CONFIDENTIAL

smith&nephew

FAO Jeremy Powell
National Institute for Health and Clinical Excellence
1rst Floor, 10 Spring Gardens
London
SW1A 2BU
Jeremy.powell@nice.org.uk

RE: Smith & Nephew's comments on the Appraisal Consultation Document (review of technology appraisal guidance 2 and 44)

Dear Mr. Powell:

Smith & Nephew wishes to thank NICE's Appraisal Committee for this opportunity to respond to NICE's October 2013 Appraisal Consultation Document (***ACD***) concerning total hip replacement (***THR***) and resurfacing arthroplasty (***RSA***) of end-stage arthritis of the hip.

In this regard, we note that the key conclusions of NICE's Appraisal Committee, as stated in the ACD (p.43), include the following:

- "Total hip replacement and resurfacing arthroplasty prostheses are recommended as treatment options for people with end-stage arthritis of the hip only if the prosthesis has a rate (or projected rate) of less than 5% at 10 years;
- If more than one type of prosthesis meeting the above criteria is suitable for a patient, the prosthesis with the lowest acquisition costs should be chosen"

NICE has set out the evidence behind these and other key conclusions of its investigations (which remain provisional) in the Evaluation Report to which the ACD refers. In particular, the Evaluation Report contains a lengthy Assessment Report prepared by external consultants (the ***Warwick Report***), as well as details of the various submissions received by NICE from Smith & Nephew dated 4 February 2013 and others. However we have subsequently been informed that the submission from Smith & Nephew in support of our Birmingham Hip Resurfacing System (BHR) RSA product was unfortunately omitted from the process due to an internal error by NICE.

The Appraisal Committee has asked for comments on, among other things, the adequacy and completeness of this evidence base. We have detailed below our views on the first three questions posed at page 1 of the ACD, namely whether:

1. all the relevant evidence has been taken into account in reaching the key conclusions set out in the ACD;

2. the summaries of clinical and cost effectiveness set out in the ACD and the Evaluation Report constitute reasonable interpretations of the evidence; and

3. the key conclusions set out in the ACD are a sound and suitable basis for providing guidance to the NHS.

1. Has all of the relevant evidence been taken into account?

No, there are several significant omissions and errors that need to be considered for a fair assessment of RSA. These include: the lack of recognition of brand as a risk factor, no account of post-operative activity levels; enhanced QALYs and favorable cost-effectiveness as noted in the paper recently published by Heintzbergen et al. [redacted]

Relevant studies were excluded and misinterpreted

In the assessment of clinical evidence selection process, the exclusion criteria were very stringent and thus failed to consider a preponderance of evidence resulting in only 16 RCTs and 8 systematic reviews being included in the assessment. A significant example of an error in the interpretation of the data within the literature cited in the assessment is the Jameson et al. study, which identifies risk factors for Hip Resurfacing Arthroplasty (RSA) (p161 of the report). The assessment report selectively presents the reported hazard ratios for gender, femoral head size, and surgeon experience but excludes prosthesis brand as a risk factor, thus failing to reveal a significant finding that was one of the main conclusions of the publication. Jameson et al., identifies BHR as the baseline standard to which other brands were compared as BHR had the lowest revision rate of the resurfacing devices tracked in the National Joint Registry for England and Wales. The results reported in the publication are presented in Figure 1.

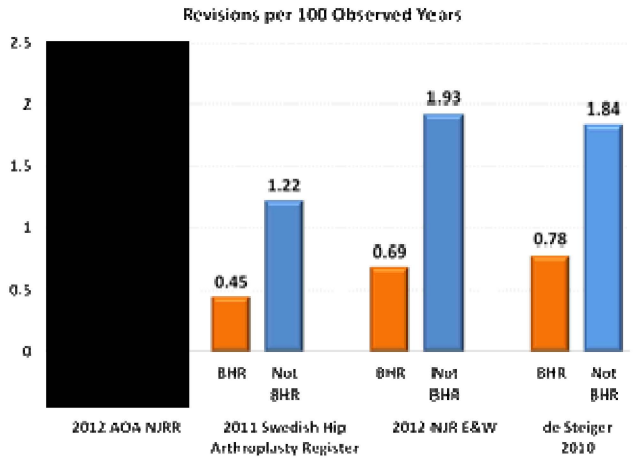
Figure 1. Independent predictors of revision following metal-on-metal hip resurfacing: A retrospective cohort study using national joint registry data (Jameson et al. The Journal of Bone and Joint Surgery 2012, 94-B:746-754).

The decision to restrict the reported analysis of registry, RCTs and publication data to RSA as a category, and to ignore the performance of individual device designs reported in peer-reviewed articles, excludes important implant survivorship data from the analysis and denies the identification of well-performing devices from the category. The conclusions drawn from the analysis are therefore incomplete and misleading for RSA and in particular BHR.

Data reported from the Australian Orthopedic Association National Joint Registry, the Swedish Hip Arthroplasty Register, the National Joint Registry for England and Wales, and a study by de Steiger et al. reporting on registry data, corroborate the contention that not all RSA prostheses display similar survivorship profiles. These data sources show a marked difference in the revision incidence rate per 100 observed implant years, Figure 2.

CONFIDENTIAL

Figure 2. RSA revision rates.



This lack of brand transparency is further accentuated in the report by the exclusion of brand as a risk factor in the summarized data in Table 49 (a summary table of registry studies for RSA). Noteworthy is the degree of hazard reported and the degree of hazard not reported by ignoring brand differences. The table reports a 1.3 hazard ratio (HR) difference between women and men, a 1.48 - 2.14 HR for smaller femoral head components, and a 1.36 HR for low volume surgeons. It ignores a higher HR by associated with brand, 1.43 – 2.82 for other resurfacing devices compared to BHR.

Exclusion of recent, evidence on cost-effectiveness

Smith & Nephew notes that a significant new cost-effectiveness analysis model for hip resurfacing has been published since the Warwick Report was compiled. We believe that NICE ought properly to take this more recent evidence into account in its conclusions.

Heintzbergen et al. overcomes several of the limitations that the Warwick Report identified as limitations to its own methods. Furthermore, it provides significant evidence that the outcomes of BHR are unique compared to an assessment of RSA outcomes calculated from a category average. [redacted]

In the evidence review portion of the report, page 130 states that it was assessed that there was "lower grade evidence" that there was no difference between QALYs or functional scores between RSA and THA. On page 243 of the report, it was stated that the applied prosthesis cost was an average cost. Rather, a weighted average cost would have better reflected prostheses costs for use in the model.

The ACD stated that "revision rate was the single most important key driver of costs and QALYs in the model." An RSA revision rate of 17.2% was used at 10 years for hip resurfacing. The actual 10-year revision rate for BHR in the Australian and Swedish arthroplasty registers was 6.7% and 6.2%, respectively, and was 5.09% at 7 years in the NJREW. The numbers used in the Warwick model would thus significantly disadvantage BHR and could be an explanation for the difference in final assessment between the Warwick model and Heintzbergen et al. paper.

Moreover, post-surgery activity levels for RSA and THR were not taken into account in the assessment in terms of both health economic arguments and the effect of higher activity on expected implant survivorship.

2. Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?

No, significant gaps in data and limitations of the analysis do not allow for an accurate cost effectiveness conclusion. The Warwick analysis is flawed because RSA was evaluated only as a category. Implant brand, reported by Jameson, et al., is a significant predictor of revision risk, however was not differentiated within model. Rather, all RSA brands were assumed to have the same risk of revision.

[redacted]

The Warwick analysis is also flawed because brand was not differentiated in the RSA segment, rather all RSA brands have been assumed to have the same risk of revision. Moreover in section 6.2.5, there are only 3 RCTs that have been analyzed and stated to have been the foundation of the comparison between RSA and THA. In two of these studies, the RSA product evaluated was Durom (Zimmer), which was subject to a recall in 2008 due to higher than anticipated revision rates. The other RCT did not specify which products were used as the brand selection as the choice was left to surgeon discretion. The lack of consideration of brand in the selection of RCT data makes these studies questionable in relevance and strength of the overall conclusions. Furthermore, the data used in the systematic reviews and the inclusion of several products in the RSA group which have subsequently been withdrawn from the market due to having high failure rates, bring into question the strength of any conclusions made in the Evaluation Report.

3. Are the provisional recommendations sound and a suitable basis for guidance to the NHS?

We support the objectives of NICE, namely to ensure that the NHS has access to the most clinically effective and cost effective hip prostheses. However, we do not support the recommendations as they currently stand. The proposed guidance first recommends the choice of implants only meeting the recommendation of less than 5% revision rate at 10 years. As regards **THA**, this recommendation may be applicable for the current products, however the second part of the recommendation being to select the least expensive implant within that group which falls below 5%, could lead to an increased cost burden to the NHS. This is because the Warwick report has viewed cost to only include implant acquisition and omitted many other important contributors to the cost (ancillary products, theater time,

QALYs, etc.). We cannot support the proposed recommendation of that implant cost acquisition solely as the determining cost factor in implant choice. It must be recognized that an implant meeting the guidance but performing better than another lower cost implant may in fact result in a total cost that is lower than the less expensive implant. This assertion is supported by our review of the paper by Heintzbergen et al. which performed a rigorous cost effective analysis in comparison to the Warwick report. In fact, the Appraisal Consultation Document itself seems conflicted in that a key conclusion that "Revision rate was the single most important key driver of costs and QALYs in the model" appears contradictory to a finding reported on page 303 "…that the cost of the prosthesis is the most important factor…". As such, we feel that the second proposed guidance of choosing the lowest cost implant, while well intentioned, is very incomplete and may perversely impact in an effort to assist the NHS in reducing their overall cost burden without further qualification.

Finally, we do not believe that the data, or analysis of the data in its current form, can support any meaningful conclusions relative to RSA. We highlighted significant shortcomings in the analysis performed for RSA including the lack of recognition of brand as a risk factor, no account of post-operative activity levels, enhanced QALYs and favorable cost-effectiveness as noted in the paper by Heintzbergen et al. Moreover, we have established that there is strong evidence contradicting the conclusion of the Warwick report that BHR is more cost effective than THA. Finally, the evidence submitted by Smith & Nephew to NICE on 4 February 2013 for the purpose of the current technology appraisal was unfortunately not taken into account in a timely manner. In light of these shortcomings it would be appropriate to perform an additional review of the RSA data with a view to assess appropriate benchmarks for RSA in this younger more active patient population.

Yours sincerely,

[redacted]

[redacted]

smith&nephew

References

De Steiger R.N., Miller L.N., Prosser G.H., Graves S.E., Davidson D.C., Stanford T.E., Poor outcome of revised resurfacing hip arthroplasty. Acta Orthop. 2010; 81(1): 72–76.

Gooch Kl., Smith D., Wasylak T., Faris PD., Marshall Da., Khong H., Hibbert JE., Parker RD., RF.,Zernicke., Beaupre R., Pearce T., Johnston WMC., Frank CB. The Alberta hip and knee replacement project: A model for health technology assessment based on comparative effectiveness of clinical pathways. International Journal of Technology Assessment in Health Care, 25:2 (2009), 113–123.

The Alberta Hip Improvement Project. Functional Outcomes for 2 Years Comparing Hip Resurfacing and Total Hip Arthroplasty. The Journal of Arthroplasty Vol. 27 No. 5 2012

Heintzbergen S., Kulin N.A., IJzerman M.J., Steuten L.M.G., Werle j., Khong H., Marshall D.A., Cost-Utility of Metal-on-Metal Hip Resurfacing Compared to Conventional Total Hip Replacement in Young Active Patients with Osteoarthritis. Value in Healthcare 2013; (16) 942-952.*

Jameson S.S., Baker P.N., Mason J., Porter M.L., Deehan D.J., Reed M.R., Independent predictors of revision following metal-on-metal hip resurfacing: a retrospective cohort study using National Joint Registry data. J Bone Joint Surg Br. 2012; 94(6):746-54.

Lavigne M., Masse V., Girard J., Roy A.G., Vendittoli P.A., Return to sport after hip resurfacing or total hip arthroplasty: a randomized study. Rev Chir Orthop Reparatrice Appar Mot. 2008 Jun;94(4):361-7.

[redacted]