8.56$, and the Cornfield exact interval, which was reported in Perondi et al. (2004), is $(0.97, 397)$. The results suggest a reduced survival with a high dose of epinephrine, but the confidence interval is very wide and includes the null value ($\theta = 1.0$).

Table 2. The results of a clinical trial of high versus standard dose of epinephrine in children with cardiac arrest

	Survival at 24h		
Treatment	Yes	No	Sum
Standard dose	7	27	34
High dose	1	33	34
Sum	8	60	68

Using `mercii`, we calculate the Baptista–Pike mid-p interval for the data in table 2. Because this was a clinical trial and not a case–control study, we use the general table headers by specifying the `noccheaders` option:

```
. mercii 7 27 1 33, noccheaders
                                                 Proportion
              |   Event   Non-event  |    Total       Event
--------------+----------------------+-----------------------
      Group 2 |       7          27  |       34       0.206
      Group 1 |       1          33  |       34       0.029
--------------+----------------------+-----------------------
        Total |       8          60  |       68       0.118
 Odds ratio estimate = 8.556
 95% Conf. interval = (1.328, 98.838) [Baptista-Pike mid-p]
```

The Baptista–Pike mid-p interval is considerably shorter than the Cornfield exact interval and does not contain $\theta = 1.0$. That result is consistent with the results from the recommended confidence intervals methods for the difference between proportions and the ratio of proportions in Fagerland, Lydersen, and Laake (forthcoming) and the recommended tests for association in Lydersen, Fagerland, and Laake (2009).

If it is required to use an exact interval, the Baptista–Pike exact interval is less conservative than the Cornfield exact interval:

```
. mercii 7 27 1 33, exact notable
 Odds ratio estimate = 8.556
 95% Conf. interval = (1.000, 195.495) [Baptista-Pike exact]
```

6 Discussion

Confidence interval estimation of the odds ratio can be greatly improved by using the Baptista–Pike method. For exact estimation, the Baptista–Pike exact interval is considerably less conservative—and thereby shorter—than the Cornfield exact interval. However, the best performing interval is the Baptista–Pike mid-p interval. It is superior to both exact and logit intervals, and it works well for small as well as large sample sizes. The Baptista–Pike mid-p interval was recommended in Fagerland, Lydersen, and Laake (forthcoming) but has not yet been available in any standard software package.

In this article, I presented the new Stata commands `merci` and `mercii`, which calculate the Cornfield exact, Cornfield mid-p, Baptista–Pike exact, and Baptista–Pike mid-p confidence intervals. The Cornfield exact interval is also available with Stata's `cc` or `cci` commands. The results from `merci`/`mercii` and `cc`/`cci` will be similar but not always identical. The confidence limits need to be calculated by an iterative algorithm, and the implementations may differ in certain aspects.

The only cases for which the two commands produce noteworthy different limits are when the 2×2 table includes one or two cell entries of 0. For example, `cci 0 10 5 5` will produce the interval $(0, 0.491)$, whereas `mercii 0 10 5 5, exact cornfield` will produce the interval $(0, 0.837)$.

For this table and for all other tables that I have discovered to give different results from `merci/mercii` and `cc/cci`, the results from `merci/mercii` are consistent with the results from StatXact® 9 (Cytel Inc.).

In conclusion, I strongly recommend the Baptista–Pike mid-p interval for the odds ratio.

7 References

Agresti, A. 1999. On logit confidence intervals for the odds ratio with small samples. *Biometrics* 55: 597–602.

Baptista, J., and M. C. Pike. 1977. Algorithm AS 115: Exact two-sided confidence limits for the odds ratio in a 2×2 table. *Journal of the Royal Statistical Society, Series C* 26: 214–220.

Cornfield, J. 1956. A statistical problem arising from retrospective studies. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, ed. J. Neyman, 135–148. Berkeley, CA: University of California Press.

Fagerland, M. W., S. Lydersen, and P. Laake. Forthcoming. Recommended confidence intervals for two independent binomial proportions. *Statistical Methods in Medical Research*.

Gart, J. J. 1966. Alternative analyses of contingency tables. *Journal of the Royal Statistical Society, Series B* 28: 164–179.

Lydersen, S., M. W. Fagerland, and P. Laake. 2009. Recommended tests for association in 2×2 tables. *Statistics in Medicine* 28: 1159–1175.

Perondi, M. B. M., A. G. Reis, E. F. Paiva, V. M. Nadkarni, and R. A. Berg. 2004. A comparison of high-dose and standard-dose epinephrine in children with cardiac arrest. *New England Journal of Medicine* 350: 1722–1730.

Woolf, B. 1955. On estimating the relation between blood group and disease. *Annals of Human Genetics* 19: 251–253.

About the author

Morten W. Fagerland is a senior researcher in biostatistics at Oslo University Hospital. His research interests include the application of statistical methods in medical research, analysis of categorical data and contingency tables, and comparisons of statistical methods using Monte Carlo simulations.