

Surgical trial design – learning curve and surgeon volume

DETERMINING WHETHER INFERIOR RESULTS ARE DUE TO THE PROCEDURE ITSELF, OR DELIVERY OF THE PROCEDURE BY THE SURGEON



**A. H. R. W. Simpson,
C. R. Howie,
J. Norrie**

*The British Editorial
Society of Bone &
Joint Surgery, London,
United Kingdom*

A fundamental difference between drug trials and surgical trials is that the latter involve a procedure and therefore the patient outcome will depend on a complex interaction between surgeon, patient and the operating environment. With regards to the surgeon's impact; the outcome will be influenced by their proficiency in the delivery of the procedure. Surgical proficiency is known to depend on both the learning curve for a procedure and the case volume of the surgeon, i.e., it is necessary to both ascend the learning curve and to practice the procedure regularly.

The learning curve varies for different procedures. For instance for hip fracture fixation the operating time does not level off until the surgeon has done 20 to 30 of these operations.¹ For reduction of uncomplicated congenitally dislocated hips it may be slightly fewer as the performance seems to level after 10 to 20 procedures and can be accelerated with appropriate simulation training.² For arthroscopy of the knee, a simulation study suggests that consultant level skills are not reached until 170 procedures have been carried out.³ Thus although the actual number might vary, there will always be a certain number of cases a surgeon needs to have performed before his/her proficiency has plateaued. From the ethical point of view, before offering a new procedure a surgeon should be confident that he/she will do sufficient cases to complete the learning curve.⁴

In some trials the condition for which the procedure is carried out and the learning curve would be expected to have been completed during orthopaedic training e.g. primary hip replacement,⁵⁻⁸ closure of knee replacement wounds.⁹ Whereas in other studies the procedure is more specialised and the condition less common (e.g. pedicle lengthening for spinal stenosis); a surgeon is

less likely to complete the learning curve during training.¹⁰ In studies with low numbers,¹¹ it is particularly important to ensure that the learning curve has been completed. A more difficult issue occurs when a randomised trial is comparing a new with an established surgical procedure, and the surgeons may well be on different points of the learning curve for the new and the established procedures. Then a crude comparison of the outcomes might be unfair to the new procedure, as it includes less experience with the new technique, and so incorporates more of the transient ascent to the plateau of that new learning curve. Some methodologists have developed the concept of expertise based trials to partially address the problem of learning curves, such as that reported in Devereaux et al.¹² Another view is that although it might be methodologically correct to adjust out the learning curve in an efficacy-based randomised trial, if one is instead interested in the comparative effectiveness of two surgical procedures, the learning curve is part and parcel of that effectiveness – in the real world, the surgeons will have to ascend that learning curve on real patients, whose outcomes should count in the overall assessment of the two procedures.

Modelling of the learning curve has been discussed by several authors,^{13,14} and can help place the findings of a study in clinical context. However, it still remains that if clinical trials include surgeons on the learning curve, they will be assessing the learning curve as well the procedure itself. Thus clinical trials aimed at assessing the procedure should ensure that the surgeons included in the trial have all completed their learning curve. Currently, even well constructed and detailed trial protocols such as the WHITE3 study do not address this aspect.¹⁵ To ensure this is the case and that it is apparent, future

■ A. H. R. W. Simpson, MA(Cantab), DM (Oxon), FRCS(England & Edinburgh), George Harrison Law Professor of Orthopaedic Surgery, University of Edinburgh, Editor-in-Chief, Bone & Joint Research, 22 Buckingham Street, London WC2N 6ET, UK.

■ C. R. Howie, MB ChB, FRCS Ed Orth, Consultant Orthopaedic Surgeon and Honorary Professor, Department of Orthopaedic Surgery, University of Edinburgh, Edinburgh, E16, 4SA, UK.

■ J. Norrie, Professor of Medical Statistics and Trial Methodology, Usher Institute of Population Health Sciences and Informatics, Nine Edinburgh BioQuater, 9 Little France Road, Edinburgh, EH16 4UX, and Director, Clinical Trials Unit, University of Edinburgh, Outpatients Building, Floor Two, Room D36, Western General Hospital, Crewe Road South, Edinburgh, EH4 2XU, UK.

Correspondence should be sent to A. H. R. W. Simpson; email: editorbjr@boneandjoint.org.uk

doi: 10.1302/2046-3758.64.BJR-2017-0051

Bone Joint Res 2017;6:194–195.

surgical trial protocols should state (1) the number of cases they consider a surgeon should have performed to complete the learning curve for that the procedure and (2) that all the surgeons contributing to that trial have completed that number of cases.

The second area that is known to affect outcome is case volume. Where a condition or procedure is rare then a multicentre study is entirely appropriate to ensure case ascertainment. However when a condition is common and a multicentre trial is conducted it is important that the total case burden for that condition at the local contributing site is known; to indicate that the site is experienced; and that case ascertainment is representative of the condition burden and not a very highly select subgroup thereof, particularly if the patient demographics do not match the general demographics of the study or local overall demographics for that condition. Furthermore it is important that for operative procedures, or indeed the conservative management of fractures that the expertise to perform the intervention is assessed.

Whilst considering the PROFHER study,¹⁶ Ghert and Mckee¹⁷ draw attention to the fact that 66 surgeons operated on 125 patients over a two and a half year period, i.e., less than one patient per surgeon per year. To judge the proficiency of the participating surgeons it would have been very helpful to have data on the number of cases they were performing in total each year. Other examples where this information would have been useful include studies looking at cemented *versus* uncemented hip replacement¹⁸ and at metal on metal resurfacing.¹⁹ Larger studies involving DVT prophylaxis have included recruitment in 27 countries but not specified the number of recruitment centres or the total number of the relevant arthroplasty performed at those centres.²⁰ Thus as well as recording the usual recruitment and exclusions data it is important to understand the total population in these multicentre studies in order to draw appropriate conclusions from the data.

The potential differences in outcome of complex surgery delivered in high-volume centres as opposed to low volume centres has also be noted to be a potentially important factor in the design of pragmatic surgical trials.²¹

Thus future surgical trial protocols should state in the protocol (1) the minimum surgical volume considered necessary to maintain proficiency and (2) the trial data should include the number of cases surgeons are doing per year *outside* the study as well as those within the study.

If this data is not reported it is impossible to judge whether any inferior results are attributable to the procedure itself or to delivery of the procedure by the surgeon.

In summary, surgical trials should state; (i) in the protocol - the number of cases in the learning curve and (ii) the recommended annual surgical volume and (iii) when reporting the trial - that contributing surgeons have completed the learning curve and (iv) the number of cases of

the study type performed by each surgeon both within and outside the study.

References

1. Bjorgul K, Novicoff WM, Saleh KJ. Learning curves in hip fracture surgery *Int Orthop* 2011;35:113-119.
2. Jain NPM, Jowett AJL, Clarke NMP. Learning curves in orthopaedic surgery: a case for super-specialisation? *Ann R Coll Surg Engl* 2007;89:143-146.
3. Price AJ, Erturan G, Akhtar K, et al. Evidence-based surgical training in orthopaedics: how many arthroscopies of the knee are needed to achieve consultant level performance? *Bone Joint J*. 2015;97-B:1309-1315.
4. Hamilton DF, Simpson AH, Howie CR, Porter DE. The role of the surgeon in the application of the scientific method to new orthopaedic devices. *Surgeon* 2013;11:117-119.
5. Langlois J, Atlan F, Scemama C, Courpied JP, Hamadouche M. A randomised controlled trial comparing highly cross-linked and contemporary annealed polyethylene after a minimal eight-year follow-up in total hip arthroplasty using cemented acetabular components. *Bone Joint J* 2015;97-B:1458-1462.
6. Jonsson BA, Kadar T, Havelin LI, et al. Oxinium modular femoral heads do not reduce polyethylene wear in cemented total hip arthroplasty at five years: a randomised trial of 120 hips using radiostereometric analysis. *Bone Joint J* 2015;97-B:1463-1469.
7. McCalden RW, Korczak A, Somerville L, Yuan X, Naudie DD. A randomised trial comparing a short and a standard-length metaphyseal engaging cementless femoral stem using radiostereometric analysis. *Bone Joint J* 2015;97-B:595-602.
8. Jassim SS, Patel S, Wardle N, et al. Five-year comparison of wear using oxidised zirconium and cobalt-chrome femoral heads in total hip arthroplasty: a multicentre randomised controlled trial. *Bone Joint J* 2015;97-B:883-889.
9. Yuenyongvivat V, Iamthanaporn K, Hongnarak T, Tangtrakulwanich B. A randomised controlled trial comparing skin closure in total knee arthroplasty in the same knee: nylon sutures versus skin staples. *Bone Joint Res* 2016;5:185-190.
10. Li P, Qian L, Wu WD, Wu CF, Ouyang J. Impact of pedicle-lengthening osteotomy on spinal canal volume and neural foramen size in three types of lumbar spinal stenosis. *Bone Joint Res* 2016;5:239-246.
11. Jönsson BY, Mjöberg B. Porous titanium granules are better than autograft bone as a bone void filler in lateral tibial plateau fractures: A randomised trial. *Bone Joint J* 2015;97-B:836-841.
12. Devereaux PJ, Bhandari M, Clarke M, et al. Need for expertise based randomised controlled trials. *BMJ* 2005;330:88.
13. Cook JA, Ramsay CR, Fayers P. Statistical evaluation of learning curve effects in surgical trials. *Clin Trials* 2004;1:421-427.
14. Papachristofi O, Jenkins D, Sharples LD. Assessment of learning curves in complex surgical interventions: a consecutive case-series study. *Trials* 2016;17:266.
15. Sims AL, Parsons N, Achten J, et al. The World Hip Trauma Evaluation Study 3: Hemi-arthroplasty Evaluation by Multicentre Investigation - WHITE 3: HEMI - An Abridged Protocol. *Bone Joint Res* 2016;5:18-25.
16. Rangan A, Handoll H, Brealey S, et al. PROFHER Trial Collaborators. Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomized clinical trial. *JAMA* 2015;313:1037-1047.
17. Ghert M, McKee M. To operate or not to operate, that is the question: the proximal humerus fracture. *Bone Joint Res* 2016;5:490-491.
18. Inngul C, Blomfeldt R, Ponzer S, Enocsson A. Cemented versus uncemented arthroplasty in patients with a displaced fracture of the femoral neck: a randomised controlled trial. *Bone Joint J* 2015;97-B:1475-1480.
19. Tice A, Kim P, Dinh L, Ryu JJ, Beaulé PE. A randomised controlled trial of cemented and cementless femoral components for metal-on-metal hip resurfacing: a bone mineral density study. *Bone Joint J* 2015;97-B:1608-1614.
20. Eriksson BL, Borris LC, Friedman RJ, et al. Rivaroxaban versus enoxaparin for thromboprophylaxis after hip arthroplasty *N Engl J Med* 2008; 358:2765-2775
21. Ford I, Norrie J. Pragmatic Trials. *N Engl J Med* 2016;375:454-463.

Funding Statement

None declared.

ICMJE Conflicts of Interest

None declared.

© 2017 Simpson, Howie and Norrie. This is an open-access article distributed under the terms of the Creative Commons Attribution licence (CC-BY-NC), which permits unrestricted use, distribution, and reproduction in any medium, but not for commercial gain, provided the original author and source are credited.