The Spine Journal 19 (2019) 1154–1161

Clinical Study

Reliability of the Neck Disability Index and Japanese Orthopedic Association questionnaires in adult cervical radiculopathy and myelopathy patients when administered by telephone or via online format

Gaurang Gupte, BS, Colleen M. Peters, MA, Jacob M. Buchowski, MD, MS, Lukas P. Zebala, MD*

Department of Orthopedics – Spine Surgery, Washington University in Saint Louis, 660 S. Euclid Ave, Campus Box 8233, St. Louis, MO 63110, USA

Received 30 October 2018; revised 1 March 2019; accepted 1 March 2019

Abstract

BACKGROUND: The internal validity of long-term studies is significantly affected by the high loss to follow-up in the spine surgery population (>20%). Phone and email-based administration of patient-reported outcomes instruments is a less cumbersome approach for increasing response rates and assessment frequency while potentially decreasing follow-up burden on patients and physicians.

PURPOSE: This study sought to validate simultaneous administration of the Neck Disability Index (NDI) and Japanese Orthopedic Association (JOA) questionnaires in patients with either cervical myelopathy and/or radiculopathy.

STUDY DESIGN/SETTING: This is a single-center, randomized crossover phone and email validation of legacy outcome measures for cervical myelopathy and/or radiculopathy patients.

PATIENT SAMPLE: The study included nonsurgical along with pre- and postsurgical cervical myelopathy and/or radiculopathy patients presenting to a tertiary spine care center.

OUTCOME MEASURES: NDI and JOA.

METHODS: Two-hundred and six patients (mean age: 58.5 years) were randomized in a 1:4 ratio to either email completion of the NDI and JOA before or after in-office completion, or to phone completion before or after in-office completion. An interval of 1 to 4 weeks was established between the administration of questionnaires. The difference between written in-office and corresponding email and phone versions was assessed with a paired *t* test. Homogeneity was assessed using intraclass correlation coefficients. Test-retest reliabilities were independently examined for postoperative patients (n=145). Recall bias was assessed in postoperative patients by calculating intraclass correlation coefficients for those with days between assessments lesser than the mean and for those greater. Differences in response rates between phone and email versions were assessed with McNemar's and Cochran-Mantel-Haenszel tests.

RESULTS: There was no significant difference between email and in-office versions (n=85) of the NDI (p=.17, Mean Paired Difference=1.34) and JOA (p=.64, Mean Paired Difference=0.11). No significant difference was seen between phone followed by in-office administration (n=32) of the NDI (p=.88, Mean Paired Difference=0.22) and JOA (p=.38, Mean Paired Difference=-0.22), nor between in-office administration followed by phone (n=44) for the NDI (p=.10, Mean Paired Difference=2.79) and JOA (p=.37, Mean Paired Difference=0.27). Intraclass coefficients (ICCs) of the email versions of the NDI and JOA were 0.88 and 0.78, respectively; of the phone-before-office versions of the NDI and JOA were 0.91 and 0.82; of the office-before-phone versions were 0.86

FDA device/drug status: Not applicable.

Author disclosures: ***GG:*** Nothing to disclose. ***CMP:*** Nothing to disclose. ***JMB:*** Royalties: Globus (F), K2M (D), Wolters Kluwer (A). Consulting: Globus (B), K2M (A), Medtronic (B), Stryker (C). Speaking/Teaching Arrangements: Globus (C). Fellowship Support: AO Spine (E per 3 academic years; paid directly to isntitution), OMeGA (D per 3 academic years; paid directly to institution). ***LPZ:*** Nothing to disclose.

* Corresponding author. Department of Orthopedics – Spine Surgery, Washington University in Saint Louis, 660 S. Euclid Ave, Campus Box 8233, St. Louis, MO 63110, USA.
Tel.: +013147474950.

E-mail addresses: zebalal@wustl.edu, lzebala@wustl.edu, akersa@wustl.edu (L.P. Zebala).

https://doi.org/10.1016/j.spinee.2019.03.002

and 0.78. Similarly, strong intraclass correlation coefficients indicative of a lack of recall bias were found for "In-Office" and external assessments completed by postsurgical patients with days between assessments lesser and greater than the mean (ICC range 0.63−0.92). No significant difference was seen in completion rates between email and in-office questionnaire completion (p=.13) and phone-before-office and in-office questionnaire completion (p=.31). However, a significant difference was found in completion rates for phone-after-office questionnaires (p<.001).

CONCLUSIONS: Administration of the NDI and JOA over phone and email in patients with cervical myelopathy or radiculopathy is valid with strong test-retest reliability and internal consistency. Phone and email administration of the NDI and JOA can reduce nonresponse rates and decrease the burden of follow-up and data acquisition. Follow-up phone reminders increase the response rate for administration of these patient-reported outcomes (PROs) via email.

Keywords: Cervical myelopathy; Cervical radiculopathy; Neck Disability Index; Outcomes; mJOA; Telephone administration; Email administration

Patient-reported outcomes (PROs) are crucial for tracking the multifaceted impact of spinal disease on patient functionality and quality of life [1]. Two key patient outcome tools, the Neck Disability Index (NDI) and the Japanese Orthopedic Association (JOA) scale, have been developed for cervical spine surgery patients. The NDI is a well-validated PRO tool adopted widely around the world that has been translated into multiple languages [2]. Originally developed in 1991 and modeled after the Oswestry Low Back Pain Index and the Roland-Morris Low Back Pain Questionnaire, its validity and consistency have been established through extensive study. It has been used as a primary measure for the development and validation of new questionnaires such as the Copenhagen Neck Functional Disability Scale, Patient-Specific Scale, Neck Pain and Disability Scale, and Functional Rating Index [3].

Similarly, the JOA scale is used to assess neurological function impairment in patients with degenerative cervical myelopathy. It tracks the severity of sensory and motor deficits. Like the NDI, its validity and internal consistency have also been assessed in multiple studies [4].

Administration of the NDI and JOA is most commonly performed using in-office paper-based questionnaires. These assessments can utilize a substantial portion of a patient's visit time and data entry is required to transcribe data from paper copies into an electronic database. This can be detrimental in the context of long-term studies, in which loss-to-follow-up rates can be as high as 20% to 50% after 5 years [5,6]. Being able to administer questionnaires to patients outside of the office environment would allow clinicians and researchers to reach patients unlikely to return for follow-up visits and thus improve response rates.

As a result, several groups have developed and validated PRO administration over phone and email to address these issues [7−9]. However, phone administration still requires mutual time availability for both patients and providers, manual data entry into an electronic record system, and calling outside of normal office hours. In contrast, electronic administration offers greater patient flexibility and reduced burden on provider office and research staff. Patients can not only independently choose a time to complete assessments but also pause and resume them at their own pace. Direct input of questionnaire results into an electronic database eliminates time-consuming manual data transcription. Computerized assessments can also automatically handle complex questionnaire skip patterns and require patients to complete all entries in order to proceed [10]. These advantages prompted us to establish the reliability of email and phone administration of the NDI and JOA in cervical spine patients through a prospective study. Our work was based on a previous study examining phone-based administration of the Scoliosis Research Society-22 (SRS-22) and Oswestry Disability Index (ODI) for adult spinal deformity patients [11].

Materials and methods

Study design

The study incorporated a two-arm randomized crossover design to examine for differences in responses for phone and email administration of NDI and JOA compared to office administration for cervical myelopathy and radiculopathy patients recruited at a single tertiary center. Based on the aforementioned study validating phone administration of the SRS-22r and ODI in adult scoliosis spine surgery patients, a power analysis was conducted to detect a 0.6 standard difference between in-office and either email or phone versions for a two-tailed significance level of 0.05 and power of 0.8 [11].This yielded 44 required patients for each arm and 88 for both arms. To account for dropout, 100 patients were slated for enrollment.

Patient selection

After receiving Institutional Review Board approval, cervical myelopathy and radiculopathy patients scheduled for clinic visits with two fellowship-trained spine surgeons were identified in the electronic medical record system at least four weeks prior their clinical visit.

Inclusion criteria consisted of nonsurgical along with pre- and postsurgical cervical myelopathy and/or radiculopathy patients over 18 years of age. This consisted of two categories of patients: "New" and "Returning" patients. The former had either been prescreened by an orthopedic nurse practitioner at our clinic knowledgeable in making this diagnosis, diagnosed by another orthopedic surgeon at our institution, or operated on by a surgeon external to our institution. "Returning" patients had been diagnosed at an earlier visit at our clinic. This category consisted of "nonoperative" patients whose care was being managed without surgery, "preoperative" patients that completed both assessments prior to surgery, and "postoperative" patients. These returning patients had already been diagnosed via both physical examination and radiography. With the exception of patients referred to clinic by our orthopedic nurse practitioner, new patients had already been diagnosed through a physical exam and radiography by another orthopedic surgeon at our institution or had been previously operated on by a surgeon external to our institution. Exclusion criteria consisted of patients with cervical tumors, trauma, non-English speakers, or patients unable to complete both questionnaires within the study timeframe.

Eligible patients at each office visit date were individually randomized to one of four groups prior to consent. "Before Office" patients would be contacted and recruited 1 to 4 weeks prior to their office visit date to obtain verbal consent for participation. "After Office" patients would be recruited and consented in-office during their follow-up visit. Patients in each arm ("Before Office" or "After Office") would receive either a phone or email assessment. Using a random number generator, a value of 1 was assigned to "Assessment Before Office" and 2 to "Assessment After Office." Assignment was performed pair wise. The number generator was run for every odd-numbered patient in the sublist. If the first patient was a 1, the second patient was assigned the opposite value of 2 and vice versa to maintain balance. The number generator was then run for the third patient onwards. A similar process was done to decide whether to email or call patients. A 3 was assigned to "Phone" and 4 was assigned to "Email". Each patient was assigned a value independent of whether they were classified into the "Assessment Before Office" or "After Office" group. If a given group was "full," it was no longer included in randomization. Any patient lacking email was assigned to the "Phone" group.

Consenting patients received either an email survey or phone survey as appropriate. Emails were sent out through REDCap, a secure, web-based application for research study data capture. It provided an intuitive interface for validated data entry; audit trails for data manipulation and export; automated export of data to statistical software; and procedures to import data from external sources [12].

"Before Office" patients had to complete assessments 1 to 4 weeks prior to their office visit date. Likewise, "After Office" patients had to complete assessments 1 to 4 weeks after their office visit. This interval was chosen based on the following. A test-retest interval of 2 to 14 days is well known in the psychometric community to be long enough to mitigate patient recall of previous questionnaire answers and short enough to occur before significant change in pathology [13,14]. On the other hand, a previous study validating phone administration of the SRS-22r and ODI utilized a 4-week time period as a cutoff. It was felt that patient pathology could potentially significantly change by 4 weeks [11]. In combination with the 1-week time period, a 4-week cutoff enabled for adequate time to gather patient responses. To further validate this time frame, we consulted our departmental statistician who concurred that a 1 to 4 week time frame was appropriate and would minimize bias risks in the study.

Any patient who had not completed an email survey was called after the first week and every 2 days after. Likewise, "phone survey" patients were contacted every 2 days. At most 3 attempts were made over the phone to reach patients. No patients were asked to add the clinic phone and/or email to their "safe caller" or "safe sender" lists. Patients receiving an assessment before their office visit were called around the time and day of week that they had picked up our first call asking them to participate. Those receiving an email or phone assessment after their office visit were asked when the best time and day of the week to call would be. This "primed" our patients to expect our call during a certain timeframe.

All calls were made from a landline at our institution to minimize the area code screen and to reach patients expecting a call with an institutional extension. Emails were sent through REDCap with an email account affiliated with our institution. Subject headers were formatted as such to improve recognition: "[Institution Name Orthopedic Department] – NDI + JOA Reliability Study"

Statistical analysis

To assess differences in scores between in-office and out-of-office administrations (phone or email) of the NDI and JOA, a paired sample *t* test was calculated. Alpha levels were defined at 0.05. Statistical analysis was conducted using R [15].

Test-retest reliability was calculated using a intraclass correlation coefficient (ICC) with a two-way random effects model with absolute-agreement to assess homogeneity between in-office and corresponding external versions of the NDI and JOA. This was performed with the "IRR" package for R with parameters "two way," "agreement" and "single" [16].

Since the study population included both surgical and nonsurgical patients, test-retest reliability was independently examined for postoperative patients, who made up the bulk of the study population. Too few non-postoperative patients were present for stratified analysis.

To assess whether recall bias existed amongst patients with a shorter duration between assessments in contrast to those with a longer duration, test-retest reliabilities were calculated and compared between post-operative patients

with a lower mean number of days between assessments and those with a greater mean number.

Version preferred was assessed using McNemar's test to examine difference in completion rates for in-office, phone and email versions along with a Cochran-Mantel-Haenszel test for assessing consistency between groups.

Results

Due to over enrollment, 161 patients completed both the in-office and external survey (85 email, 76 phone). As shown by the patient demographics in Table 1, the "Before Office" and "After Office" email groups were combined to perform a matched pair *t* test since no significant difference was found regarding age (p=.92) and gender (p= .38). However, the phone groups were analyzed separately since they differed in age (p=.02) but not gender (p= .42). As quantified by Table 2, returning postoperative patients were the largest patient population (145 out of 161) in this study.

As indicated by the assessment results in Table 3, no significant difference was detected between either the email versions (n=85) of the NDI (Mean Paired Difference=1.34, p=.17) or the JOA (Mean Paired Difference=0.11, p=.64). Both email versions demonstrated excellent homogeneity with ICCs of 0.88 and 0.78, respectively.

As highlighted in Table 4, no significant difference was seen between "Phone before Office" and "In-Office" administration (n=32) of the NDI (Mean Paired Difference=0.22, p=.88) and JOA (Mean Paired Difference=−0.22, p=.38). Excellent homogeneity was observed with ICCs of 0.91 and 0.82 for the NDI and JOA, respectively. Furthermore, as showcased in Table 5, no significant difference was seen between "In-Office" and "Phone-After-Office" administration followed by phone (n=44) for the NDI (p=.10, Mean Paired Difference=2.79) and JOA (p=.37, Mean Paired Difference=0.27). Excellent ICC homogeneity was observed for the NDI with 0.86 and JOA at 0.78. Separate analysis of the postsurgical patient subset in Table 6 found excellent ICC homogeneity and no significant difference in between "In-Office" and external assessment for patients with days between assessments lesser than and greater than the mean days between assessments.

As demonstrated in Table 7, no significant difference was seen in completion rates between email and in-office questionnaire completion (p=.13). Similarly, as indicated in Table 8, no significant difference was found for completion rates for phone-before-office and in-office questionnaire completion (p=.31). However, as highlighted in Table 9, a significant difference was found in completion

Table 1
Patient characteristics for phone and email groups

	Phone group			Email group		
	Phone first	Office first	p Value	Email first	Office first	p Value
Age Mean±SD	63.0±10.1	56.3±12.5	.02*	58.1±11.8	57.9±10.3	.92
Gender						
Male	13	22	.42	23	30	.38
Female	19	22		17	15	

Table 2
Patient visit type by assessment type

Patient visit type	Email (combined)	Phone then office	Office then phone	Total
New (N.P referral)	1	0	1	2
New (intrainstitution surgeon referral)	1	0	2	3
New (previous surgery by outside surgeon)	1	0	0	1
Returning no-operation	4	2	0	6
Returning preoperation	3	0	1	4
Returning postoperation	75	30	40	145
Total	85	32	44	161

Table 3
Summary of paired *t* test for the differences between the office and corresponding email versions of the NDI and JOA

	N	Mean days between assessments	Mean score	SD	Mean of paired differences	p Value	SEM	ICC (95% CI)
Email NDI	85	13.55	28.28	19.47	1.34	.17	1.01	0.88 (0.81–0.92)
Office NDI			28.54	17.77				
Email JOA			14.44	2.80	0.11	.64	0.20	0.78 (0.68–0.85)
Office JOA			14.46	2.63				

Table 4
Summary of paired *t* test for the differences between the phone-before-office and corresponding in-office versions of the NDI and JOA

	n	Mean days between assessments	Mean score	SD	Mean of paired differences	p Value	SEM	ICC (95% CI)
Phone (preoffice) NDI	32	14.00	29.81	19.79	0.22	.88	1.48	0.91 (0.82–0.96)
Office NDI			29.59	19.24				
Phone (preoffice) JOA			13.97	2.13	−0.22	.38	0.24	0.82 (0.67–0.91)
Office JOA			14.19	2.43				

Table 5
Summary of paired *t* test for the differences between the phone-after-office and corresponding in-office versions of the NDI and JOA

	N	Mean days between assessments	Mean score	SD	Mean of paired differences	p Value	SEM	ICC (95% CI)
Phone (postoffice) NDI	44	14.27	35.64	21.11	2.79	.10	1.27	0.86 (0.75–0.92)
Office NDI			38.42	19.62				
Phone (postoffice) JOA			13.45	2.97	0.27	.37	0.21	0.78 (0.63–0.87)
Office JOA			13.73	2.99				

Table 6
Summary of paired *t* test and intraclass correlation coefficients for returning postoperative patients by assessment type and days between assessments

	Days between assessments	Measure	Number of patients	p Value	ICC (95% CI)
Email	<14 Days	NDI	44	.06	0.86 (0.75–0.92)
		JOA		.33	0.79 (0.65–0.88)
	>=14 Days	NDI	31	.73	0.91 (0.82–0.95)
		JOA		.34	0.63 (0.36–0.80)
Phone then office	<14 Days	NDI	18	.45	0.89 (0.74–0.96)
		JOA		.66	0.81 (0.57–0.93)
	>=14 Days	NDI	12	.28	0.92 (0.74–0.97)
		JOA		.79	0.88 (0.65–0.97)
Office then phone	<14 Days	NDI	23	.12	0.92 (0.83–0.97)
		JOA		.57	0.74 (0.48–0.88)
	>=14 Days	NDI	17	1.00	0.80 (0.53–0.93)
		JOA		.57	0.79 (0.52–0.92)

Table 7
Completion rates for email assessments

	Completed email questionnaire	No complete email questionnaire	Total	McNemar yest statistic p value
Completed in-office questionnaire	85 (81%)	13 (12%)	98	.13
No complete in-office questionnaire	7 (7%)	0 (0%)	7	
Total	92	13	105	

Table 8
Completion rates for phone-before-office assessments

	Completed phone questionnaire	No complete phone questionnaire	Total	McNemar test statistic p value
Completed in-office questionnaire	32 (89%)	2 (6%)	34	.31
No complete in-office questionnaire	2 (6%)	0 (0%)	2	
Total	34	2	36	

rates for phone-after-office questionnaires (p<.001) that were lower (70%) than those for the email group (81%) and the phone-before office group (89%). A Cochran-Mantel-Haenszel test found these groups to be consistent (p=.18).

Discussion

Administration of the NDI and JOA is currently limited to in-office assessments during patient follow-up visits. This is a significant limitation for long-term studies, which

Table 9
Completion rates for phone-after-office assessments

	Completed phone questionnaire	No complete phone questionnaire	Total	McNemar test statistic p value
Completed in-office questionnaire	44 (70%)	17 (27%)	61	<.001
No complete in-office questionnaire	2 (3%)	0 (0%)	2	
Total	46	17	63	

face significant loss to follow-up potentially resulting in nonresponder bias. Although the extent of this bias is controversial as several studies have shown no statistically significant difference in outcomes between responders and nonresponders, others have found nonresponders to be younger and experience worse outcomes [17–19]. Reliance on in-office paper copy administration is also cumbersome and burdens office and research staff. As an alternative, phone and email administration of several PROs has been tested and validated.

In this study, we demonstrate the reliability of phone and email administration of the NDI and JOA in myelopathy and radiculopathy patients. Our validation methodology is consistent with that of the International Society of Patient Outcomes Research in that modifications were minor, equivalence testing incorporating ICCs was performed, and a random crossover study design was utilized [20]. No statistically significant difference was found between in-office and email or phone administration of the NDI and JOA. Both phone and email versions of the NDI showcase strong ICCs that are similar to the originally reported two-day test-retest reliability of the NDI at 0.89 [3]. While the test-retest reliability of the JOA had not yet been established in prior literature to the best of our knowledge, its excellent ICCs are similar to those found during the phone validation of the SRS-22 and ODI and email validation of the SF-36 and HAQ [21,10,11]. Thus, either email or phone administration can be performed for patients unable to attend their scheduled follow-up visit.

Our results differ from prior literature in several respects. While our overall phone response rate (77%) is congruous with that of previous studies, our email response rate is significantly higher (88%) [22,23,7]. This contrasts previous studies in orthopedic patient populations exhibiting lower email assessment response rates ranging from 24% to 42% [7,23,24,8]. We attribute our high response rate to the reminder phone calls (up to 3) to patients who had consented to answering an email or a phone-call along with the one-click simplicity of REDCap enabling patients to directly access surveys on the web portal via either mobile or PC through a unique link emailed to them instead of a login prompt. These reminder calls enabled us to directly talk with patients, which thus increased the persuasive power of our reminder calls. If patients indicated that our email may have been "lost" in their inbox or spam folder, we resent it and bumped it to the top of their inbox. If patients had still not responded to the email, we called them again two days later. Furthermore, as found by Rolfson et al., response rates to their email questionnaire were only 49% after the first reminder but increased to 81%, as we found, after sending patients a reminder letter over "regular" mail [24]. It seems that multiple reminders are necessary to obtain a high response rate. The aforementioned studies we found with low email response rates did not use a multiple reminder system. Email administration was also significantly more convenient for patients and research staff. Patients were able to answer questionnaires at nonbusiness hours. Direct transfer of assessment results into our electronic database precluded additional manual data entry. However, a potential limitation of REDCap administered email questionnaires is that they can be started and stopped, whereas in-office administration questionnaire is performed as a single-setting event at a specific time. This can degrade assessment validity for certain clinical contexts and outcome data. Although we did not calculate the research staff time cost per completed assessment, our experience with the ease of email over phone roughly coincided with the analysis of Harewood et al. who quantified research staff time cost per completed e-mail assessment at $0.71 and for each completed phone survey at $2.08 [22]. While we did not keep track of the number of phone-only patients in our study, this group was small (<5) and there were likely phone-only patients in the phone-first group that we were not aware of since this status was only ascertained at the time of office visit for patients who would receive the assessment afterwards. This perhaps explains why our "phone-assessment before office" group contains a significantly older population (mean age 63) than the "phone-assessment after office" group (mean age 56). Limitations in the study include a broad study population consisting of both surgical and nonsurgical patients. We however believe that this was controlled through the randomized crossover design of the study along with paired *t* test and ICC analysis that significantly reduced the effect of individual variability. The nonpostoperative patients also make a small proportion of our overall population. In context of our tight statistics proving the reliability of administering these tests over phone or email for a broad swath of patients, inclusion of this small population would not affect the overall result since the statistics are being driven by the large postoperative group. Furthermore, strong test-retest reliability for postoperative patients, along with our overall statistics for all groups, underscore that a confounding bias is not present. The choice of a 1- to 4-week interval also represented several tradeoffs. The 1-week interval was selected as an optimal equilibrium between the tradeoffs of recall bias and minimization of assessment score drift due to

chronic disease processes. Indeed, as has been noted in literature, a 2- to 14-day test-retest interval is effective for mitigating recall bias [13]. The 4-week interval was another optimal equilibrium between the tradeoffs of having sufficient time to follow up with patients who had not yet completed their external (email or phone) assessment and minimizing score drift. Our mean days between assessments for email and both phone groups, which very closely approximated 14 days for all groups, fell into a "sweet spot" in the middle of this interval. While it might be claimed that a Hawthorne effect might exist in the form of higher PROs and higher test-retest reliability amongst our study patients who were aware of testing, the impact of participation in the study on PROs is nonunique. If a Hawthorne effect existed, it would still apply whether the study was being conducted or not. Since our in-office NDI and JOA forms being administered as standard of care already contain written text asking patients to "please answer all questions completely" since "it is in [their] best interest and will assist [their] doctor with [their] care" and thus lead patients to presume that their surgeon will read their answers at the time of care, patients might be incentivized to report higher PROs than in actuality to receive greater attention. In other words, the link between participation in the study and the impact on PROs is neither unique nor significant. PRO inflation is caused by factors outside of our study. Any effect on recall would have also been mitigated by the time period between assessments.

In summary, we demonstrate a methodology for reducing loss to follow-up and improving study validity by demonstrating no statistical significance and high homogeneity in both the email and phone versions of the NDI and JOA when compared to in-office versions. We also find email to be a more convenient and reliable form of assessment than phone. Administration of the NDI and JOA over phone or email also offers additional opportunities. Studies can be completed sooner by reducing the number of patients recruited to account for follow-up loss. Additionally, email and phone assessments can also enable new opportunities for assessing patient populations less likely to follow-up, such as trauma patients, that have been understudied due to dependence on in-office assessments.

References

[1] Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res 2008;17:179–93. https://doi.org/10.1007/s11136-007-9295-0.

[2] MacDermid JC, Walton DM, Avery S, Blanchard A, Etruw E, McAlpine C, et al. Measurement properties of the neck disability index: a systematic review. J Orthop Sport Phys Ther 2009;39:400–C12. https://doi.org/10.2519/jospt.2009.2930.

[3] Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. J Manip Physiol Ther 2008;31:491–502. https://doi.org/10.1016/j.jmpt.2008.08.006.

[4] Tetreault L, Kopjar B, Nouri A, Arnold P, Barbagallo G, Bartels R, et al. The modified Japanese Orthopaedic Association scale: establishing criteria for mild, moderate and severe impairment in patients with degenerative cervical myelopathy. Eur Spine J 2017;26:78–84. https://doi.org/10.1007/s00586-016-4660-8.

[5] Dettori J. Loss to follow-up. Evid Based Spine Care J 2011;2:7–10. https://doi.org/10.1055/s-0030-1267080.

[6] Von Allmen RS, Weiss S, Tevaearai HT, Kuemmerli C, Tinner C, Carrel TP, et al. Completeness of follow-up determines validity of study findings: results of a prospective repeated measures cohort study. PLoS One 2015;10:1–13. https://doi.org/10.1371/journal.pone.0140817.

[7] Palmen LN, Schrier JCM, Scholten R, Jansen JHW, Koe S. Is it too early to move to full electronic PROM data collection ? A randomized controlled trial comparing PROM's after hallux valgus captured by e-mail, traditional mail and telephone. Foot Ankle Surg 2016;22:46–9. https://doi.org/10.1016/j.fas.2015.05.001.

[8] Schwartzenberger J, Presson A, Lyle A, Farrell AO, Tyser AR. Remote collection of patient-reported outcomes following outpatient hand surgery: a randomized trial of telephone, mail, and e-mail. J Hand Surg Am 2017;42:693–9. https://doi.org/10.1016/j.jhsa.2017.05.002.

[9] Loban A, Mandefield L, Hind D, Bradburn M. A randomized trial found online questionnaires supplemented by postal reminders generated a cost-effective and generalizable sample but don't forget the reminders. J Clin Epidemiol 2017;92:116–25. https://doi.org/10.1016/j.jclinepi.2017.08.003.

[10] Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. Value Health 2008;11:322–33. https://doi.org/10.1111/j.1524-4733.2007.00231.x.

[11] Bokshan SL, Godzik J, Dalton J, Jaffe J, Lenke LG, Kelly MP. Reliability of the revised Scoliosis Research Society-22 and Oswestry Disability Index (ODI) questionnaires in adult spinal deformity when administered by telephone. Spine J 2016;16:1042–6. https://doi.org/10.1016/j.spinee.2016.03.022.

[12] Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81. https://doi.org/10.1016/j.jbi.2008.08.010.

[13] Brouwer CNM, Schilder ÆAGM, Veenhoven RH, Grobbee ÆDE. Reliability and validity of functional health status and health-related quality of life questionnaires in children with recurrent acute otitis media 2007:1357–73. doi:10.1007/s11136-007-9242-0.

[14] Carreon LY, Glassman SD, Campbell MJ, Anderson PA. Neck Disability Index, short form-36 physical component summary, and pain scales for neck and arm pain : the minimum clinically important difference and substantial clinical benefit after cervical spine fusion. Spine J 2010;10:469–74. https://doi.org/10.1016/j.spinee.2010.02.007.

[15] R Core Team. R: A Language and Environment for Statistical Computing 2018.

[16] Gamer M, Lemon J, <puspendra.pusp22@gmail.com>IFPS. irr: Various Coefficients of Interrater Reliability and Agreement 2012.

[17] Solberg TK, Sørlie A, Sjaavik K, Nygaard ØP. Would loss to follow-up bias the outcome evaluation of patients operated for degenerative disorders of the lumbar spine ? A study of responding and non-responding cohort participants from a clinical spine surgery registry. Acta Orthop 2011: 3674. https://doi.org/10.3109/17453674.2010.548024.

[18] Imam MA, Barke S, Stafford GH, Parkin D, Field RE. Loss to follow-up after total hip replacement : a source of bias in patient reported outcome measures and registry datasets? 2014;24:465–72. doi:10.5301/hipint.5000141.

[19] Hutchings A, Neuburger J, Grosse Frie K, Black N, van der Meulen J. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. Health Qual Life Outcomes 2012;10:34. https://doi.org/10.1186/1477-7525-10-34.

[20] Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO good research practices task force report. Value Health 2009;12:419–29. https://doi.org/10.1111/j.1524-4733.2008.00470.x.

[21] Kato S, Oshima Y, Oka H, Chikuda H, Takeshita Y. Comparison of the Japanese Orthopaedic Association (JOA) Score and Modified JOA (mJOA) score for the assessment of cervical myelopathy: a multicenter observational study. PLoS One 2015;10(4):e0123022. https://doi.org/10.1371/journal.pone.0123022.

[22] Harewood GC, Yacavone RF, Locke GR, Wiersema MJ. Prospective comparison of endoscopy patient satisfaction surveys: e-mail versus standard mail versus telephone. Am J Gastroenterol 2001;96(12):3312–7.

[23] Nota SPFT, Strooker JA, Ring D. Differences in response rates between mail, e-mail, and telephone follow-up in hand surgery research. Hand (N Y) 2014;9(4):504–10. https://doi.org/10.1007/s11552-014-9618-x.

[24] Rolfson O, Salomonsson R, Dahlberg LE, Garellick G. Internet-based follow-up questionnaire for measuring patient-reported outcome after total hip replacement surgery-reliability and response rate. Value Health 2011;14:316–21. https://doi.org/10.1016/j.jval.2010.08.004.