



Technical Guide

Confidence Intervals

Contents

Introduction	2
Software options	2
Directory of methods	3
Appendix 1 – Byar's method	6
Appendix 2 – χ^2 exact method	7
Appendix 3 – Wilson Score method	8
Appendix 4 – Dobson's method	9
Appendix 5 – Confidence intervals for odds and odds ratios	10
Appendix 6 – Confidence intervals for a regression slope	12
Appendix 7 – Simulated confidence intervals for the concentration index	13
PHE Technical Guides	14
References	14

Introduction

Confidence intervals are calculated around many different types of statistic used in public health analysis. The background to the use of confidence intervals and the reasons for using them are explained in [APHO Technical Briefing 3 – Commonly used public health statistics and their confidence intervals](#).¹

Different methods must be used for different types of statistic, reflecting the different statistical distributions underpinning the statistics, particularly where counts or sample sizes are small (approximate methods are often available for large samples, but it is rarely necessary to use them). Many of the methods are described in various PHE or PHO publications and tools; the purpose of this guide is to act as a directory, linking to existing resources where available. Where the documentation is not available in other PHE resources, the details are given in appendices to this guide.

Software options

Almost all the methods listed or described below can be implemented in a variety of different software environments. It is possible to implement every one of these in *R*, Stata, Excel and other statistical packages. Most can be implemented in SQL. However, some are easier than others. There are three essential types of methodology:

- 1 Byar's method, which is purely arithmetical, and can be implemented easily, anywhere: it is particularly useful for applications in SQL, which does not have statistical lookup tables built in (although it is possible to use *R* scripts within SQL Server 2016/2017).
- 2 Simulation (repeated resampling) methods. These can be easily implemented in statistical packages such as Stata and *R*, and also in Excel using VBA scripts, although this tends to be slower to run.

All other methods, which do not involve simulation, but do involve lookups of statistical distributions, eg normal, *t*, Poisson, χ^2 . All these methods can easily be implemented in Excel, and several of them are, as referred to below, in the [PHE Tool for calculating common public health statistics and their confidence intervals](#)¹ and in the PHEIndicatorMethods *R* package.² They can also be implemented in Stata, SAS, StatsDirect, or other statistical packages.

Directory of methods

Counts

Byar's method (Appendix 1) when the numerator (count) is at least 10. When the numerator is less than 10, use the exact χ^2 method (Appendix 2). For small numerators, Byar's method, which is an approximation to the Poisson distribution, can be less accurate and an exact method based on the Poisson distribution can be used. For 95% confidence intervals, Byar's method is within 0.2% of the exact value for numerators of 10 or more. For 99.8% confidence intervals it is within 1.5% of the exact value for numerators of at least 10, but it always errs on the conservative side, ie confidence limits are slightly wider than the exact ones.

This applies to counts that are not constrained to a theoretical maximum, ie they are not numerators of proportions. For the numerator of a proportion, use the Wilson Score method.

The method is available in Excel in the [PHE Tool for calculating common public health statistics and their confidence intervals](#).¹ Note that in the PHE tool there is no separate worksheet for counts but the sheet for rates can be used.

It is also available in the PHEIndicatorMethods R package.²

Proportions (often presented as percentages)

Wilson Score method (Appendix 3).

The method is available in Excel in the [PHE Tool for calculating common public health statistics and their confidence intervals](#)¹ and in the PHEIndicatorMethods R package.²

Crude rates

Byar's method (Appendix 1) when the numerator (count) is at least 10. When the numerator is less than 10, use the exact χ^2 method (Appendix 2).

The method is available in Excel in the [PHE Tool for calculating common public health statistics and their confidence intervals](#)¹ and in the PHEIndicatorMethods R package.²

Indirectly standardised rates and ratios

Byar's method (Appendix 1) when the numerator (observed count) is at least 10. When the numerator is less than 10, use the exact χ^2 method (Appendix 2).

The method is available in Excel in the [PHE Tool for calculating common public health statistics and their confidence intervals](#)¹ and in the PHEIndicatorMethods R package.²

Directly standardised rates

There are two methods used within PHE: Dobson's method (with Byar's method) and Tiwari's³ modified Gamma method. They give extremely similar results. The Dobson method is used in the Health Intelligence Division of PHE and by NHS Digital. The Tiwari method is used by the Cancer Analysis Team in PHE. Work commissioned in 2017 by PHE has shown that there is little to choose between the two methods but for very small numbers

(counts between 10 and 25) Dobson's method (with Byar's method) gives very slightly more accurate coverage (ie for 95% confidence intervals, the intervals are slightly closer to the stated 95% coverage). For consistency, it is recommended that Dobson's method is used for any new indicators, but the Tiwari method gives confidence intervals that are sufficiently accurate.

Dobson's method is described in Appendix 4, Byar's method in Appendix 1.

The method is available in Excel in the [PHE Tool for calculating common public health statistics and their confidence intervals](#)¹ and in the PHEIndicatorMethods R package.²

Odds (used for excess winter deaths and proportional analysis)

Wilson Score method (adapted).

Odds are directly related to proportions, so the standard method for proportions can be adapted (Appendix 5) and has been shown to give more accurate confidence intervals than alternatives.

Odds ratio

Logit limits.

See appendix 5.

Life expectancy

Chiang's and Silcocks' method.

Described by Eayres and Williams⁴ and implemented in the [PHE Life expectancy calculator](#).⁵

Regression slope

t-distribution.

See Appendix 6.

Slope index of inequality and relative index of inequality

The method used in PHE is one based on simulation. The confidence intervals for the individual points on which the SII is based are used to estimate the statistical distribution for each one, and then the SII is repeatedly calculated (say 100,000 times) based on random samples from those distributions of the underlying points. This gives a picture of the statistical distribution for the SII, from which confidence intervals can be estimated accurately.

The method is set out step-by-step and implemented by VBA scripts in the [PHE Slope index of inequality tool](#).⁶ R scripts are also available in PHE to carry out the calculations.

The slope index of inequality (SII) is a regression slope, and before the simulation method was developed the standard published method used for calculating confidence intervals around the SII in PHE publications was the *t*-distribution method for a regression slope. This method is still applied in the [PHE Inequalities Calculation Tool](#).⁷ However, because this

method takes no account of the stability of the individual points used to calculate the SII (ie the confidence intervals around the individual points) so it is unnecessarily crude: it tends to be overly conservative, but not always: on occasions it gives confidence intervals that are too narrow.

Concentration index

There is no distributional information about the concentration index, Gini coefficient or other Lorenz-curve based indicators but confidence intervals can be calculated for the concentration index by simulation in the same way as for the slope index of inequality, as set out in the previous section. Details of the application of this method to the concentration index are provided in Appendix 7.

Appendix 1 – Byar's method

A rate of events r is given by:

$$r = \frac{O}{n}$$

where:

O is the numerator number of observed events;

n is the denominator population-years at risk.

The $100(1-\alpha)\%$ confidence limits for the rate r are given by:

$$r_{lower} = \frac{O_{lower}}{n}$$

$$r_{upper} = \frac{O_{upper}}{n}$$

where:

O_{lower} and O_{upper} are the lower and upper confidence limits for the observed number of events.

Using Byar's method⁸ the $100(1-\alpha)\%$ confidence limits for the observed number of events are given by:

$$O_{lower} = O \times \left(1 - \frac{1}{9O} - \frac{z}{3\sqrt{O}} \right)^3$$
$$O_{upper} = (O+1) \times \left(1 - \frac{1}{9(O+1)} + \frac{z}{3\sqrt{(O+1)}} \right)^3$$

where:

z is the $100(1-\alpha/2)$ th percentile value from the Standard Normal distribution.

For example, for a 95% confidence interval, $\alpha = 0.05$ and $z = 1.96$ (ie the 97.5th percentile value from the Standard Normal distribution).

For counts, the last pair of formulae are all that are required.

For indirectly standardised ratios, such as SMRs, the denominator n is not the population-years at risk, but is the expected number of events – otherwise the formulae are the same.

Appendix 2 – χ^2 exact method

Using the link between the Poisson and χ^2 distributions⁹ the equations for O_{lower} and O_{upper} in Appendix 1 can be replaced by:

$$O_{\text{lower}} = \frac{\chi^2_{\text{lower}}}{2}$$

$$O_{\text{upper}} = \frac{\chi^2_{\text{upper}}}{2}$$

where:

χ^2_{lower} is the 100(1- α /2)th percentile value from the χ^2 distribution with 2O degrees of freedom;

χ^2_{upper} is the 100(α /2)th percentile value from the χ^2 distribution with 2O+2 degrees of freedom.

Appendix 3 – Wilson Score method

The proportion p is given by:

$$p = \frac{O}{n}$$

where:

O is the numerator observed number of individuals in the sample/population having the specified characteristics;

n is the denominator total number of individuals in the sample/population.

Using the Wilson Score method^{10,11} the $100(1-\alpha)\%$ confidence limits for the proportion p are given by:

$$p_{lower} = \frac{(2O + z^2 - z\sqrt{z^2 + 4Oq})}{2(n + z^2)}$$

$$p_{upper} = \frac{(2O + z^2 + z\sqrt{z^2 + 4Oq})}{2(n + z^2)}$$

where:

q is $1-p$;

z is the $100(1-\alpha/2)$ th percentile value from the Standard Normal distribution.

For example, for a 95% confidence interval, $\alpha = 0.05$ and $z = 1.96$ (i.e. the 97.5th percentile value from the Standard Normal distribution).

Appendix 4 – Dobson’s method

The directly standardised rate (DSR) is given by:

$$DSR = \frac{1}{\sum_I w_i} \times \sum_I \frac{w_i O_i}{n_i}$$

where:

O_i is the observed number of events in the local or subject population in age group i ;

n_i is the number of individuals in the local or subject denominator population in age group i , or the population \times period at risk (e.g. 'person-years');

w_i is the number (or proportion) of individuals in the reference or standard population in age group i .

The $100(1-\alpha)\%$ confidence limits for the directly standardised rate DSR are given by:¹²

$$DSR_{lower} = DSR - \sqrt{\frac{Var(DSR)}{Var(O)}} \times (O_{lower} - O)$$
$$DSR_{upper} = DSR + \sqrt{\frac{Var(DSR)}{Var(O)}} \times (O_{upper} - O)$$

where:

O is the total observed count of events in the local or subject population;

O_{lower} and O_{upper} are the lower and upper confidence limits for the observed count of events, given by the formulae for Byar’s method in Appendix 1;

$Var(O)$ is the variance of the total observed count O ;

$Var(DSR)$ is the variance of the directly standardised rate.

The variances of the observed count O and the DSR are estimated by:

$$Var(O) = \sum_I O_i$$
$$Var(DSR) = \frac{1}{\left(\sum_I w_i\right)^2} \times \sum_I \frac{w_i^2 O_i}{n_i^2}$$

Appendix 5 – Confidence intervals for odds and odds ratios

Odds are directly related to proportions: the numerator is the same (the number of cases), but for odds, the denominator is the number of non-cases, whereas for proportions, the denominator is the total number of cases and non-cases.

	(Local) group of interest	(National) reference group
Cases	a	b
Non-cases	c	d

For proportions, we have local proportion = $\frac{a}{a+c}$ and reference proportion = $\frac{b}{b+d}$.

For odds, we have local odds = $\frac{a}{c}$ and reference odds = $\frac{b}{d}$.

The ratio of the proportions, or relative risk, is $\frac{a(b+d)}{b(a+c)}$.

The odds ratio, OR, is $\frac{ad}{bc}$.

The symmetric confidence interval for odds, $\frac{a}{c}$, can be derived from the Wilson Score method for proportions:

$$\text{lower limit} = \frac{2a + z^2 - z\sqrt{z^2 + \frac{4ac}{a+c}}}{2c + z^2 + z\sqrt{z^2 + \frac{4ac}{a+c}}}$$

$$\text{upper limit} = \frac{2a + z^2 + z\sqrt{z^2 + \frac{4ac}{a+c}}}{2c + z^2 - z\sqrt{z^2 + \frac{4ac}{a+c}}}$$

where z is the $100\left(1 - \frac{\alpha}{2}\right)$ th percentile value from the Standard Normal distribution (for 95% confidence intervals use $z = 1.9600$).

For odds ratios, logit limits¹³ are used:

$$\text{lower limit} = e^{\ln(OR) - z\sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}}$$

$$\text{upper limit} = e^{\ln(OR) + z\sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}}$$

Note that in the case where the odds ratio is used to compare observed and expected odds, we can often assume that the expected odds is fixed, with no variance, as we do for the observed count in an SMR, so the confidence interval is calculated around the observed odds (as above) and the resulting confidence limits each divided by the expected odds to give the confidence interval for the odds ratio.

The logit is a mathematical function defined as the log of the odds:

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$$

Where p is a proportion

Appendix 6 – Confidence intervals for a regression slope

When calculating a regression slope (gradient), we have n points, each with an x coordinate and a y coordinate.

The regression slope, b , is calculated as

$$b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

A $100(1 - \alpha)\%$ symmetric confidence interval around b is given by

$$\begin{aligned} \text{lower limit} &= b - t_{n-2, \alpha} \sqrt{\frac{n \sum y_i^2 - (\sum y_i)^2 - b^2(n \sum x_i^2 - (\sum x_i)^2)}{(n-2)(n \sum x_i^2 - (\sum x_i)^2)}} \\ \text{upper limit} &= b + t_{n-2, \alpha} \sqrt{\frac{n \sum y_i^2 - (\sum y_i)^2 - b^2(n \sum x_i^2 - (\sum x_i)^2)}{(n-2)(n \sum x_i^2 - (\sum x_i)^2)}} \end{aligned}$$

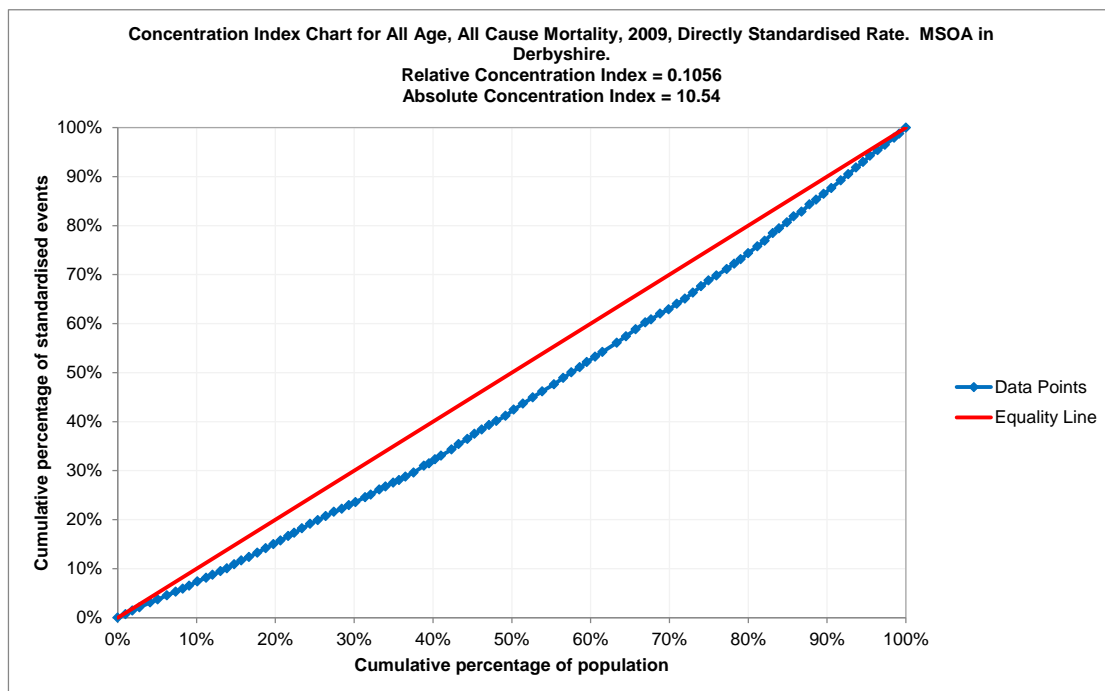
where $t_{n-2, \alpha}$ is the $\alpha\%$ quantile of Student's t -distribution with $n - 2$ degrees of freedom and all summations are from $i = 1$ to n .

Note that the method assumes that the points are independent of each other. If they are not independent observations (for example if the points are rolling average time series data – ie the periods overlap) then this method cannot be used to calculate a meaningful confidence interval for the slope.

Appendix 7 – Simulated confidence intervals for the concentration index

The concentration index is calculated from a Lorenz plot of the cumulative share of health in the population (on the y-axis; eg life expectancy) against the cumulative proportion of the population ranked by a socioeconomic variable (on the x-axis; eg Indices of Deprivation). If health is distributed perfectly equally across the socioeconomic groups, the Lorenz curve coincides with the diagonal – the equality line illustrated in the figure below. The further the curve is from the diagonal equality line, the greater the degree of inequality. The concentration index is the area between the Lorenz curve and the line of equality, measured and reported as a proportion of the total area beneath (or above) the line of equality.

The concentration index shows how unevenly health is distributed according to population share. It takes a value between zero and 1 (or 100%), where zero indicates perfect equality and 1 indicates 'ultimate inequality' (ie the hypothetical situation where all the 'good health' is in the least deprived group). However, in the context of health the range of attainable inequality will usually be much smaller.



Confidence intervals for the concentration index can be calculated using simulation either in an Excel VBA script or using statistical software.

Each value on the y-axis is derived from the indicator values for the individual areas. As long as we have some information about the underlying variability of those indicator values, ie from distributional assumptions or calculated variance or confidence intervals, it is possible to simulate other plausible values of the concentration index for each area by sampling repeatedly from the underlying values' distributions. Randomised values are generated, and concentration index values calculated a large number of times (eg 10,000, 100,000 or more) and the results from each repetition recorded, giving a distribution of

plausible values. The 2.5th and 97.5th centile values from that distribution can be used as estimated 95% confidence intervals for the concentration index.

PHE Technical Guides

This document forms part of a suite of PHE technical guides that are available on the Fingertips website: <https://fingertips.phe.org.uk/profile/guidance>

References

- ¹ Eayres D. *Technical Briefing 3: Commonly used public health statistics and their confidence intervals*. APHO; 2008. Available at [https://fingertips.phe.org.uk/documents/APHO Tech Briefing 3 Common PH Stats and CIs.pdf](https://fingertips.phe.org.uk/documents/APHO_Tech_Briefing_3_Common_PH_Stats_and_CIs.pdf). Tool to accompany briefing available at [https://fingertips.phe.org.uk/documents/PHE Tool for common PH Stats and CIs.xlsx](https://fingertips.phe.org.uk/documents/PHE_Tool_for_common_PH_Stats_and_CIs.xlsx)
- ² Anderson G. PHEindicatormethods: *Common Public Health Statistics and their Confidence Intervals*. R package version 1.0.0. Public Health England, 2018. To be made publicly available on CRAN or GitLab mid-2018.
- ³ Tiwari RC, Clegg LX, Zou Z. Efficient interval estimation for age-adjusted cancer rates. *Statistical Methods in Medical Research*, 2006;**15**:547–569.
- ⁴ Eayres DP, Williams ES. Evaluation of methodologies for small area life expectancy estimation. *J Epidemiol Community Health*, 2004;**58**:243-249.
- ⁵ PHE Life expectancy calculator. [https://fingertips.phe.org.uk/documents/PHE Life Expectancy Calculator.xlsm](https://fingertips.phe.org.uk/documents/PHE_Life_Expectancy_Calculator.xlsm)
- ⁶ PHE Slope Index of Inequality Tool. [https://fingertips.phe.org.uk/documents/PHE Slope Index of Inequality Tool \(Simulated CIs\).xlsm](https://fingertips.phe.org.uk/documents/PHE_Slope_Index_of_Inequality_Tool_(Simulated_CIs).xlsm)
- ⁷ PHE Inequalities Calculation Tool. [https://fingertips.phe.org.uk/documents/PHE Inequalities Calculation Tool.xls](https://fingertips.phe.org.uk/documents/PHE_Inequalities_Calculation_Tool.xls)
- ⁸ Breslow NE, Day NE. *Statistical methods in cancer research, volume II: The design and analysis of cohort studies*. Lyon: International Agency for Research on Cancer, World Health Organisation; 1987.
- ⁹ Armitage P, Berry G. *Statistical Methods in Medical Research*. 3rd edition. Oxford, Blackwell Science Ltd, 1994.
- ¹⁰ Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 1927; 22: 209–12.
- ¹¹ Newcombe RG, Altman DG. Proportions and their differences. In Altman DG et al. (eds). *Statistics with confidence*. 2nd edition. London: BMJ Books; 2000: 46–8.
- ¹² Dobson AJ, Kuulasmaa K, Eberle E, Scherer J. Confidence intervals for weighted sums of Poisson parameters. *Statistics in Medicine*, 1991;**10**:457-462.
- ¹³ Armitage P, Berry G. *Statistical Methods in Medical Research*. 3rd edition. Oxford, Blackwell Science Ltd, 1994:131-132.