

20 October 2016

Going far beyond the evidence

- [Helen H Handoll](#), Reader in Orthopaedics Teesside University
- Other Contributors:
 - Stephen D Brealey, Senior Trial Manager
 - Ada Keding, Statistician
 - Amar Rangan, Professor of Orthopaedic Surgery and Orthopaedic Surgeon

Sir,

We read this paper with interest. The publication on the same day of this meta-analysis and our article on defining the fracture population in the PROFHER trial¹ adds fuel to the debate on the interpretation of the results of pragmatic trials such as PROFHER. We suggest that our article,¹ the trial report² and a recent 'User's guide' to the medical literature relating to surgical trials featuring PROFHER³ provide strong support to the ready applicability of our trial findings. What is the place then of the findings and conclusions of Sabharwal et al's 'meta-analysis'?⁴ Below, we set out the main reasons why their meta-analysis is misconceived and why their conclusions, presented so boldly in their title, are invalid.

Inadequate selection of outcomes and presentation of the evidence. The selection of the four outcomes provides an incomplete and inadequate summary of the evidence. In particular, the Constant score reported is the sole measure of functional outcome. The authors do, however, acknowledge the validity of the patient-reported measures of function reported by PROFHER and more recently performed trials. Similarly selective is the restriction of adverse events to AVN, OA and non/malunion. This partiality is even more pronounced where the humeral head is replaced.

Inappropriate subgroup analyses. The stated virtue of this meta-analysis is being the first meta-analysis for this comparison that "has used subgroup analysis based on the complexity of fractures to compare patient outcomes". Despite a systematic review and trial literature replete with warnings on the dangers of subgroup analyses, the authors have taken a bold approach and in doing so have provided a misleading picture of the evidence. The selection *a priori* of subgroup variables is a recognised strength, but the classification of a subgroup variable based on a 50% predominance threshold for mixed fracture populations is misconceived, and confounding is unavoidable. The authors have also failed to recognise the limitations of summary data and to accept that the data are just not sufficient to support a valid subgroup analysis based on fracture type. Nor have they noted that the subgroup analyses presented for the PROFHER trial split between one- and two-part *versus* three- and four-part (or involvement or not of the tuberosities) do not indicate a difference in treatment effect (patient-reported function measured via the Oxford Shoulder Score) in the different subgroups.²

Incorrect interpretation of the results. Caution is required in the interpretation of the meta-analysis and even more so in the interpretation of subgroup analyses, the individual subgroup results inevitably being based on a reduced sample size. Having judged using GRADE the quality of the evidence for the overall analyses to be 'low' for health utility or 'very low' for the other three outcomes, this raises the question why the authors thought that it was safe to present the results for subgroups as evidence for underpinning their conclusion. This is particularly questionable for the

health utility data favouring replacement surgery that are from one trial of 49 participants with four-part fractures.

Although there are other errors, such as faulty reporting in the PRISMA diagram, that serve to undermine our confidence in the conduct and reporting of this review, we have kept our focus in this letter on the key flaws that invalidate the conclusions. Given the serious flaws in the conduct and reporting of the review, it is particularly unfortunate that their unsubstantiated conclusion is expressed in a headline-grabbing title.

H. Handoll, Reader in Orthopaedics,
Teesside University,
S.D. Brealey, Senior Trial Manager,
A. Keding, Statistician,
University of York,
A. Rangan, Professor in Orthopaedics and Consultant Surgeon,
James Cook University Hospital and University of York,
UK.

1. **Handoll HH, Brealey SD, Jefferson L, et al.** Defining the fracture population in a pragmatic multicentre randomised controlled trial: PROFHER and the Neer classification of proximal humeral fractures. *Bone Joint Res* 2016;5:481-489.
2. **Rangan A, Handoll H, Brealey S, et al.** Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomized clinical trial. *JAMA* 2015;313:1037-1047.
3. **Evaniew N, Carrasco-Labra A, Devereaux PJ, et al.** How to Use a Randomized Clinical Trial Addressing a Surgical Procedure: Users' Guide to the Medical Literature. *JAMA Surg* 2016;151:657-662.
4. **Sabharwal S, Patel NK, Griffiths D, et al.** Trials based on specific fracture configuration and surgical procedures likely to be more relevant for decision making in the management of fractures of the proximal humerus: findings of a meta-analysis. *Bone Joint Res* 2016;5:470-480.

Conflict of Interest:
None declared.