


## Do the findings of case series studies vary significantly according to methodological characteristics?

K Dalziel,<sup>1</sup> A Round,<sup>1</sup> K Stein,<sup>1\*</sup> R Garside,<sup>1</sup>  
E Castelnuovo<sup>1</sup> and L Payne<sup>2</sup>

 Peninsula Technology Assessment Group, Peninsula Medical School,  
Universities of Exeter and Plymouth, Exeter, UK

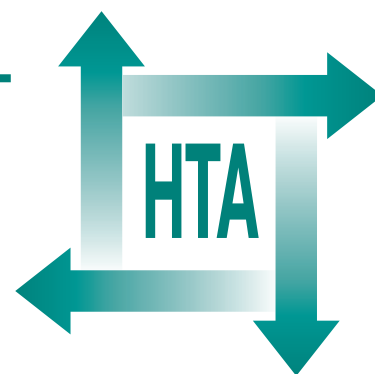
<sup>2</sup> Wessex Institute for Health, University of Southampton, UK

\* Corresponding author

### *Executive summary*

*Health Technology Assessment 2005; Vol. 9: No. 2*

Health Technology Assessment  
NHS R&D HTA Programme





## Executive summary

### Objectives

This study had three aims:

1. To review the use of case series in National Institute for Clinical Excellence (NICE) Health Technology Assessment (HTA) reports.
2. To review systematically the methodological literature for papers relating to the validity of aspects of case series design.
3. To investigate characteristics and findings of case series using examples from the UK's Health Technology Assessment programme.

### Background

Although randomised controlled trials (RCTs) offer the most robust evidence for effectiveness, this level of data is not always available for health technology assessments. Given that policy decisions still need to be made even in the absence of RCT evidence, it is important to try to understand the elements of case series design that determine their quality. Although a simple hierarchy of evidence will place case series as a weak form as evidence, individual studies, just like individual RCTs, may vary widely in quality and different studies of the same intervention may produce widely different estimates of outcome frequency. The validity of any study, whatever its form, will depend on the quality of its design, execution and interpretation. Nevertheless, case series studies are the most vulnerable to bias and confounding. RCTs attempt to minimise challenges to internal validity through minimising selection, performance, detection and attrition biases. However, this may lead to problems of external validity if strict exclusion criteria lead to a population being assessed, which is very different to that treated in clinical practice.

The aspects of quality that influence the validity of RCTs have been empirically studied, and it is generally agreed that adequate blinding, concealment and randomisation methods are crucial. A number of different scales and checklists for quality exist, but not all of them are empirically based or rigorously developed. As the authors were not aware of agreed aspects of quality for case series that were important, this study

aimed to look at what types of quality measure had been used in NICE HTAs and to search the literature systematically to see if empirical studies of case series had been published.

While comparisons of the results from RCTs and other study designs have been undertaken, they have been restricted to observational studies with control groups. These yield conflicting results, with non-randomised studies variously showing greater treatment effects, similar treatment effects and lower treatment effects in different subject areas investigated. The evidence suggests that non-randomised controlled evidence shows more variance than RCTs and the direction of effect is unpredictable. As we were not aware of such investigations of case series and RCT results, we aimed to investigate this.

### Review of the use of case series in NICE HTAs

Currently completed NICE HTAs were obtained from the NICE website. Of the 47 completed HTAs, 14 (30%) had included information from case series studies.

In two cases no RCTs were identified and the other 12 reports also included data from between two and 70 RCTs. The number of case series included ranged from two to 159. Inclusion criteria for case series included study size and length of follow-up. Various quality criteria were applied ( $n = 9$ ), with the CRD Report criteria being used in three cases. Data from case series were used to confirm RCT results, to inform an economic model, to explore variation and for meta-analysis.

We found that there was no consensus on which case series to include in HTAs, how to use them or how to assess their quality, despite them being used in 30% of NICE HTAs.

### Systematic review of methodological literature

We carried out searches in electronic databases, handsearched journals and examined the



bibliographies of papers in order to find studies that assessed aspects of case series design, analysis or quality in relation to study validity. No empirical studies were found. However, it is known that searches that are sensitive enough to identify case series are difficult to design with appropriate specificity and it is possible that we failed to locate such studies.

## Investigation of characteristics and findings of case series studies

A number of hypotheses relating to the design of case series studies were developed *a priori*. These were empirically investigated using four case examples from existing reports produced as part of the UK's HTA programme.

We included HTAs that had at least 40 case series studies available, included at least one good-quality RCT and contained information on the age of participants as a minimum description of the included population. We identified three reports on four topics – functional endoscopic sinus surgery for nasal polyps, spinal cord stimulation for chronic back pain, percutaneous transluminal coronary angioplasty (PTCA) and coronary artery bypass grafting (CABG) for chronic angina.

Data were extracted on outcome measures and study population characteristics.

Analysis was undertaken on a between-study level within each review. For each hypothesis, continuous variable data were explored through scatter plots and robust regression. Regression analysis weighted by sample size were also performed. Binary data were explored through *t*-tests and Mann–Whitney tests. Analysis of variance (ANOVA) was also performed, weighted for sample size. Multivariate analysis using disease severity, age and male sex was performed using multivariate robust regression or ANOVA as appropriate.

Comparisons between cases series and RCTs were performed using the intervention arms of RCTs as a comparator. There were only enough data to do this for PTCA and CABG. Meta-analysis of RCT data was compared with weighted robust regressions using the intervention as the confounding factor and estimating the coefficient size.

Poor reporting of case series characteristics severely constrained analysis and there were

insufficient data to investigate all the hypotheses. Findings were not consistent across the different topics and were subject to considerable uncertainty.

No relationship was found between sample size and outcome frequency. No relationship was found between prospective data collection and outcome frequency. One analysis each (in different topic areas) found a significant association between multi-centre studies and outcome, between independent outcome measurement and outcome frequency and between earlier publication and outcome frequency. Length of follow-up was found to be significantly associated with outcome frequency in three analyses. One topic area had scored case series for quality and this was found to be associated with outcome. However, this quality score contained items which we investigated separately in this review, without evidence of impact.

Compared with RCT evidence, which showed no difference between PTCA and CABG, case series estimates of mortality showed a 1–2% increase in mortality for CABG. For angina recurrence, neither case series nor RCT data showed any difference between the two interventions.

## Limitations

We found no previous studies empirically investigating methodological characteristics of case series. However, it is possible that the search strategy failed to find relevant studies.

All the examples in our analysis were surgical interventions, which are prone to additional confounding factors owing to difficulties of standardisation compared with drug treatment. Our findings may not be generalisable outside the interventions studied.

The case series reports included generally exhibited poor reporting of methodological characteristics. This constrained our analysis.

The use of several methods of analysis has led to apparently discrepant results. Given the number of analyses performed, the usual level of significance ( $p = 0.05$ ) should be viewed with caution.

The most important limitation of our study is the small number of cases on which our findings are based. The results are therefore tentative and should be viewed with caution. ►

## Conclusions

Case series are incorporated in a significant proportion of health technology assessments.

A wide range of quality criteria have been used to appraise the quality of case series and decide on their inclusion in reviews of studies using this design. In this small series of case studies drawn from HTAs carried out for the NHS HTA programme, we found little evidence to support the use of many of the factors included in quality assessment tools. Importantly, we found no relationship between study size and outcome across the four examples studied.

Isolated examples of a potentially important relationship between other methodological factors and outcome were shown, such as blinding of outcome measurement, but these were not shown consistently across the small number of examples studied.

Comparison of case series and RCT data was possible in only two examples studied but demonstrated a greater range in outcomes reported in case series, reflecting the likelihood that this design includes different populations. However, outcomes were not better in case series, contrary to expectations.

Estimates of comparative efficacy of alternative techniques by comparing case series studies were shown to be different from analyses based on RCTs. However, it is not clear from

this whether this is an effect of confounding or indicates different efficacy in different populations.

This study is based on a very small sample of studies and should therefore be considered as exploratory. Further investigation of the relationship between methodological features and outcome is justified given the frequency of use of case series in health technology assessments.

## Need for further research

Further research into the methodological features of case series and their outcome is justified in a wider sample of technologies and larger sets of case series.

Value of information analyses including case series could be explored.

Further exploration of the differences between case series and RCT results, preferably using registry or comprehensive case series data, would be valuable.

## Publication

Dalziel K, Round A, Stein K, Garside R, Castelnuovo E, Payne L. Do the findings of case series studies vary significantly according to methodological characteristics? *Health Technol Assess* 2005;**9**(2).



**INAHTA**

### **How to obtain copies of this and other HTA Programme reports.**

An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (<http://www.hta.ac.uk>). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

- fax (with **credit card** or **official purchase order**)
- post (with **credit card** or **official purchase order** or **cheque**)
- phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

### **Contact details are as follows:**

HTA Despatch  
c/o Direct Mail Works Ltd  
4 Oakwood Business Centre  
Downley, HAVANT PO9 2NP, UK

Email: [orders@hta.ac.uk](mailto:orders@hta.ac.uk)  
Tel: 02392 492 000  
Fax: 02392 478 555  
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

### **Payment methods**

#### *Paying by cheque*

If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

#### *Paying by credit card*

The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

#### *Paying by official purchase order*

You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

### **How do I get a copy of HTA on CD?**

Please use the form on the HTA website ([www.hta.ac.uk/htacd.htm](http://www.hta.ac.uk/htacd.htm)). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

---

The website also provides information about the HTA Programme and lists the membership of the various committees.

# NHS R&D HTA Programme

The research findings from the NHS R&D Health Technology Assessment (HTA) Programme directly influence key decision-making bodies such as the National Institute for Clinical Excellence (NICE) and the National Screening Committee (NSC) who rely on HTA outputs to help raise standards of care. HTA findings also help to improve the quality of the service in the NHS indirectly in that they form a key component of the 'National Knowledge Service' that is being developed to improve the evidence of clinical practice throughout the NHS.

The HTA Programme was set up in 1993. Its role is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined to include all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care, rather than settings of care.

The HTA programme commissions research only on topics where it has identified key gaps in the evidence needed by the NHS. Suggestions for topics are actively sought from people working in the NHS, the public, consumer groups and professional bodies such as Royal Colleges and NHS Trusts.

Research suggestions are carefully considered by panels of independent experts (including consumers) whose advice results in a ranked list of recommended research priorities. The HTA Programme then commissions the research team best suited to undertake the work, in the manner most appropriate to find the relevant answers. Some projects may take only months, others need several years to answer the research questions adequately. They may involve synthesising existing evidence or designing a trial to produce new evidence where none currently exists.

Additionally, through its Technology Assessment Report (TAR) call-off contract, the HTA Programme is able to commission bespoke reports, principally for NICE, but also for other policy customers, such as a National Clinical Director. TARs bring together evidence on key aspects of the use of specific technologies and usually have to be completed within a limited time period.

## Criteria for inclusion in the HTA monograph series

Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this monograph was commissioned by the HTA Programme as project number 02/33/01. As funder, by devising a commissioning brief, the HTA Programme specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme or the Department of Health.

Editor-in-Chief: Professor Tom Walley  
Series Editors: Dr Peter Davidson, Professor John Gabbay, Dr Chris Hyde,  
Dr Ruairidh Milne, Dr Rob Riemsma and Dr Ken Stein  
Managing Editors: Sally Bailey and Caroline Ciupek

ISSN 1366-5278

© Queen's Printer and Controller of HMSO 2005

This monograph may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to NCCHTA, Mailpoint 728, Boldrewood, University of Southampton, Southampton, SO16 7PX, UK.

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.

Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.