Empirical comparison of Discrete Choice Experiment and Best-Worst Scaling to estimate stakeholders' risk tolerance for hip replacement surgery

Joris D. van Dijk, BSc[1]; Karin C.G.M. Groothuis-Oudshoorn, PhD[1]; Deborah A. Marshall, PhD[2]; Maarten J. IJzerman, PhD[1]
(1) University of Twente, dept. Health Technology and Services Research, Enschede, the Netherlands
(2) Alberta Bone & Joint Health Centre and University of Calgary, Calgary, AB, Canada

OBJECTIVES

Empirical comparison of two preference elicitation methods estimating the risk tolerance for total hip arthroplasty and total hip resurfacing arthroplasty;

- Discrete choice experiment (DCE) &
- Profile case best-worst scaling (BWS).

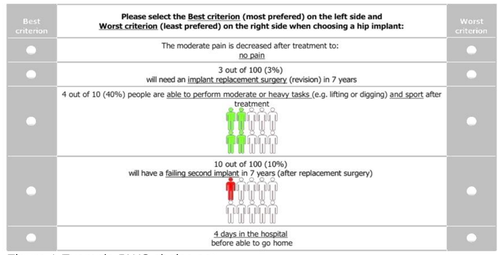

Figure 1 Example BWS choice set

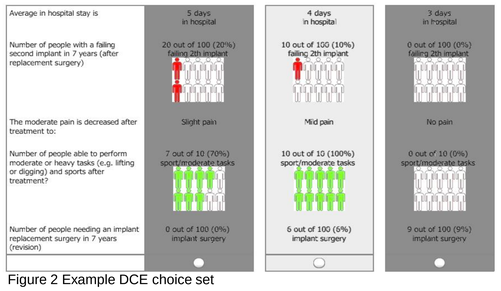

Figure 2 Example DCE choice set

Table 1: Scenarios used for MAR estimation*

Attributes	Base case scenario	Scenario 1	2	3	4
Chance first revision in 7 years	3.9% <60 years 4.2% < 70 years [1]	6%	6%	9%	9%
Chance second revision per 7 years	31% [2]	20%	20%	30%	30%
Moderate pain decrease to	Slight pain	No pain	Slight pain	Slight pain	Slight pain
Able to perform daily tasks	39% [3]	100%	40%	40%	0%
Average in hospital stay is	4.7 days [5]	3 days	5 days	5 days	6 days

*The utility of the four attributes of each scenario was substracted from the utility of the base case scenario resulting in a risk utility allowing calculation of the corresponding maximum acceptable risk

Table 2: Time and difficulty to complete the choice sets

	Median time per task (p5-p95)	Mean difficulty (95% Confidence interval)
DCE	24 s. (8-69 s.)	2.51 (2.42-2.60)
BWS	28 s. (11-80 s.)	2.57 (2.48-2.66)
	p < 0.01	p = 0.059

METHODS

Online survey under males in general population in the US between 45-65 years.

- Socio-demographic questions.
- Two randomized blocks with 12 DCE and 8 BWS questions (e.g. in Figure 1 & 2).

Difficulty choice sets was rated & time to complete the choices was monitored.
Choice sets were analysed using conditional logit analysis.
The risk tolerance; the maximum acceptable risk (MAR) for first or second revision surgery, was estimated using four hypothetical scenarios (shown in Table 1).

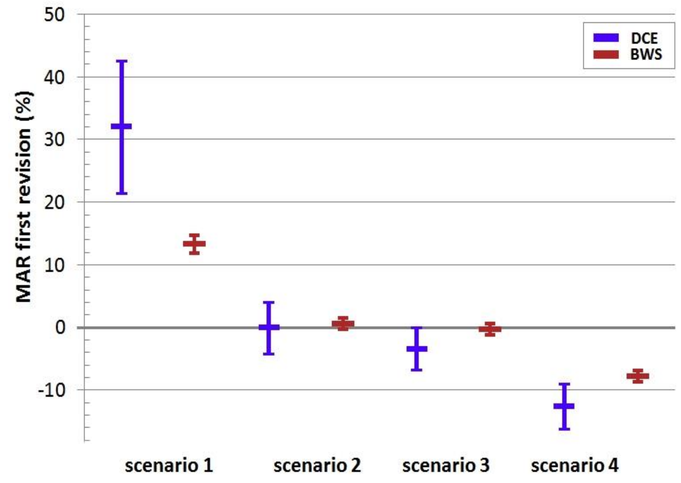

Figure 3: Estimated MARs (including 95% confidence intervals) for the chance on a first revision in seven years by DCE and BWS for all four hypothetical scenarios.

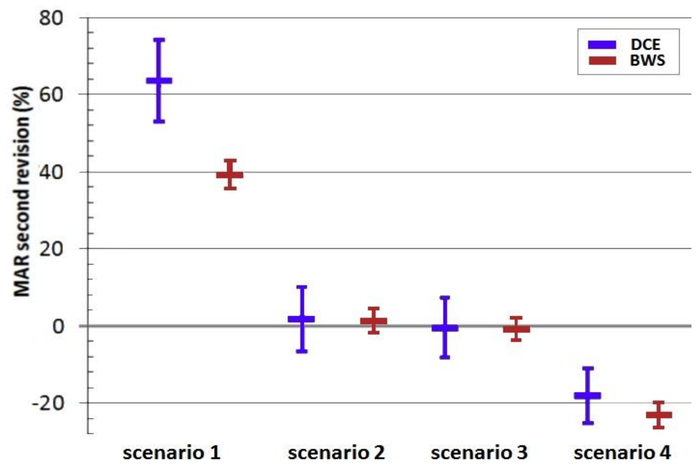

Figure 4: Estimated MARs (including 95% confidence intervals) for the chance on a second revision in seven years after first revision by DCE and BWS for all four scenarios.

RESULTS

429 respondents completed the online survey.
Determined part worth utilities were the same as expected, except for hospital stay by DCE. Yet, the part worth utilities for changing to higher levels differ between both methods (see Figure 5).

Also the MAR differed between both methods for the different scenarios; ranging from 0% to 19% MAR difference between the methods for a first revision (see Figure 3 and 4).

BWS choice sets took more time to complete (see Table 2).
BWS choice sets were rated as more difficult (see Table 2).

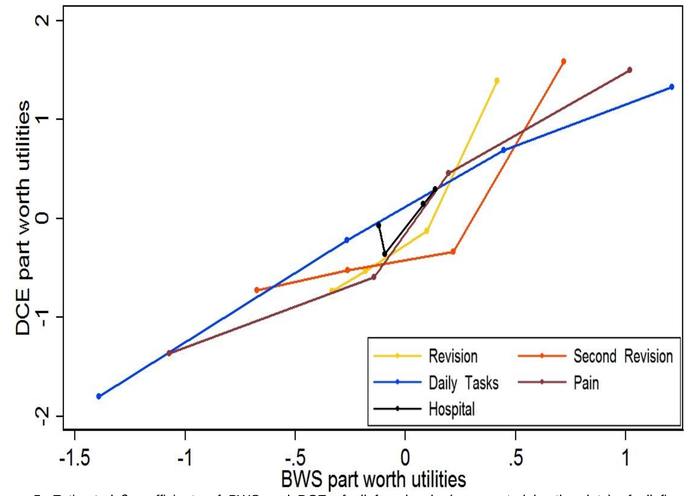

Figure 5: Estimated β-coefficients of BWS and DCE of all four levels (represented by the dots) of all five attributes (connected through the lines). The first levels, with the highest part worth utilities, are the dots in the most upper right corner.

CONCLUSIONS

Both methods produce different results, might leading to different outcomes.
So the choice between methods to elicit stakeholder preferences may induce decision sensitivity. Yet, both seem to be consistent in predicting risk tolerance if the benefits are changed.

Looking at MARs, DCE seems to be more sensitive for a change in benefits and risks while the MAR estimates obtained through BWS have considerable lower uncertainty than DCE.

Strengths

- Empirical research comparing the outcomes of BWS and DCE within persons, using a large group of respondents.
- Survey was constructed using international guidelines with a minimal number of advised tasks for best comparison.

Limitations

- Attribute selection; possible overlap in the attributes chance on first and second revision.
- Linearity was assumed between the part worth utilities for the MAR estimation.

Future research

- Additional empirical comparisons between multiple preference elicitation methods.
- Validation of these methods by comparing it to actual patient decisions.

UNIVERSITY OF TWENTE.

References
1 National Joint Registry for England and Wales, 2011
2 Australian Orthopaedic Association National Joint Replacement Registry, 2010
3 Venditoli P, et al., 2006
4 Alberta Bone and Joint Health Institute, 2006

Poster presenter: Joris van Dijk, BSc.
For a poster copy or additional information:
jorisvdijk@gmail.com