VALUE IN HEALTH 14 (2011) 316–321

available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.elsevier.com/locate/jval

Internet-Based Follow-Up Questionnaire for Measuring Patient-Reported Outcome after Total Hip Replacement Surgery—Reliability and Response Rate

Ola Rolfson, MD, PhD[a,*], *Roger Salomonsson, MSc*[b], *Leif E. Dahlberg, MD, PhD*[c], *Göran Garellick, MD, PhD*[a]

[a] *Department of Orthopaedics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden*

[b] *The Swedish Hip Arthroplasty Register, Gothenburg, Sweden*

[c] *Joint and Soft Tissue Unit, Department of Clinical Sciences, Skåne University Hospital, Lund University, Malmö, Sweden*

Keywords:
Patient-reported outcome measure
Internet questionnaire
EQ-5D
Health-care quality register

ABSTRACT

Objective: This randomized methodologic study sought to test the reliability of an Internet questionnaire and investigate the differences in response rates between traditional pen-and-paper questionnaires and Internet questionnaires for measuring patient-reported outcome after total hip replacement surgery.

Methods: From the Swedish Hip Arthroplasty Register, 2400 patients were chosen at random but stratified by age, sex, and diagnosis for inclusion in a 4-year follow-up using the health-related quality of life tool EQ-5D and visual analogue scales for pain and satisfaction. The patients were randomized to answer the follow-up model protocol either via a password-protected Internet questionnaire or via a mailed pen-and-paper questionnaire.

Results: A reliability test for the Internet follow-up instrument showed adequate correlation. However, the Internet group and the pen-and-paper group differed significantly ($P < 0.001$) with a 92% response rate in the latter and 49% in the former. Adjusted to the normal age distribution of the total hip replacement population, the Internet response rate was 34%.

Conclusions: The patient-administered Internet questionnaire alone does not give a sufficient response rate in the total hip replacement population to replace the pen-and-paper questionnaire. However, the system is reliable and could be used for measuring patient-reported outcome if supplemented with traditional pen-and-paper questionnaires for Internet nonrespondents. It is expected that this answer procedure will soon predominate in view of the general development of Internet functions. Register work may then become less resource-consuming and the results may be analyzed in real time.

Funding: This study is an analysis of the databases in the Swedish Hip Arthroplasty Register, which is funded by the Swedish National Board of Health and Welfare, the Swedish Association of Local Authorities and Regions and the County Council of the Western Region of Sweden. Biomet Europe has contributed financially to this study. The publication of the results is not contingent on the sponsor's approval. The Swedish Hip Arthroplasty Register continuously collects prospective observational data regarding all total hip replacement surgery in Sweden. The Registry holds a general approval from the Local Ethical Review Board in Göteborg (decision S 067-02, March 19, 2002).

* *Address correspondence to:* Ola Rolfson, Department of Orthopaedics, Mölndals Hospital, Sahlgrenska University Hospital, SE-431 82, Mölndal, Sweden.

E-mail address: ola.rolfson@vgregion.se.

1098-3015/$36.00 – see front matter Copyright © 2011, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

doi:10.1016/j.jval.2010.08.004

Introduction

To monitor the quality of health care systems, patient-reported outcome measures (PROMs) should be collected and analyzed before and after medical interventions, optimally with both disease-specific and generic instruments. In Sweden, the Swedish Association of Local Authorities and Regions and the National Board of Health and Welfare are increasingly demanding that health care system provide PROMs.

Traditionally pen-and-paper questionnaires are used to gather this information and data are collected, recorded, and computerized manually. The administration of pen-and-paper forms is not only time-consuming and costly but it also constitutes a risk of errors when entering data. Another disadvantage of pen-and-paper questionnaires is that missing values often compromise data quality. This is a problem in routine use in health care quality registers but also in trials managed by clinical research organizations [1,2].

The development of information technology and software along with an exponential increase in the use of the Internet now allows new modalities for collecting PROMs, and these need to be compared to the traditional method [3]. Studies have shown comparable results between pen-and-paper versions and patient-administered computer versions of questionnaires in a variety of areas [4–7].

The Swedish Hip Arthroplasty Register is a nationwide health quality-register that has been collecting data regarding all primary total hip replacement (THR) surgery and hip revision surgery performed in Sweden since 1979, and prospective PROMs were introduced in 2002 [8]. The idea of collecting PROMs originated in the mid-1990s but due to a volume problem it was not realized until safe Internet-based applications for data entry had been developed.

The aim of the Registry is not only to monitor and improve the results of THR surgery but also to improve the tools for collecting data. For this purpose the health-related quality of life (HRQoL) questionnaire, including EQ-5D, Charnley category, and pain visual analogue scale (VAS), was adapted to an Internet-based touch-screen application for preoperative use in hospital clinics. This system has been tested internally for reliability and validity and is very effective. The advantages include immediate online access to the results, no missing values, and decreased risk of systematic errors (e.g., illegible handwriting and incorrect manual registration). The system is also less laborious, with no need for manual registration of questionnaires.

According to recent reports from the World Internet Institute in collaboration with the Internet Infrastructure Foundation, 90% of Swedes older than age 16 years had access to the Internet in 2008 and 81% are using it [9,10]. The largest increase in the past few years is among younger retired people (ages 65–74 years). Similar observations have been made in other countries with high Internet use (e.g., Canada, Australia, New Zealand, and the United States) [11].

Our objective was to develop an application for an Internet-based follow-up questionnaire for PROMs, test the reliability of the Internet questionnaire, and investigate the feasibility of replacing the traditional pen-and-paper questionnaire.

Methods and Patients

Register, demographic classification, and outcome measures

All units performing THR surgery in Sweden continuously report to the Swedish Hip Arthroplasty Registry. A standardized follow-up protocol, including patient-reported outcome parameters, is used. All patients are asked to complete a 10-item questionnaire, including Charnley's functional categories (A, B, and C) [12]; a disease-specific pain VAS where zero = no pain and 100 = unbearable pain; and the generic EQ-5D measurement. This is done preoperatively and at 1 and 6 years postoperatively, with the intention to repeat the measurement at 10 years postoperatively. The EQ-5D [13] is a HRQoL instrument that evaluates subjects in five dimensions, namely mobility, self-care, usual activities, pain/discomfort, and anxiety/depression [13,14]. Each dimension is divided into three levels of inclining severity generating 243 possible response combinations. The EQ-5D can be presented as a health profile or as a global health index with a weighted total value for HRQoL. The minimum value is -0.594 and the maximum is 1.0. The Charnley categories permit correction of scores due to differing musculoskeletal comorbidity burdens. Supplementing the follow-up instrument, a VAS (zero = satisfied and 100 = dissatisfied) addressing satisfaction with the outcome of the intervention has been added. At the clinical follow-ups (1, 6, and 10 years) the questionnaire is mailed to each patient. The general response rate to the Registry after one reminder is 92%.

Patients, stratification, and randomization

This study includes patients derived from the Swedish Hip Arthroplasty Register for an "extra" 4-year follow-up after THR surgery. All patients had baseline data (from the questionnaire answered preoperatively). Four different age cohorts of 600 patients each were created: younger than 50 years at operation, 50 to 59 years, 60 to 75 years, and older than 75 years. Each cohort was also stratified regarding diagnosis and sex (80% osteoarthritis and 20% other diagnosis, 60% women and 40% men, which represents the overall distribution in the Register). Following the stratification parameters, patients were selected at random from the Register database and also randomized to answer either via the traditional pen-and-paper questionnaire or via the Internet questionnaire. In the total sample 30% lived in metropolitan regions, 35% in fairly large cities, 7% in densely populated areas, 21% in areas of intermediate population, and 8% in sparsely populated areas. The distribution between the two groups after randomization randomization was largely identical.

For the pen-and-paper questionnaire group the form and an information letter were sent with a stamped, addressed envelope enclosed. The Internet group received an information letter with a personal password and an Internet address. Nonrespondents in both groups received a reminder after 8 weeks. A second reminder was sent out after another 4 weeks offering the nonrespondents the possibility of answering via the alternative mode (crossover) (i.e., the nonrespondents of the Internet question-

naire group received a pen-and-paper form and the non-respondents of the pen-and-paper group received personal access information for the Internet questionnaire).

Additional questions about the type of Internet connection used and the experience of security when answering via Internet were added to the Internet questionnaire.

Reliability test

A reliability test was performed among the 100 first answerers to the Internet questionnaire. They were asked to reanswer the Internet questionnaire a short time later.

The Internet application

A Web-based questionnaire application (Patient Information Via Internet [PIVI]) was developed for performing online patient surveys. The database is implemented through a terms catalogue so that it can accommodate new questionnaire designs without programming or restructuring. It also permits reuse of standardized questions, helpful when using EQ-5D or other strictly standardized forms.

Patients (or any other respondents) are allocated through an administrative page by entering the number to be allocated. The application then generates a list of serial numbers and corresponding passwords (eight characters long, six letters, and two digits in mixed case). This list can, and should, be merged with patient data and constitutes the link between the PIVI-generated serial number and, for example, a personal identity number. The advantage of this approach is that the PIVI database contains only anonymous data. All sensitive data are managed offline.

A patient receives a printed letter with a Web address and a password, together with instructions on how to open the questionnaire in a Web browser where the survey is outlined. A patient logs in with the enclosed password, answers the questions all on one Web page, and submits the questionnaire. Any unanswered mandatory questions are marked with a red frame, clearly stating what needs to be corrected before the questionnaire can be submitted. Once submitted, the questionnaire cannot be updated by a patient.

Survey results can be retrieved continuously in many ways through the Web service. For example, with Microsoft Excel (Microsoft Corp., Redmond, WA, USA), they are delivered in XML format through a strongly encrypted connection, which in turn is established with proper authorization to the Web service. When the survey is completed, this retrieved data can be merged with the rest of the patient data through the PIVI-generated serial number (Fig. 1).

Statistical methods

Randomization was carried out with Tave's Minimization [15]. Probability values less than 0.05 represented a significant difference. The Mann-Whitney U test, Fisher's exact test, or logistic regression was used to evaluate any difference between the groups as appropriate. Patient-reported outcome measurements are presented as the mean and the median with standard deviation. Spearman's correlation was used for reliability tests. SPSS version 17.0, 2008 (SPSS Inc, Chicago, IL, USA) was used for all statistical analyses.

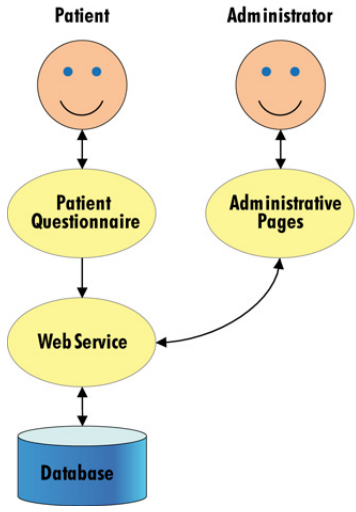

Fig. 1 – Schematics of the Patient Information Via Internet application. Questionnaires dynamically generated from the database; that is, no programming is required to change survey content. When respondents are allocated, unique respondent identifiers are returned together with strong passwords. This means that respondents are not identifiable from the data stored in the database alone. Built with ASP.NET 2.0/C# and MS Access.

Results

After stratification and randomization there was a delay of 12 months due to difficulties in obtaining the proper certificate to set up and publish the Internet questionnaire in the university's Web domain. At the time of mailing, 49 of the patients randomized to Internet questionnaire group and 61 in the pen-and-paper questionnaire group had either died or could not be found in the national population register (due to emigration, identity secrecy, or some other reason). The mean time from operation to answering the 4-year follow-up protocol was 47.6 months (median 48 ± 7.1 months).

After the first reminder, the 92% response rate in the pen-and-paper questionnaire group (exactly the response frequency the Register has had in earlier measurements) significantly differed from the 49% response rate in the Internet questionnaire group ($P < 0.01$, based on Fisher's exact test) (see Table 1). The total frequency in the Internet group rose to 81% when the nonrespondents were offered the possibility to answer via the pen-and-paper questionnaire. The difference in response frequency between the four age groups was not significant in the pen-and-paper questionnaire group ($P = 0.093$, based on logistic regression) but was significant in the Internet group ($P < 0.001$, based on logistic regression). Response rates declined with increasing age in the Internet group (Table 1).

Fifty-two percent of men in the Internet questionnaire group answered and 93% in the pen-and-paper group answered. The

Table 1 – Response rates after first reminder; with the crossover opportunity, and with the combination of answering modes in different age categories according to stratification.

Age category	n	Response rate after first reminder (%)	Response rate with crossover (%)	Total response rate (%)
Pen-and-paper questionnaire				
<50 y	293	88	2	90
50–59 y	296	92	1	93
60–75 y	287	97	0	97
>75 y	275	91	0	91
Total	1151	92	1	93
Internet questionnaire				
<50 y	294	71	15	86
50–59 y	289	62	27	89
60–75 y	288	37	40	77
>75 y	268	23	48	71
Total	1139	49	32	81

corresponding response frequency among the women was 47% and 92%, respectively. The differences in response rates between men and women were not significant (P = 0.09 and P = 0.74 respectively, based on Fisher's exact test).

As presented in Table 2, there were no significant differences in response rates within the two groups with regard to population density (P > 0.05, based on logistic regression).

Adjusting from age-cohort grouping to normal age distribution in the Swedish Hip Arthroplasty Register, the Internet response rate was computed to 34% and the pen-and-paper version to 94% (Table 3).

Sixty-four percent of those who answered via the Internet used their own computers, 72% used a broadband connection, 21% used a dial-up connection, and the rest did not know what Internet connection they used. The majority (67%) stated that they felt secure in answering the questionnaire via the Internet, whereas 12% felt insecure, and 21% had no view on this.

Table 2 – Response rate after first reminder in randomization groups divided into subgroups according to population density. A logistic regression including age and sex was used to identify any differences in response rates with respect to population density within the two groups.

Population density area	Pen-and-paper questionnaire		Internet questionnaire	
	Response rate (%)	Distribution	Response rate (%)	Distribution
Metropolitan areas	89	0.31	52	0.29
Fairly large cities	92	0.34	50	0.35
Densely populated areas	92	0.055	47	0.08
Areas of intermediate population density	94	0.21	42	0.20
Sparsely populated areas	91	0.085	53	0.08
	P = 0.22	1	P = 0.24	1

Table 3 – Response rates after first reminder by randomization group compared to the ordinary response rate to the Swedish Hip Arthroplasty Register. Because the randomization groups were stratified with respect to age, the expected response rate was calculated by adjusting to the normal age distribution of the Register.

Age distribution in cohorts and register	Internet questionnaire (n = 1151)	Pen-and-paper questionnaire (n = 1139)	Register (n = 12,672)
<50 y	0.258	0.255	0.044
50–59 y	0.254	0.257	0.135
60–75 y	0.253	0.249	0.488
>75 y	0.235	0.239	0.333
Total	1	1	1
Response rate after first reminder (%)	49	92	92
Age-adjusted response rate (%)	34	94	NA

NA, not applicable.

All respondents' answers in the two groups are compared in Tables 4 and 5. Mean EQ-5D was lower among the respondents in the pen-and-paper group than among those in the Internet group (P = 0.002, based on Mann-Whitney U test). Adjusting for age and sex removed this difference in EQ-5D index (P = 0.15, based on logistic regression). Mean pain VAS was higher in the Internet group than in the pen-and-paper group (P = 0.018, Mann-Whitney U test) and this difference remained after adjusting for age and sex (P = 0.002, based on logistic regression). Despite this adjustment, no significant differences between the groups with regard to satisfaction VAS were detected (P = 0.132, based on Mann-Whitney U test and P = 0.53, based on logistic regression).

The EQ-5D index differed significantly when comparing the patients who responded as they were initially randomized. Pen-and-paper answerers in the pen-and-paper group reported lower EQ-5D index (mean 0.74, median 0.80, and SD 0.28) than Internet answerers who reported via the Internet questionnaire (mean 0.80, median 0.85, and SD 0.27) (P < 0.001, based on Mann-Whitney U test). When age, sex, and EQ-5D index were included in the same regression analysis model, the difference was still significant (P = 0.001, based on logistic regression). Pain VAS was lower (mean 15, median 8, and SD 19) among the pen-and-paper respondents among than the Internet-respondents (mean 19, median 9, and SD 21) (P = 0.001, based on Mann-Whitney U test). A similar adjustment for age and sex showed that the difference in pain VAS persisted (P = 0.001, based on logistic regression). Satisfaction VAS did not differ between the groups (pen-and-paper mean 17, median 9, and SD 22, Internet mean 16, median 92, and SD 21) (P = 0.94, based on Mann-Whitney U test), and adjusting for age and sex did not affect this observation (P = 0.94, based on logistic regression).

The probability of responding in any way for the two alternatives, start with pen-and-paper questionnaire or start with Internet questionnaire, is illustrated in Fig. 2. The logistic regression model from which these probability curves are created includes age and randomization group as dependent variables and any response as independent variable.

Table 4 – Comparisons of EQ-5D, pain visual analogue scale (VAS), and satisfaction VAS among all respondents (regardless of final mode of response) in the two randomization groups. Results are presented as the mean and the median with standard deviation (SD). The P value was computed with Mann-Whitney U test.

	Pen-and-paper group		Internet group		*P*
	Mean (SD)	Median	Mean (SD)	Median	
EQ-5D	0.74 (0.28)	0.80	0.76 (0.28)	0.80	0.002
Pain VAS	15 (18)	8	18 (21)	10	0.018
Satisfaction VAS	17 (22)	9	17 (21)	9	0.132

SD, standard deviation.

Reliability test

To test the reliability of the questionnaire the 100 patients who answered the Internet questionnaire first were asked to reanswer it 3 weeks after the first mail was delivered. Seventy patients completed the retest form. Spearman's correlation for the EQ-5D index was 0.82, for VAS pain 0.83, and for VAS satisfaction 0.63.

Discussion

At present the Swedish Hip Arthroplasty Register handles approximately 20,000 postoperative questionnaires yearly. This burden will successively increase as the stock of patients scheduled for long-term follow-up grows. Fully implemented, the follow-up program will generate about 35,000 questionnaires yearly. Manual handling of these questionnaires requires exhaustive work. The National Board of Health and Welfare and the Swedish Association of Local Authorities and Regions require that PROMs for individual patients should be reported for all national health quality registers in Sweden (at present 71). However, partly due to lack of resources, most registers currently do not include PROMs. Thus there is a great need for an effective and safe system for collecting PROMs, and this is offered via the Internet questionnaires. Further, a computerized answering system is scientifically and methodologically attractive because it supplies better data quality with no missing values or systematic errors. Such a system also facilitates immediate feedback and online analysis to the participating clinics.

Table 5 – Beta coefficient (β), standard error (SE), and P value for three different logistic regression models comparing the two randomization groups among patients who answered the questionnaire, with respect to EQ-5D, pain visual analogue scale (VAS), and satisfaction VAS respectively. Age at primary operation, and sex, were taken into account in the comparisons.

Variable	β	SE	*P*
EQ-5D	0.237	0.163	0.15
Constant	0.110	0.284	
Sex	−0.0145	0.0922	0.88
Age at operation	−0.00687	0.00318	0.031
Pain VAS	0.007	0.002	0.002
Constant	0.177	0.249	
Sex	−0.016	0.092	0.860
Age at operation	−0.007	0.003	0.030
Satisfaction VAS	0.001	0.002	0.533
Constant	0.307	0.244	
Sex	−0.021	0.092	0.817
Age at operation	−0.007	0.003	0.022

Our study shows that an Internet questionnaire is at present not sufficient to replace the pen-and paper form for the THR population. The low response rate could partly be explained by the fact that the eight-digit password was complicated with upper- and lower-case letters mixed with numerals. Similarly, the Internet link provided was complicated. A simplified Web address and possibility to log-in via digital E-identification for secure E-service may increase the response rate.

This study also made us aware that some Internet-group patients were offended by being asked to answer via the Internet, and this may explain the lower response rate in the Internet group even after the crossover.

The insignificant sex difference in Internet response rates among these age groups is consistent with recent observations showing small sex-related differences in the use of Internet [10]. From a sex-equality perspective the use of the Internet questionnaire seems to be nonproblematic. Similarly, there were no statistical differences in response rates depending on population density. However, there were significant differences between the different respondents. The Internet respondents were not only younger but also reported higher HRQoL regardless of age. This suggests that healthier patients in the THR population are more frequent Internet users. An interesting finding is that the Internet respondents reported more pain even after adjustment for age and sex. This did not apply to satisfaction for which the same design of visual analogue scale is used. Our interpretation is that the Internet respondents in the THR population had more active lifestyles and were then more likely to be hampered by the limitations of their artificial joint.

Analysing the content of the questionnaires after combining the answering modalities indicated no significant differences for satisfaction or EQ-5D index after adjusting for age and sex. This rather suggests that the content of the answer is not affected by the type of administration, only the tendency to answer. However, pain remained significantly higher among all Internet group respondents even after adjusting for age and sex. This implies that the proportion of the THR population that would normally respond to the traditional questionnaire but in fact responded neither to the Internet questionnaire nor the crossover opportunity typically report lower

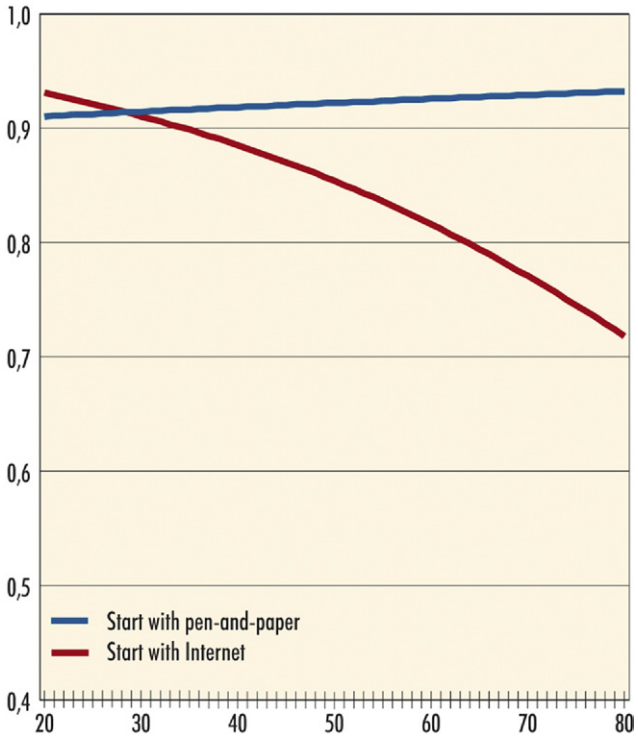

Fig. 2 – Probability of final response based on a logistic regression model including age at operation and randomization group. The curves for the probability to respond for the two different ways of administrating the questionnaire cross at age 28 y at operation.

pain at follow-up. In line with the discussion above about the Internet respondents, the Internet group nonrespondents who would normally respond may tend to be less active and demanding and therefore experience less discomfort from the hip replacement.

Strictly using an Internet-based method for collecting PROMs in the THR population would bias the results because older patients and those with more severe comorbidities do not respond. For this reason it is at present of great importance to give the opportunity to combine the methods of answering.

The Internet questionnaire system requires an initial investment. Still we think that at present a combined system, where the patients may answer either via Internet or in the traditional way, is feasible. The continuous but slowing increase in Internet use suggests a future improved response rate among the THR population using the Internet application. However, even at limited Internet response rates, the combined answering system appears cost-saving for large volumes. Accordingly, the Swedish Hip Arthroplasty Register plans shortly to introduce a system that will allow the patient a choice of answering mode.

Conclusions

This patient-administered Internet questionnaire alone gave an insufficient response rate in the THR population to replace pen-and-paper questionnaires. However the system is reliable and could be used for measuring patient-reported outcomes if supplemented with traditional pen-and-paper questionnaires for Internet nonrespondents. In view of the general development of Internet functions, it is expected that this dual answer procedure will soon predominate. Register work may then become less resource-consuming and the results may be analyzed in real time.

REFERENCES

[1] Stone AA, Shiffman S, Schwartz JE, et al. Patient non-compliance with paper diaries. BMJ 2002;324:1193–4.
[2] Wilson AS, Kitas GD, Carruthers DM, et al. Computerized information-gathering in specialist rheumatology clinics: an initial evaluation of an electronic version of the Short Form 36. Rheumatology (Oxford) 2002;41:268–73.
[3] Gutteling JJ, de Man RA, Busschbach JJ, et al. Overview of research on health-related quality of life in patients with chronic liver disease. Neth J Med 2007;65:227–34.
[4] Velikova G, Wright EP, Smith AB, et al. Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. J Clin Oncol 1999;17:998–1007.
[5] Kleinman L, Leidy NK, Crawley J, et al. A comparative trial of paper-and-pencil versus computer administration of the Quality of Life in Reflux and Dyspepsia (QOLRAD) questionnaire. Med Care 2001;39:181–9.
[6] Cook AJ, Roberts DA, Henderson MD, et al. Electronic pain questionnaires: a randomized, crossover comparison with paper questionnaires for chronic pain assessment. Pain 2004;110:310–7.
[7] Saleh KJ, Radosevich DM, Kassim RA, et al. Comparison of commonly used orthopaedic outcome measures using palm-top computers and paper surveys. J Orthop Res 2002; 20:1146–51.
[8] Herberts P, Malchau H, Garellick G. The Swedish Hip Arthroplasty Register Annual Report 2003. Department of Orthopaedics, Sahlgrenska University Hospital. Gothenburg, 2004. Available from: http://www.shpr.se/Libraries/Documents/AnnualReport2004-EN.sflb.ashx [Accessed January 6, 2011].
[9] Findahl O. Svenskarna och Internet 2008 (The Internet in Sweden 2008). World Internet Institute, 2008. Available from: http://www.wii.se/publicerat/rapporter.html [Accessed January 6, 2011].
[10] Findahl O. Svenskarna och Internet 2009 (The Internet in Sweden 2009). World Internet Institute, 2009. Available at: http://www.wii.se/publicerat/rapporter.html [Accessed January 6, 2011].
[11] World Internet Project, International Report 2009. 2008.
[12] Charnley J. Low friction arthroplasty of the hip: theory and practice. Berlin; Springer-Verlag, 1979.
[13] The EuroQol Group. EuroQol—a new facility for the measurement of health related quality of life. Health Policy 1990;16:10.
[14] Carr-Hill RA. Health related quality of life measurement—Euro style. Health Policy 1992;20:321–8.
[15] Altman DG, Bland JM. Treatment allocation by minimisation. BMJ 2005;330:843.