



The paradox of self-rated health following joint replacement surgery

Thomas Perneger^{1,2} · Anne Lübbecke^{2,3}

Accepted: 27 September 2018 / Published online: 15 October 2018
© Springer Nature Switzerland AG 2018

Abstract

Purpose Self-rated health is a commonly used patient-reported outcome, but its responsiveness to is not well documented. We examined the ability of self-rated health to capture health changes attributable to a highly effective surgical intervention.

Methods Prospective study of patients with severe osteoarthritis of the hip ($N=990$) or knee ($N=907$) who underwent total hip replacement (THA) or total knee replacement (TKA). Self-rated health was assessed pre-operatively and 1 year after surgery on a scale between “excellent” and “poor,” along with other health items (other 11 items of the SF12 questionnaire) and multi-item Pain and Function scales.

Results On average, self-rated health was unchanged by surgery. In both THA and TKA cohorts, of 10 patients, 6 rated their health the same after surgery as before, 2 gave a higher rating, and 2 gave a lower rating. In contrast, major improvements were observed for all other SF12 items, and for the Pain and Function scales, in both cohorts of patients. Nevertheless, both before and after surgery, self-rated health was associated with the other SF12 items and with Pain and Function scores. These associations were stronger after surgery than before.

Conclusions Self-rated health was not responsive to major improvements in health, documented by other instruments, attributable to joint replacement surgery. However, self-rated health was even more strongly associated with concurrent assessments of more specific health problems after surgery than before. Caution is advised in interpreting changes in self-rated health following health-altering interventions.

Keywords Self-rated health · Responsiveness to change · Joint replacement surgery · SF12 · WOMAC

Introduction

A commonly used patient-reported outcome is a person’s assessment of his or her health as excellent, very good, good, fair, or poor [1–3]. This “self-rated health” is simple to obtain, and yet has notable predictive properties: it predicts subsequent mortality [4, 5], use of health services [6] and health care expenditures [7], and is associated with various health limitations [8–10].

Whether self-rated health captures essentially the current health state of a person, which is prone to change over time,

or whether it reflects an enduring perception of one’s health, remains debated [11]. A situation in which this issue can be addressed is that of symptomatic patients who undergo a highly effective treatment, such as joint replacement surgery for advanced osteoarthritis. Symptoms will improve on average, but whether health ratings follow is unclear. We have examined the changes in self-rated health from before surgery to 1 year after surgery among patients who underwent arthroplasty of the hip or knee, and compared them to changes of other more specific health assessments.

Methods

Study design and sample

This prospective study was based on the Geneva Arthroplasty Registry [12], which includes all patients undergoing total arthroplasty of the hip or knee at a Swiss teaching hospital. Registry procedures and the use of the data for

✉ Thomas Perneger
thomas.perneger@hcuge.ch

¹ Division of Clinical Epidemiology, Geneva University Hospitals, 4 rue Gabrielle-Perret-Gentil, 1211 Geneva, Switzerland

² University of Geneva, Geneva, Switzerland

³ Division of Orthopedic Surgery and Traumatology, Geneva University Hospitals, Geneva, Switzerland

research were approved by the Research Ethics Commission. The study included patients who underwent primary elective total hip arthroplasty (THA) or total knee arthroplasty (TKA) between 2010 and 2016.

Study variables

The main outcome variable was self-rated health, an item of the Short Form 12-item (SF12) questionnaire [13], adapted into French [14]. Other health measures included the other SF-12 items, and Pain and Function scores (from 0 = worst, to 100 = best) assessed by the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [15]. Descriptive variables included patient sex, age, body mass index, smoking status, and pre-operative numerical scales for pain (between 0 and 10) and activity level (from 0 to 10 = high activity) [16].

Analysis

We included all eligible patients without a priori sample size determination. We compared baseline characteristics of THA and TKA patients; subsequent analyses were done separately for the 2 cohorts. We described self-rated health before and after surgery, and obtained a measure of change, as the number of categories by which self-rated health changed from before surgery to after surgery. To assess reliability, we obtained two-way mixed intraclass correlation coefficients (ICC) for absolute agreement.

We examined the change from before to after surgery for all SF12 items. Because the number of response options varies from 2 to 6 for the different items, all were dichotomized to facilitate comparisons. We report proportions with the positive rating before and after surgery, and a McNemar odds ratio for change. To explore if self-rated health relates in a similar way to the other measures before and after surgery, we obtained odds ratios for a positive rating on each of the 11 SF12 items in relation to “excellent” or “very good” self-rated health. We also examined means of WOMAC

scores across health ratings before and after surgery. Analyses were performed using SPSS version 22.

Results

The study included 990 patients with THA and 907 patients with TKA who returned SF12 and WOMAC questionnaires before and 1 year after surgery. TKA patients were older than THA patients [mean 71.5 years (SD 9.4) vs. 68.6 (11.9), $p < 0.001$], and more likely to be female (65.5% vs. 53.7%, $p < 0.001$), obese (43.9% vs. 23.4%, $p < 0.001$), and non-smokers (89.3% vs. 72.9%, $p < 0.001$). Pre-operative pain levels (means 6.2 vs. 6.1, $p = 0.66$) and activity levels (means 3.6 vs. 3.6, $p = 0.49$) were similar.

Self-rated health

In both cohorts, and both before and after surgery, the distributions of self-rated health were symmetrical (Table 1). Among THA patients, more rated their health as “excellent” after surgery than before (7.8% vs. 3.9%), but fewer rated health as “very good” (19.7% vs. 22.6%). Among TKA patients, the distributions were similar before and after surgery. ICCs were at 0.60 and 0.56 in the 2 cohorts.

The change in self-rated health categories, from before to after surgery, was symmetrical in both cohorts of patients (Fig. 1). Self-rated health remained unchanged in 61.1% of patients with THA, and 60.7% of patients with TKA.

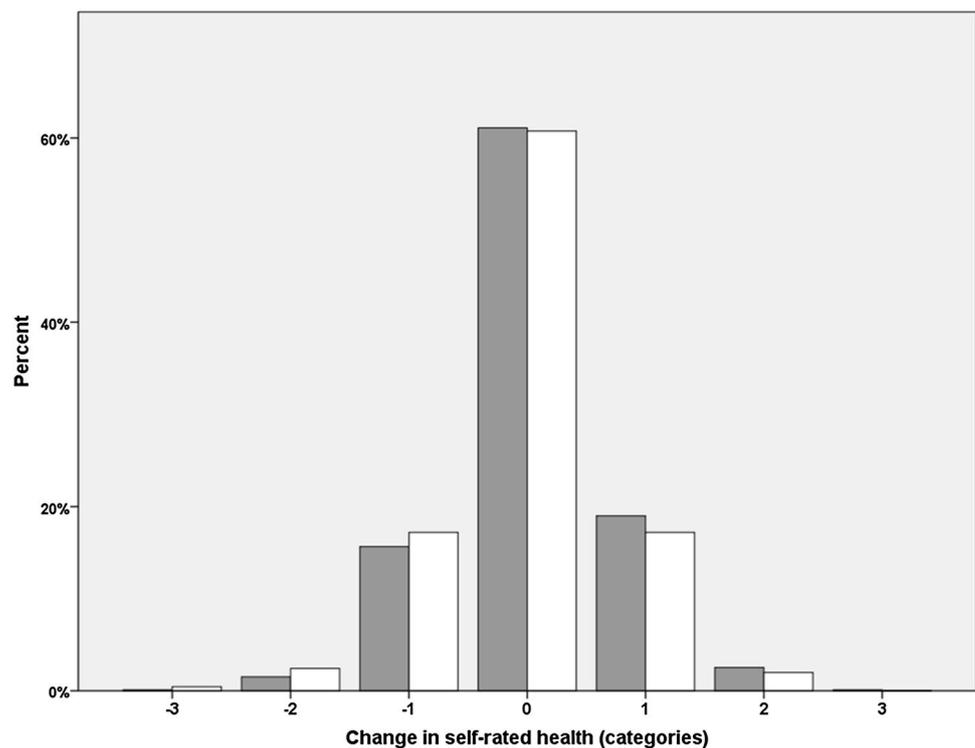
Changes over time of SF12 items

A positive rating on self-rated health (“excellent” or “very good”) remained at the same level before and after surgery in both cohorts of patients, and the McNemar odds ratios were near one (Table 2). In contrast, all other SF12 items improved following surgery, in both cohorts.

Table 1 Distributions of responses to the self-rated health item before and after total hip arthroplasty or total knee arthroplasty, Geneva Arthroplasty Registry, 2010–2016

	Total hip arthroplasty ($N=990$)		Total knee arthroplasty ($N=907$)	
	Pre-operative	At 1 year	Pre-operative	At 1 year
Self-rated health, N (%)				
Excellent	39 (3.9)	77 (7.8)	41 (4.5)	40 (4.4)
Very good	224 (22.6)	195 (19.7)	166 (18.3)	163 (18.0)
Good	566 (57.2)	557 (56.3)	529 (58.3)	522 (57.6)
Fair	138 (13.9)	144 (14.5)	149 (16.4)	156 (17.2)
Poor	23 (2.3)	17 (1.7)	22 (2.4)	26 (2.9)
Intraclass correlation coefficient (95% CI)	0.60 (0.55–0.63)		0.56 (0.51–0.60)	

Fig. 1 Distribution of change in self-rated health on a 5-category scale (excellent, very good, good, fair, poor), expressed in number of categories, between the pre-operative assessment and 1 year after surgery, in patients with total hip arthroplasty (gray bars) and total knee arthroplasty (white bars)



Associations with other health measures

The other eleven SF12 items were associated with “excellent” or “very good health,” both before and after surgery (Table 3). Of note, in both cohorts of patients, the odds ratios were weaker pre-operatively than post-operatively. The proportion of variance in self-rated health explained by all other 11 items (Nagelkerke R^2) increased from 0.16 to 0.34 in THA patients, and from 0.21 to 0.36 in TKA patients.

Pain and Function scores were associated with self-rated health, but in post-operative assessments, the means were considerably higher at any given level of self-rated health than before surgery; this was seen for both scales, in both cohorts (Table 4). The variance explained (eta squared) increased as well, in both cohorts.

Discussion

In these cohorts of patients with severe osteoarthritis of the hip or knee, self-rated health was insensitive to the effects of joint replacement surgery. On average, patients were helped by joint replacement surgery, as reflected by favorable changes in all aspects of health measured by the other SF12 items and Pain and Function scales. Yet, self-rated health remained at the same level 1 year after surgery as before the operation: of 10 patients, roughly six rated their health the same after surgery as before, 2 gave a higher rating, and 2 gave a lower rating. This was seen in both cohorts.

Does this mean that self-rated health is unrelated to specific symptoms? No—both before and after surgery, self-rated health was associated with the patients’ symptoms and limitations. Hence the paradox: self-rated health is associated with the burden of symptoms, yet when the symptoms change self-rated health does not. We believe that the most likely explanation is that joint replacement surgery has two effects: on the one hand surgery improves symptoms by a direct causal mechanism, on the other hand surgery (or, more specifically, the pre-surgical vs post-surgical context) alters the contributions of various determinants of self-rated health, whether the symptoms that were improved by surgery or other determinants that were not.

Furthermore, associations between symptoms and self-rated health were stronger after surgery than before. This was reported previously [17]. Possibly, in the pre-operative period, when patients experience severe symptoms, their self-rated health may be more dependent on factors other than symptoms captured by SF12 items and WOMAC scales. Context-dependent changes in determinants of self-rated health have not received much attention thus far.

Our findings also support the notion that people’s characterisation of their health includes a stable component, that may take into account one’s ability to cope with current symptoms, and the perspective to maintain health in future [11]. In the pre-operative period, patients who suffer from osteoarthritis may view their pain and limitations as fixable problems, which do not affect their health in a broad sense. In the post-operative period, patients may have

Table 2 Changes between assessments made pre-operatively and 1 year after surgery for 12 dichotomized items of the SF12 Health Survey, with conditional odds ratios, in patients with total hip arthroplasty and total knee arthroplasty

Abbreviated item content	Positive answer	Negative answer	Total hip arthroplasty			Total knee arthroplasty		
			Positive answer		McNemar odds ratio (95% CI)	Positive answer		McNemar odds ratio (95% CI)
			Before surgery (%)	1 year after surgery (%)		Before surgery (%)	1 year after surgery (%)	
Self-rated health	Excellent or very good	Good, fair or poor	26.6	27.5	1.1 (0.8–1.5)	22.8	22.4	0.9 (0.7–1.3)
Limited in moderate activities	Not at all	Yes, a little, or yes, a lot	18.2	51.3	9.9 (7.0–13.8)	24.0	37.3	2.5 (2.0–3.3)
Limited in climbing several flights of stairs	Not at all	Yes, a little, or yes, a lot	12.3	41.0	10.5 (7.2–15.2)	10.7	25.2	4.0 (2.9–5.6)
Due to physical health, accomplished less	No	Yes	15.7	63.3	12.5 (9.1–17.2)	22.4	52.3	5.1 (3.9–6.6)
Due to physical health, was limited in kind of activities	No	Yes	18.0	58.4	11.0 (8.0–15.2)	25.8	48.5	3.7 (2.9–4.8)
Pain interfered with work	Not at all or a little bit	Moderately, quite a bit or Extremely	12.3	63.6	26.4 (16.9–41.3)	13.3	50.5	11.5 (8.0–16.5)
Due to emotional problems, accomplished less	No	Yes	45.4	65.9	3.2 (2.6–4.1)	44.8	60.9	2.6 (2.0–3.2)
Due to emotional problems, was less careful	No	Yes	41.9	68.4	5.0 (3.9–6.6)	47.1	59.6	2.1 (1.7–2.7)
Felt calm and peaceful	All, most or a good bit of the time	Some, a little or none of the time	53.7	69.6	3.1 (2.4–4.0)	55.0	64.5	2.0 (1.5–2.5)
Had a lot of energy	All, most or a good bit of the time	Some, a little or none of the time	25.5	41.4	3.4 (2.6–4.5)	27.2	37.7	2.1 (1.6–2.7)
Felt downhearted and blue	A little or none of the time	All, most, a good bit or some of the time	45.4	57.2	2.5 (1.9–3.3)	44.1	54.0	2.0 (1.5–2.6)
Physical or emotional problems interfered with social activities	A little or none of the time	All, most or some of the time	38.4	64.1	5.4 (4.1–7.1)	44.0	57.7	2.4 (1.9–3.1)

become accustomed to the new state of things, and do not consider themselves as substantially healthier than before. Such adaptive phenomena are often described as response shift [18–21]. It is possible that many patient-reported outcome measures are context-dependent, which would call for caution in interpreting any post-intervention change—or lack thereof.

Another possible explanation for lack of responsiveness is poor reliability. Self-rated health is a single item, and as such must capture a fair amount of noise. In our cohorts, the ICCs were 0.56 and 0.60, which is acceptable for a single

item, and similar to previous reports [22]. Furthermore, the strong associations of self-rated health with other health scales indicate that lack of reliability was not the chief reason for the observed stability.

The availability of two distinct cohorts of patients, and the comparability of the phenomena observed in each, suggests that our findings may be generalizable. Furthermore, the sample size was sufficiently large to yield stable estimates, the data collection procedures were standardized, and the instruments well known and reputed valid. The main weakness of our study is the lack of qualitative data

Table 3 Associations (odds ratios and 95% confidence intervals) between “excellent” or “very good” self-rated health and positive ratings on the other 11 items of the SF12 Health Survey, before and after surgery, in patients with total hip arthroplasty and total knee arthroplasty

Abbreviated item content	Total hip arthroplasty		Total knee arthroplasty	
	Before surgery	1 year after surgery	Before surgery	1 year after surgery
Limited in moderate activities	2.9 (2.1–4.1)	6.1 (4.4–8.6)	3.0 (2.1–4.2)	5.5 (3.9–7.7)
Limited in climbing several flights of stairs	2.9 (2.0–4.3)	5.1 (3.8–6.9)	2.0 (1.3–3.1)	5.3 (3.8–7.5)
Due to physical health, accomplished less	2.0 (1.4–2.8)	4.0 (2.8–5.7)	1.9 (1.4–2.7)	6.2 (4.2–9.3)
Due to physical health, was limited in kind of activities	1.9 (1.3–2.7)	4.0 (2.9–5.6)	1.9 (1.4–2.7)	3.7 (2.6–5.2)
Pain interfered with work	2.5 (1.7–3.7)	6.3 (4.2–9.4)	3.1 (2.1–4.6)	6.9 (4.6–10.2)
Due to emotional problems, accomplished less	2.9 (2.1–3.8)	5.6 (3.7–8.3)	2.7 (1.9–3.7)	5.2 (3.4–7.9)
Due to emotional problems, was less careful	2.8 (2.1–3.8)	7.0 (4.5–10.9)	3.3 (2.4–4.6)	7.0 (4.5–11.1)
Felt calm and peaceful	2.7 (2.0–3.7)	4.9 (3.3–7.4)	4.2 (2.9–6.0)	6.4 (4.0–10.2)
Had a lot of energy	3.2 (2.3–4.3)	5.4 (4.0–7.3)	3.9 (2.8–5.5)	7.3 (5.1–10.4)
Felt downhearted and blue	2.9 (2.2–3.9)	4.8 (3.4–6.8)	3.7 (2.7–5.2)	4.7 (3.2–6.8)
Physical or emotional problems interfered with social activities	2.4 (1.8–3.2)	6.1 (4.1–9.1)	2.4 (1.8–3.3)	5.6 (3.7–8.4)

Table 4 Mean values (standard deviations) of WOMAC Pain and Function scores, according to self-rated health, before and after surgery, in patients with total hip arthroplasty and total knee arthroplasty

	Total hip arthroplasty		Total knee arthroplasty	
	Pre-operative	One year after surgery	Pre-operative	One year after surgery
WOMAC Pain score, mean (SD)				
Self-rated health				
Excellent	51.5 (15.6)	93.4 (12.8)	48.2 (21.6)	89.9 (15.8)
Very good	46.4 (17.1)	92.4 (12.8)	43.9 (16.0)	85.7 (15.4)
Good	39.5 (17.5)	85.5 (17.6)	39.9 (17.0)	75.2 (19.8)
Fair	30.0 (16.2)	66.2 (25.6)	31.4 (15.4)	56.0 (21.4)
Poor	21.8 (15.9)	70.3 (31.3)	23.2 (15.0)	38.1 (26.9)
Eta squared	0.11	0.18	0.08	0.26
WOMAC Function score, mean (SD)				
Self-rated health				
Excellent	54.9 (17.9)	91.8 (16.0)	56.4 (20.0)	88.4 (16.0)
Very good	47.7 (18.6)	90.4 (14.2)	50.6 (18.2)	83.0 (17.3)
Good	40.2 (17.9)	79.7 (18.7)	44.2 (18.4)	72.0 (19.8)
Fair	29.5 (15.8)	55.6 (22.7)	33.1 (16.7)	51.5 (20.3)
Poor	23.1 (16.0)	50.0 (25.1)	23.9 (14.9)	36.0 (23.9)
Eta squared	0.12	0.28	0.12	0.27

about the reasons for patients’ ratings of their health, and thus a limited capacity to explain the observed stability of self-rated health. Qualitative research about the meaning of self-rated health (e.g., [23]) may help identify factors that contribute the most to health ratings in particular contexts.

Funding The Geneva Arthroplasty Registry is supported in part by the Fondation pour la Recherche Ostéo-Articulaire.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

- Jylhä, M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social Science & Medicine*, 69, 307–316.
- Fayers, P. M., & Sprangers, M. A. G. (2002). Understanding self-rated health. *Lancet*, 359, 187–188.
- Garbarski, D. (2016). Research in and prospects for the measurement of health using self-rated health. *Public Opinion Quarterly*, 80, 977–997.
- DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2005). Mortality prediction with a single general self-rated health question: A meta-analysis. *Journal of General Internal Medicine*, 21, 267–275.
- Lima-Costa, F. M., Cesar, C. C., Chor, D., & Proietti, F. A. (2012). Self-rated health compared with objectively measured health

- status as a tool for mortality risk screening in older adults: 10-year follow-up of the Bambui Cohort Study on Aging. *American Journal of Epidemiology*, 175, 228–235.
6. DeSalvo, K. B., Fan, V. S., McDonnell, M. B., et al. (2005). Predicting mortality and healthcare utilization with a single question. *Health Services Research*, 40, 1233–1246.
 7. DeSalvo, K. B., Jones, T. M., Peabody, J., et al. (2009). Health care expenditure prediction with a single item, self-rated health measure. *Medical Care*, 47, 440–447.
 8. Latham, K., & Peek, C. W. (2013). Self-rated health and morbidity onset among late midlife US adults. *The Journals of Gerontology*, 68, 107–116.
 9. Perneger, T. V., Gayet-Ageron, A., Courvoisier, D. S., Agoritsas, T., & Cullati, S. (2013). Self-rated health: Analysis of distances and transitions between response options. *Quality of Life Research*, 22, 2761–2768.
 10. Meng, X., & D'Arcy, C. (2016). Determinants of self-rated health among Canadian seniors over time: A longitudinal population-based study. *Social Indicators Research*, 126, 1343–1353.
 11. Bailis, D. S., Segall, A., & Chipperfield, J. G. (2003). Two views of self-rated general health status. *Social Science & Medicine*, 56, 203–217.
 12. Lübbecke, A., Garavaglia, G., Barea, C., & Hoffmeyer, P. Why do we need hospital-based registries? The Geneva Hip Arthroplasty Registry. EFORT publication 2010. Accessed February 26, 2018, from http://www.rpa.spot.pt/getdoc/0071e52c-7cdf-4d4e-86f7-5266c8b15bfb/The-Geneva-Hip-Arthroplasty-Registry_E-BOOK_Table-.aspx.
 13. Gandek, B., Ware, J. E., Aaronson, N. K., et al. (1998). Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project. International quality of life assessment. *Journal of Clinical Epidemiology*, 51, 1171–1178.
 14. Leplège, A., Ecosse, E., Verdier, A., & Perneger, T. V. (1998). The French SF-36 Health Survey: Translation, cultural adaptation, and preliminary psychometric evaluation. *Journal of Clinical Epidemiology*, 51, 1013–1023.
 15. Whitehouse, S. L., Lingard, E. A., Katz, J. N., & Learmonth, I. D. (2003). Development and testing of a reduced WOMAC function scale. *The Journal of Bone and Joint Surgery (British)*, 85, 706–711.
 16. Zahiri, C. A., Schmalzried, T. P., Szuszczewicz, E. S., & Amstutz, H. C. (1998). Assessing activity in joint replacement patients. *The Journal of Arthroplasty*, 13, 890–895.
 17. Perruccio, A. V., Davis, A. M., Hogg-Johnson, S., & Badley, E. M. (2011). Importance of self-rated health and mental well-being in predicting health outcomes following total joint replacement surgery for osteoarthritis. *Arthritis Care & Research*, 63:973–981.
 18. Sprangers, M. A., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: A theoretical model. *Social Science & Medicine*, 48, 1507–1515.
 19. Barclay-Goddard, R., Epstein, J. D., & Mayo, N. E. (2009). Response shift: A brief overview and proposed research priorities. *Quality of Life Research*, 18, 335–348.
 20. Ubel, P. A., Peeters, Y., & Smith, D. (2010). Abandoning the language of “response shift”: A plea for conceptual clarity in distinguishing scale recalibration from true changes in quality of life. *Quality of Life Research*, 19, 465–471.
 21. Razmjou, H., Yee, A., Ford, M., & Finkelstein, J. A. (2006). Response shift in outcome assessment in patients undergoing total knee arthroplasty. *JBJS*, 88, 2590–2595.
 22. Zajacova, A., & Dowd, J. B. (2011). Reliability of self-rated health in US adults. *American Journal of Epidemiology*, 174, 977–983.
 23. Garbarski, D., Dykema, J., Croes, K. D., & Edwards, D. F. (2017). How participants report their health status: cognitive interviews of self-rated health across race/ethnicity, gender, age, and educational attainment. *BMC Public Health*, 17, 771.