

ORIGINAL ARTICLE

Simple nomograms to calculate sample size in diagnostic studies

S Carley, S Dosman, S R Jones, M Harrison



For a larger version of
fig 1 see the EMJ website

See end of article for
authors' affiliations

Correspondence to:
Dr S Carley, Manchester
Royal Infirmary,
Manchester M13 9WL,
UK; simon.carley@
mmc.nhs.uk

Accepted
11 December 2003

Emerg Med J 2005;22:180–181. doi: 10.1136/emj.2003.011148

Objectives: To produce an easily understood and accessible tool for use by researchers in diagnostic studies. Diagnostic studies should have sample size calculations performed, but in practice, they are performed infrequently. This may be due to a reluctance on the part of researchers to use mathematical formulae.

Methods: Using a spreadsheet, we derived nomograms for calculating the number of patients required to determine the precision of a test's sensitivity or specificity.

Results: The nomograms could be easily used to determine the sensitivity and specificity of a test.

Conclusions: In addition to being easy to use, the nomogram allows deduction of a missing parameter (number of patients, confidence intervals, prevalence, or sensitivity/specificity) if the other three are known. The nomogram can also be used retrospectively by the reader of published research as a rough estimating tool for sample size calculations.

In a previous paper,¹ we described a method of calculating sample size in diagnostic tests by determining the precision of the expected sensitivity and specificity. However, our experience suggests that many colleagues are reluctant to use the mathematical formula we described. Nomograms have been used in trials of therapy to aid calculation and understanding. The nomogram designed by Gore² for trials of therapy is an example of an easily understood and accessible tool that can be used by reader and researcher alike.

Sample size estimation in diagnostic tests may take two forms.⁴ Firstly, the number of subjects needed to test the hypothesis that a particular parameter will exceed a pre-determined level can be estimated (that is, is the sensitivity of the new test within 10% of the reference test?). This formal statistical approach is used when the researcher needs to specify equivalence or the difference between two tests.

Secondly, the number of subjects needed to define an expected level of sensitivity and specificity together with the precision of that estimate (that is, the confidence intervals) can be calculated. The mathematical approach differs depending upon whether the researcher is estimating precision (the second method) or testing a hypothesis (the first method). The methods described here are taken from the work by Buderer and are based on the researcher estimating the number of patients required to determine the precision of the result.³

We sought to develop a similar nomogram to estimate sample size in diagnostic trials based upon determination of sample size precision. We present nomograms for the prospective calculation of sample size in studies evaluating single diagnostic tests.

METHODS

Derivation of the nomogram

The calculations by Buderer³ were used to plot the nomogram. These calculations were described in detail in our

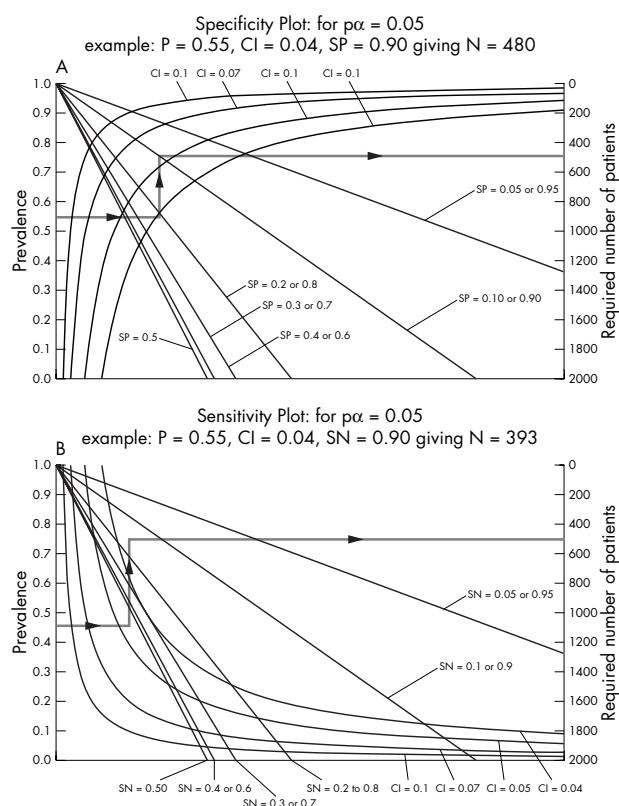


Figure 1 (A) Specificity plot for $\alpha=0.05$. Example: $P=0.55$, $CI=0.04$, $SP=0.90$ giving $n=480$. (B) Sensitivity plot for $\alpha=0.05$. Example: $P=0.55$, $CI=0.04$, $SN=0.90$ giving $n=393$. A larger version of fig 1 is available on the EMJ website (<http://emjonline.com/supplemental>).

previous paper.¹ In order to simplify the nomogram, and by convention, we fixed the probability of finding a false positive result at 5% or less. This figure is taken at the conventional level from therapeutic studies; no agreed convention currently exists for diagnostic tests. The nomograms were created using a Microsoft Excel spreadsheet.

This method should only be used if more than five subjects have each possible outcome (true positive, false positive, true negative, false negative).

RESULTS

Use of the nomograms

Fig 1 shows nomograms for the calculation of sensitivity (A) and specificity (B). For both nomograms the following method is used.

1. Use a straight edge to draw a horizontal line from the estimated prevalence to the required confidence interval (for example you estimate that a test has 80% sensitivity; you want to have a fairly precise estimate and so set your confidence intervals for 5%).
2. Draw a line from that intersection vertically until it meets the expected sensitivity or specificity.
3. Draw a horizontal line from that intersection to the right hand axis where the number of patients required may be estimated.

An example of how to use the nomograms is shown as a grey, arrowed line. Simple movement of the rule between different values on each axis can illustrate how a change in

any single parameter may influence the precision of the results.

DISCUSSION

The nomogram has four elements to it (number of patients, confidence intervals, prevalence and sensitivity/specificity). If the researcher knows any three of these, they will be able to estimate the fourth. This allows the nomogram to be used retrospectively by the reader of published research as a rough estimating tool for sample size calculations. However, we must again stress the importance of the a priori sample size calculation when planning a study.

Authors' affiliations

S Carley, Manchester Royal Infirmary, Manchester, UK

S Dosman, Cambridge, UK

S R Jones, Hope Hospital, Salford, Manchester, UK

M Harrison, North Staffordshire Hospital, UK

Competing interests: none declared

REFERENCES

- 1 **Jones S**, Carley S, Harrison M. An introduction to power and sample size estimation. *Emerg Med J* 2003;**20**:453–8.
- 2 **Gore SM ADG**. How large a sample. In: Gore SM ADG, eds. *Statistics in practice*. London: BMJ Publishing, 2001.
- 3 **Buderer NM**. Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Acad Emerg Med* 1996;**3**:895–900.
- 4 **Arkin CF**, Wachtel MS. How many patients are necessary to assess test performance. *JAMA* 1990;**263**:275–8.