



■ RESEARCH

Are the patient-rated wrist evaluation (PRWE) and the disabilities of the arm, shoulder and hand (DASH) questionnaire used in distal radial fractures truly valid and reliable?

Y. V. Kleinlugtenbelt,
R. G. Krol,
M. Bhandari,
J. C. Goslings,
R. W. Poolman,
V. A. B. Scholtes

Joint Research OLVG
East, Amsterdam, The
Netherlands

■ Y. V. Kleinlugtenbelt, MD, PhD, Orthopaedic Surgeon, Department of Orthopaedic and Trauma Surgery, Joint Research OLVG Oost, PO Box 95500, 1090 HM, Amsterdam, the Netherlands; Department of Orthopaedic and Trauma Surgery, Deventer Ziekenhuis, Nico Bolkesteinaan 75, 7416 SE Deventer, the Netherlands; and Division of Orthopaedic Surgery, Department of Surgery, McMaster University, 1200 Main St, Hamilton, Ontario L8N 3Z5, Canada.

■ R. G. Krol, MD, Resident Radiologist,

■ R. W. Poolman, MD, PhD, Orthopaedic Surgeon,

■ V. A. B. Scholtes, PhD, Program Manager, Department of Orthopaedic and Trauma Surgery, Joint Research OLVG Oost, PO Box 95500, 1090 HM, Amsterdam, The Netherlands.

■ M. Bhandari, MD, PhD, Orthopaedic Surgeon and Professor of Surgery, Division of Orthopaedic Surgery, McMaster University, 1200 Main St. West, Room 4E17, Hamilton, Ontario L8N 3Z5, Canada.

■ J. C. Goslings, MD, PhD, Trauma Surgeon and Professor of Surgery, Department of Surgery, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.

Correspondence should be sent to Y. V. Kleinlugtenbelt; email: ijdokib@hotmail.com

doi: 10.1302/2046-3758.71.BJR-2017-0081.R1

Bone Joint Res 2018;7:36–45.

Objectives

The patient-rated wrist evaluation (PRWE) and the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire are patient-reported outcome measures (PROMs) used for clinical and research purposes. Methodological high-quality clinimetric studies that determine the measurement properties of these PROMs when used in patients with a distal radial fracture are lacking. This study aimed to validate the PRWE and DASH in Dutch patients with a displaced distal radial fracture (DRF).

Methods

The intraclass correlation coefficient (ICC) was used for test-retest reliability, between PROMs completed twice with a two-week interval at six to eight months after DRF. Internal consistency was determined using Cronbach's α for the dimensions found in the factor analysis. The measurement error was expressed by the smallest detectable change (SDC). A semi-structured interview was conducted between eight and 12 weeks after DRF to assess the content validity.

Results

A total of 119 patients (mean age 58 years (SD 15)), 74% female, completed PROMs at a mean time of six months (SD 1) post-fracture. One overall meaningful dimension was found for the PRWE and the DASH. Internal consistency was excellent for both PROMs (Cronbach's α 0.96 (PRWE) and 0.97 (DASH)). Test-retest reliability was good for the PRWE (ICC 0.87) and excellent for the DASH (ICC 0.91). The SDC was 20 for the PRWE and 14 for the DASH. No floor or ceiling effects were found. The content validity was good for both questionnaires.

Conclusion

The PRWE and DASH are valid and reliable PROMs in assessing function and disability in Dutch patients with a displaced DRF. However, due to the high SDC, the PRWE and DASH are less useful for individual patients with a distal radial fracture in clinical practice.

Cite this article: *Bone Joint Res* 2018;7:36–45.

Keywords: Patient-rated wrist evaluation (PRWE), Disabilities of the Arm Shoulder and Hand (DASH) questionnaire, Distal radial fracture, Validity, Reliability

Article focus

- To determine which PROMs have the best measurement properties for evaluation of functional outcome in patients with distal radius fractures.
- Validation of the PRWE and DASH in patients with distal radius fractures according to the COSMIN standard.
- To determine whether the PRWE and DASH are useful for individual patients

with distal radius fractures in clinical practice.

Key messages

- The PRWE and the DASH are valid and reliable PROMs in assessing function and disability in patients with displaced distal radius fractures.
- The PRWE and DASH are less useful for individual patients with distal radius fractures in clinical practice.

Strengths and limitations

- Strength: We used the COSMIN standards for adequate study design and statistical analysis.
- Limitation: The Dutch language version of the PRWE and DASH were used.

Introduction

In order to conduct high-quality clinical studies in the treatment of patients with a distal radial fracture, and to exchange results globally in a standardized way, there must be consensus on the use of outcome measures. Instruments such as patient-reported outcome measures (PROMs) are gaining importance in clinical trials of fracture treatment.¹ The methodological quality of these instruments is important; they should be valid and reliable. Ideally, this should be determined prior to use, as the quality of such instruments directly affects the quality of the information obtained with these instruments.² If not, one risks imprecise or biased results, potentially leading to incorrect conclusions.³ To assess the methodological quality of a PROM, standards are needed. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) group set these standards for adequate study design and statistical analysis.² The group also developed a checklist in an international Delphi study in which consensus was reached on terminology, definitions, and a taxonomy of measurement properties of PROMs.³

Recently, we performed a systematic review in which we used this COSMIN checklist to determine the methodological quality of studies that evaluated measurement properties of various PROMs for patients with a distal radial fracture.⁴ The patient-rated wrist evaluation (PRWE) and the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire were most extensively evaluated in terms of measurement properties. However, strong evidence supporting any available PROMs in patients with a distal radial fracture is lacking. We found that, overall, the measurement properties are good, but the methodological quality of the studies assessing these properties is low.⁵⁻¹⁶ Based on this review, we currently risk imprecise or biased results when using these PROMs in, for instance, clinical studies, and may base our knowledge on incorrect conclusions. This review has shown that studies of higher methodological quality are needed to determine the quality of measurement properties. If the methodological quality of clinimetric studies continues to improve, PROMs can be selected more carefully.⁴

In the current study, we aim to examine in greater depth which PROMs have the best measurement properties for evaluation of functional outcome in patients with a distal radial fracture. We will therefore attempt to establish the content validity, test-retest reliability, internal consistency, measurement error, and floor and ceiling effects of the Dutch PRWE and DASH in patients with a

distal radial fracture. The measurement properties will be assessed according to the recently formulated COSMIN standards.²

Materials and Methods

Study design. A multicentre prospective cross-sectional clinimetric study was performed between July 2012 and April 2013 at the orthopaedic and surgery departments of three participating hospitals. Ethics approval was obtained from the local medical ethical committee at all three Dutch hospitals (WO 12.064). Written informed consent was obtained from all patients.

Study patients. Patients were eligible for inclusion if they were 18 years of age or older, and were presenting in the emergency department with a displaced distal radial fracture that required reduction. Both conservatively and surgically treated patients were included. Patients were excluded if they 1) had a prior fracture or pathology of the ipsilateral distal radius, 2) had multiple fractures, 3) had cognitive impairment, or 4) were unable to understand the Dutch language.

We aimed to include at least 20 patients to assess the content validity, which is double the number that is required.¹⁷ This group was consulted at between eight and 12 weeks post-fracture.

In addition, we aimed to include at least 100 patients to assess the reliability domain, which is required in order to achieve a study status of high quality according to the COSMIN guidelines.¹⁷ These patients were assessed at between four and eight months post-fracture, and did not participate in the content validation. The type of fracture was scored on radiographs according to AO classification.¹⁸

The PRWE is a self-administered, patient-specific questionnaire that consists of 15 items. It was designed to measure wrist pain and disability in activities of daily living, and consists of two subscales: pain and function. The pain subscale comprises five items, while the subscale function is divided into six specific activities and four usual activities. Both subscales are summed and scored on a ten-point ordinal scale.^{7,19} The 'pain' subscale score is the sum of the five items. The 'function' subscale score is calculated by the sum of the ten items divided by two. The total score of the PRWE is the sum of the scores of both subscales. A score of 100 represents the worst functional score, whereas 0 represents no disability.

In 2004, the PRWE was modified to become the PRWHE (patient-rated wrist/hand evaluation). The PRWHE consists of the same items and scoring system as the PRWE, with minor changes.²⁰ In the PRWHE, the term 'wrist' was replaced by 'wrist/hand'. Also, two aesthetic items, that are not part of the scoring system, were added. Therefore, measurement properties of these two items were not assessed in this study. The PRWHE-Dutch Language Version (PRWHE-DLV) was used in this study.²¹

The DASH is a self-administered questionnaire, developed to evaluate symptoms and physical function of the whole upper limb. It is scored in two components: the main disability/symptom section and two optional sections. The main component of the DASH is a 30-item scale concerning the patient's health status during the preceding week: 21 items regarding the degree of difficulty in performing certain physical activities; five items pertaining to the severity of pain, activity-related pain, tingling, weakness, and stiffness; and four items concerning the effect that the upper limb problem has had on social activities, work, sleep, and self-image. Each item is scored on a five-point ordinal scale. To calculate the main DASH score, all completed responses are summed and the mean is calculated. The total score is calculated by subtracting 1 from the mean and multiplying by 25 $((n-1)*25)$. This gives a total score ranging from best to worst on a scale of 0 to 100.²² At least 27 of the 30 items must be completed to calculate a score.

Both optional sections - high performance sport/music and work - consist of four items, scored on a five-point ordinal scale and calculated similarly. However, all five items must be answered, as the percentage of missing items must not exceed 10%. The DASH-Dutch Language Version (DASH-DLV) was used in this study.²³

Assessment of measurement properties: validity. Content validity examines the degree to which the content of a health-related-patient-reported outcome measurement (HR-PROM) is an adequate reflection of the construct to be measured. For all measurement instruments, it is important that the content validity is assessed by experts. For PROMs, patients - particularly representatives of the target population - are the experts. They are the most appropriate assessors of the relevance of the items of the questionnaire.²⁴

Content validity was assessed by phone interview. Eight weeks after incurring a distal radial fracture, patients were asked by phone to participate in the study. When patients agreed to participate, a semi-structured interview was conducted, in which the patient was asked about the functional problems that were the most limiting to them. Each interview started with the standard question: "What are your main limitations in both physical and social functioning after breaking your wrist?" If necessary, examples were given (e.g. carrying heavy objects, personal hygiene).

After the interview, a list was compiled of the functional problems named during the interview. Using the results of all of the interviews, a list of the 20 most common functional problems was created and compared with the items used in the PRWE and the DASH questionnaire.

Content validity is based on judgement and no statistical testing is involved.²⁴ We considered the content validity to be good if 75% of the items of the PRWE and DASH matched the problems mentioned in the interviews.

Assessment of measurement properties: reliability. All patients who were eligible for inclusion received an information letter, between four and eight months after sustaining a distal radial fracture, in which they were asked to participate in this study. Participants completed a web-based questionnaire containing the PRWE and the DASH at home. If participants did not have access to the internet, they could alternatively receive a paper version. Two weeks after completing the questionnaire for the first time, patients received an email or letter in which they were asked to complete the questionnaire for the second time at home. In this two-week interval, no major changes in the health status were expected and immediate recall would be unlikely.

The digital and paper versions of the DASH and PRWE were identical. All items on the web-based questionnaire required a response, thus no unanswered items were expected. If a patient who received a paper version had not answered all of the items, he/she was contacted by phone and asked to provide a response to the omitted items. If the patient did not want to answer any of these items, they were excluded from the study.

The internal consistency is the degree of inter-relatedness among items.³ If items in a scale are summarized in a total score, the items should be sufficiently correlated. This correlation is established by the internal consistency and indicates whether the items seem to measure the same construct.²⁵ If one item measures something else, it will have a lower item-total correlation than the other items. The internal consistency was assessed by using the first measurement of the test-retest reliability.

First, exploratory factor analysis was performed to determine whether the PROM consisted of only one overall dimension, or of more than one dimension. Factor analysis was assessed by calculating eigenvalues. An eigenvalue of one or higher indicated a dimension. The eigenvalues were presented in a scree plot. The relative contribution of different dimensions was judged based on the 'elbow' in the scree plot and the percentage of variance.²⁴ If the PROM consisted of two or more dimensions, factor loading was assessed. Factor loading represents the correlation between the items in the PROM and the underlying dimensions. We considered factor loadings of at least 0.50 to be meaningful.²⁶

Internal consistency was determined by calculating the Cronbach's α for the dimensions found in the factor analysis. If the Cronbach's α has a value of > 0.70 , items are considered sufficiently correlated.²⁷ However, values > 0.95 can indicate that the instrument contains too many items that assess the same underlying construct.²⁸ The QuickDASH is a shortened version of the DASH outcome measure. Instead of 30 items, the QuickDASH uses only 11 items. The Cronbach's α for the 11 items of the QuickDASH was also determined.

The test-retest reliability is the proportion of the total variance in the measurements, due to true differences

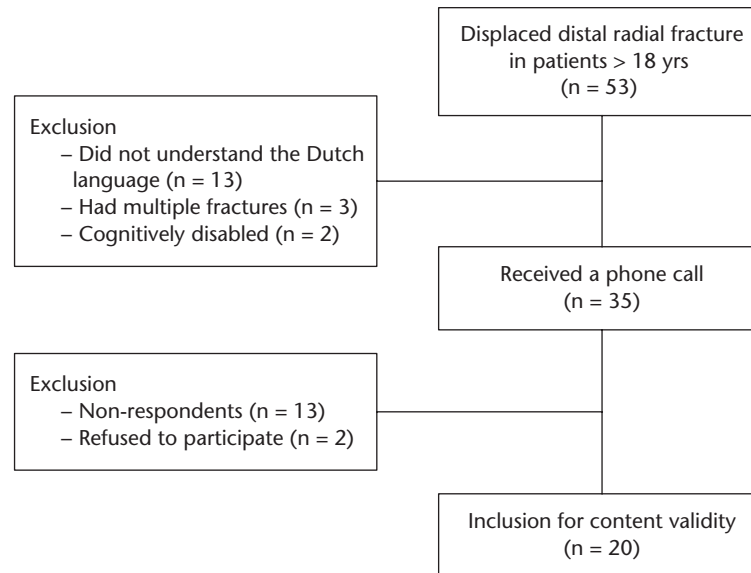


Fig. 1

Flow chart showing patient recruitment and exclusion for the assessment of content validity.

between patients over time.³ This refers to the degree to which the measurement instrument is free from measurement error. It also estimates the extent to which scores for patients who have not changed are the same for repeated measurements at different timepoints.^{3,29} High reliability is important for discriminative purposes if one wants to distinguish outcomes among patients. The test-retest reliability was assessed by calculating the intraclass correlation coefficient (ICC) with a 95% confidence interval (CI). The ICC is a relative parameter and will always have a value between 0 and 1. Higher values represent higher reliability. An ICC of > 0.70 is considered acceptable, > 0.80 is considered good, and > 0.90 is considered excellent.²⁹ Only patients who completed the optional modules of the DASH twice were included for the test-retest assessment of these modules.

The measurement error is the systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured.³ When the measurement error is low or zero, the difference measured is due to true differences. To express the measurement error, the standard error of the measurement (SEM) and the smallest detectable change (SDC) can be used. The SEM represents the standard deviation of repeated measures of one individual. The SDC represents the minimal change that must occur in order for the scale to affirm that it is a real change, rather than a measurement error. The SEM was calculated from the square root of the variance between the measurements and the error variance of the ICC. For a conventional confidence level of 95%, the SDC was calculated as $1.96 \times \sqrt{2} \times SEM$.²⁴

The presence of floor or ceiling effects may have a negative effect on the quality of the instrument. If patients score primarily in the extremes, the responsiveness may

be limited. Floor or ceiling effects were considered to be present if more than 15% of the respondents achieved the minimum or maximum possible score.³⁰ When taking the SDC into account, we should consider floor and ceiling effects more broadly. If a score is closer to the SDC than the maximum or minimum score, a change beyond the measurement error cannot be measured. We also assessed the percentage of patients within the SDC range from both extremes. Floor and ceiling effects were assessed by using the first measurement of the test-retest reliability.

Statistical analysis. Statistical analysis was conducted using the SPSS software, version 18 (IBM Corp., Armonk, New York).

Results

Validity. In total, 35 patients met the inclusion criteria and received a phone call (Fig. 1). Two patients refused participation in the study and 13 patients did not respond. Consequently, 20 patients were included in the study to assess the content validity and were interviewed.

The mean age was 59.30 years (SD 13.61). Most included patients were women (80%). Half of the patients had an AO subtype C1 fracture and most of the patients (65%) received conservative treatment (Table I). The mean time between the fracture and the semi-structured interview was 9.85 weeks (SD 1.98).

A total of 74 problems were mentioned in the semi-structured interviews. All 15 items (100%) of the PRWHE were named in the interviews. Therefore, we considered the content validity of the PRWHE to be good. Only one DASH item, 'gardening', was not mentioned in the interviews. Ninety-six per cent of the DASH questions were mentioned in the interviews, which we considered to be

Table I. Characteristics of included patients for internal consistency, test-retest reliability and content validity

	Internal consistency	Test-retest reliability	Work module		Sport module		Content validity
Patients, n	119	109	Int cons	Test-retest	Int cons	Test-retest	20
Female, n (%)	88 (74.0)	82 (75.0)	60 (71.4)	55 (71.4)	52 (74.3)	46 (78.0)	16 (80.0)
Mean age, yrs (SD)	58.40 (15.32)	58.76 (15.12)	53.74 (13.76)	54.35 (13.65)	55.37 (15.24)	55.83 (14.89)	59.30 (13.61)
AO subtype, n (%)							
A2	25 (21.0)	23 (21.1)	18 (21.4)	17 (22.1)	14 (20.0)	12 (20.3)	4 (20.0)
A3	15 (12.6)	13 (11.9)	10 (11.9)	8 (10.4)	10 (14.3)	7 (11.9)	2 (10.0)
B1	1 (0.8)	1 (0.9)	1 (1.2)	1 (1.3)	1 (1.4)	1 (1.7)	0 (0.0)
B2	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (5.0)
B3	6 (5.0)	5 (4.6)	5 (6.0)	4 (5.2)	3 (4.3)	2 (3.4)	2 (10.0)
C1	35 (29.4)	34 (31.2)	23 (27.4)	23 (29.9)	17 (24.3)	16 (27.1)	10 (50.0)
C2	30 (25.2)	27 (24.8)	23 (27.4)	21 (27.3)	22 (31.4)	19 (32.2)	1 (5.0)
C3	7 (5.9)	6 (5.5)	4 (4.8)	3 (3.9)	3 (4.3)	2 (3.4)	0 (0.0)
Treatment, n (%)							
Conservative	73 (61.3)	67 (61.5)	51 (60.7)	48 (62.3)	43 (61.4)	36 (61.0)	13 (65.0)
K-wire fixation	1 (0.8)	1 (0.9)	1 (1.2)	1 (1.3)	1 (1.4)	1 (1.7)	0 (0.0)
ORIF (volar plate)	39 (32.8)	36 (33.0)	30 (35.7)	27 (35.1)	23 (32.9)	20 (33.9)	6 (30.0)
ORIF (dorsal plate)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
External fixation	6 (5.0)	5 (4.6)	2 (2.4)	1 (1.3)	3 (4.3)	2 (3.4)	1 (5.0)

K-wire, Kirschner wire; ORIF, open reduction internal fixation; Int cons, internal consistency

good content validity. Both work and sport were mentioned in the semi-structured interview. Therefore, we consider the content validity of both optional modules of the DASH to be good.

Reliability. Of the 297 patients who entered the emergency room at one of the three participating hospitals during the study period with a displaced distal radial fracture requiring reduction, a total of 119 met the inclusion criteria and completed the questionnaire the first time after a mean of 6.15 months (SD 1.00). Most included patients were women (75%). Ten patients did not respond to the second measurement. Therefore, 119 patients were included for the assessment of internal consistency, and 109 patients were included to assess test-retest reliability and measurement error (Fig. 2). The mean time between the first and second assessments was 18.66 days (SD 7.27). No major change in health status took place in any of the patients between the measurement points, so no patients were excluded for that reason. Since not all patients were employed and/or played a sport or instrument, the optional work and sport modules of the DASH were assessed with fewer patients. The exact number of patients is described in Table I, as are their characteristics for each assessment.

There were no missing items in the online completed questionnaires. In total, 23 patients completed the questionnaires on paper. There were no missing items in the paper PRWE. One item was left unanswered in the DASH in two instances. Even after a phone consult, the answer could not be obtained as these patients did not wish to answer the sexuality item of the DASH. Both patients were excluded from further assessment.

The factor analysis was performed for both the PRWE and the DASH. Only one dimension was extracted from the PRWE, which explained 66.26% of the total variance.

The 'elbow' in the scree plot was seen at the second component (Fig. 3). The Cronbach's α of the total PRWE was 0.96 (Table II), indicating excellent internal consistency and redundancy. Removing items from the questionnaire did not result in a higher Cronbach's α . The Cronbach's α for the 11 items of the QuickDASH when using our data would be 0.91, also indicating an excellent internal consistency and no redundancy.

Five dimensions were extracted from the DASH. Item loading (Table III) showed that the first dimension consisted of items typically asking about strength. The second dimension is more specific on function. The third dimension consisted of items pertaining to pain and disabilities. The fourth and fifth dimensions consisted of two or more specific items.

The first dimension explained 55.71% of the total variance, and the scree plot showed an 'elbow' at component two, indicating only one overall meaningful dimension (Fig. 4). Therefore, despite extracting five dimensions, we assessed the Cronbach's α only for the total DASH and the optional modules.

The Cronbach's α of the total DASH, the optional work section, and the sport/music section were 0.97, 0.94, and 0.96, respectively, indicating excellent internal consistency. Removing items from the questionnaire did not result in a higher Cronbach's α .

Table II shows the distribution of the data at the first and second measurements. The ICC of the PRWE was 0.87, which indicates good reliability. The SEM was 7.40, with a SDC of 20.51. The ICC of the total DASH, the optional work module, and the sport module were 0.91, 0.87, and 0.87, respectively, indicating excellent reliability for the total DASH and good reliability for the optional modules. The SEMs were 5.10, 5.08, and 11.18, respectively; the SDCs were 14.12, 14.08, and 30.99, respectively.

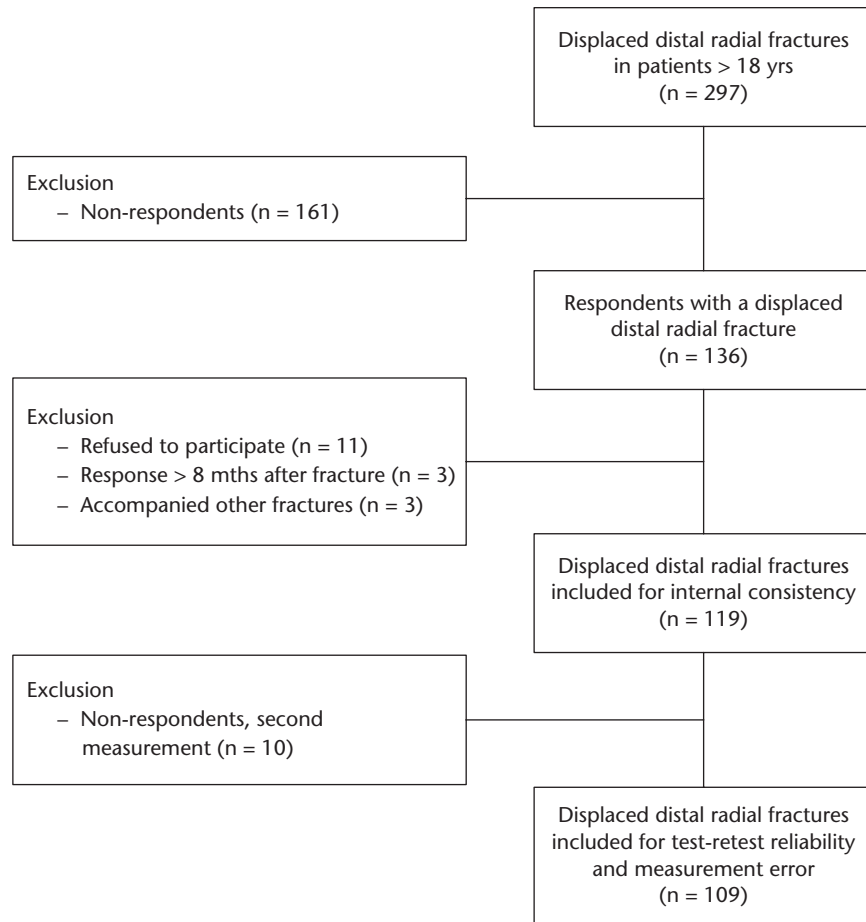


Fig. 2

Flowchart showing patient recruitment and exclusion for the assessment of internal consistency, test-retest reliability and measurement error.

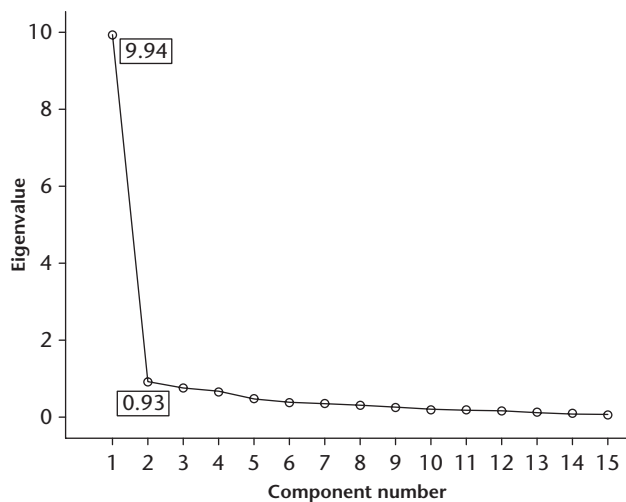


Fig. 3

Screeplot PRWE. Only one dimension was extracted of the PRWE, which explained 66.26% of the total variance. The “elbow” in the scree plot was seen at the second component.

No floor or ceiling effects were found in the PRWE or total DASH questionnaires. Floor or ceiling effects were found in both optional modules of the DASH

questionnaire (Table IV). When taking the SDC into consideration, more than 15% of the scores of both the PRWE and the total DASH were within the range of the SDC from the lowest possible score (45% and 46%), showing a clear floor effect. Only the score for the optional sport module of the DASH questionnaire was within the range of the SDC from the highest possible score, showing a ceiling effect.

Discussion

Although the PRWE and the DASH are the most thoroughly studied PROMs, the quality of these two PROMS was not supported with strong levels of evidence on any of the measurement properties to evaluate patients with a distal radial fracture.⁴ This study, in which the COSMIN standards were followed to ensure high methodological quality, provides strong evidence that both the PRWE and the DASH questionnaires have good content validity, and are reliable and internally consistent instruments for the assessment of patients with a distal radial fracture.

Using semi-structured interviews with patients with a distal radial fracture, we found good content validity for both the PRWE and the DASH. In a previous study, only

Table II. Results for internal consistency, reliability, and measurement error

PROM	Cr- α	Mean T1 (sd)		Mean T2 (sd)	ICC (95% CI)	SEM	SDC
		n = 119	n = 109	n = 109			
PRWE	0.96	26.92 (21.16)	25.97 (20.37)	24.95 (20.73)	0.87 (0.82 to 0.91)	7.38	20.47
DASH	0.97	19.55 (17.70)	18.83 (16.59)	19.36 (17.93)	0.91 (0.87 to 0.94)	5.10	14.12
DASH, Work	0.94	15.40 (20.61)	14.98 (18.74)	14.06 (19.93)	0.87 (0.80 to 0.92)	5.08	14.08
DASH, Sport	0.96	33.66 (32.26)	33.39 (32.57)	28.92 (30.01)	0.87 (0.79 to 0.92)	11.18	30.99

PROM, patient-reported outcome measure; Cr- α , Cronbach's α ; Mean T1, mean score at timepoint 1; Mean T2, mean score at timepoint 2; ICC, intraclass correlation; SEM, standard error of measurement; SDC, smallest detectable change; PRWE, Patient-rated wrist evaluation; DASH, the Disabilities of the Arm, Shoulder and Hand questionnaire

Table III. Factor analysis for the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire

Question	Component				
	1	2	3	4	5
1	0.76*	0.24	0.26	0.14	0.15
2	0.14	0.81*	0.15	0.06	0.11
3	0.33	0.48	0.39	0.35	0.26
4	0.35	0.46	0.43	0.31	0.34
5	0.62*	0.34	0.34	0.28	0.05
6	0.53*	0.51*	0.17	0.21	0.08
7	0.61*	0.35	0.40	0.31	0.11
8	0.62*	0.36	0.24	0.35	0.29
9	0.45	0.58*	0.22	0.31	0.20
10	0.73*	0.10	0.21	0.34	0.24
11	0.86*	0.07	0.15	0.20	0.23
12	0.41	0.53*	0.18	0.48	0.13
13	0.20	0.64*	0.09	0.52*	0.23
14	0.14	0.57*	0.25	0.42	0.17
15	0.28	0.42	0.21	0.58*	0.39
16	0.27	0.68*	0.24	0.37	0.20
17	0.20	0.45	0.31	0.59*	0.25
18	0.57*	0.49	0.32	0.07	0.19
19	0.61*	0.40	0.32	0.20	0.19
20	0.29	0.15	0.19	0.75*	0.12
21	0.27	0.27	0.27	0.65*	0.24
22	0.20	0.12	0.77*	0.22	-0.14
23	0.28	0.34	0.68*	0.16	0.07
24	0.27	0.42	0.57*	0.05	0.53*
25	0.51*	0.27	0.53*	-0.03	0.38
26	0.17	0.08	0.04	0.29	0.70*
27	0.53*	0.14	0.44	0.23	0.40
28	0.19	0.19	0.59*	0.24	0.28
29	0.30	0.31	0.12	0.23	0.68*
30	0.32	0.03	0.58*	0.28	0.22

*Factor loadings > 0.50 are appropriate

the developers of the PRWE assessed the content validity of the PRWE for patients with a distal radial fracture.⁷ They assessed the content validity by interviewing experts in the relevant field of medicine rather than interviewing patients with a distal radial fracture. However, the most appropriate assessors of the relevance of items on a questionnaire are the representatives of the target population.²⁴ To our knowledge, this is the first study in which the assessment of the content validity of the DASH in patients with a distal radial fracture was performed. A frequent remark was that the affected distal radial fracture was not the dominant side. Another remark was about cutting food. Most people cut food with their right hand. Five female patients stated that they experienced difficulty with putting on their bra. Overall, functional

problems experienced by patients with a distal radial fracture consisted of activities for which both hands are needed.

To our knowledge, this study is the first in which a factor analysis of the PRWE and the DASH was performed specifically in patients with a distal radial fracture. The PRWE was developed as a one-dimensional questionnaire. However, that dimension consists of two subscales (pain and function). Exploratory factor analysis extracted only one dimension of the PRWE. Therefore, Cronbach's α was assessed only for the total 15 items on the PRWE. The DASH was also developed as a one-dimensional questionnaire, with two optional modules. However, we derived five dimensions. Despite extracting five dimensions, we assessed Cronbach's α only for the total DASH, as the total score is calculated by using all 30 questions. Component one explained 57.71% of the total variance and the scree plot showed an 'elbow' at component two, indicating only one overall meaningful dimension.

For both PRWE and DASH, a high Cronbach's α was found, 0.96 and 0.97 respectively, comparable with previous validation studies.^{5,6,9-15} However, a Cronbach's $\alpha > 0.95$ could indicate item redundancy. This suggests that some items can be removed when using one of these measurement instruments in patients with a distal radial fracture.

The QuickDASH is a shortened version of the DASH outcome measure. Instead of 30 items, the QuickDASH uses 11 items. Cronbach's α for the QuickDASH when using our data would be 0.91, also indicating an excellent internal consistency and no redundancy. Based on our results, one might recommend using the QuickDASH instead of the DASH, but further investigation is needed in order to confirm this conclusion.

We determined a good (ICC = 0.87) and excellent (ICC = 0.91) reliability for the PRWE and DASH questionnaires, respectively, which are comparable with other studies. This study is the third to report on measurement error of the PRWE, and the first for the DASH in patients with a distal radial fracture. In our study, the SDC of the DASH (14.1) was clearly lower than the SDC of the PRWE (20.5), which could imply that the DASH is more useful, especially in clinical practice.

Kim and Kang¹² reported a SDC of 4.4 for the PRWE in 63 patients with a distal radial fracture. We found a much

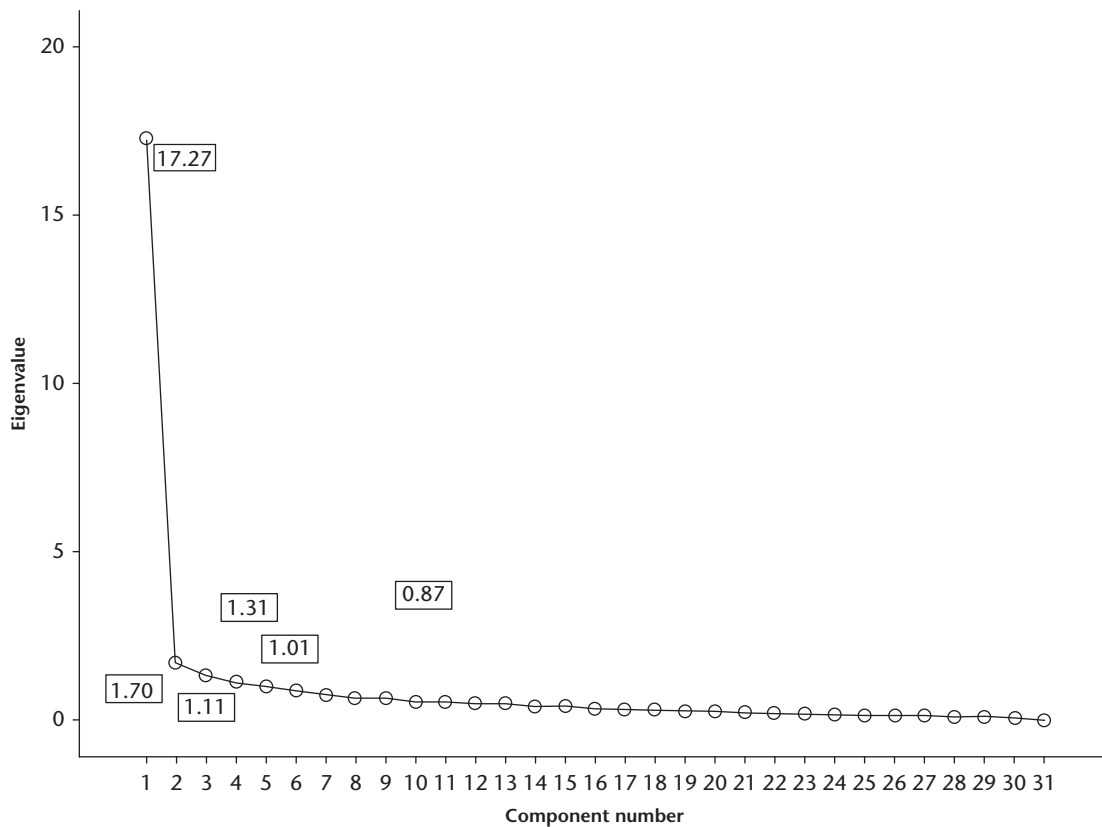


Fig. 4

Screeplot DASH. Five dimensions were extracted of the DASH. Dimension one explained 55.71% of the total variance and the “elbow” in the scree plot was seen at the second component, insinuating only one overall meaningful dimension.

Table IV. Absolute floor and ceiling scores, the smallest detectable change (SDC) ranges, and the scores that fell within the SDC range for both extremes of both patient-rated wrist/hand evaluation (PRWHE) and the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire scores

PROM	Absolute (%)		SDC range		Patients within SDC range (%)	
	Floor	Ceiling	Floor	Ceiling	Floor	Ceiling
PRWHE	5.9	0.8	0.00 to 0.47	79.53 to 100.00	45.4	1.7
DASH	4.2	0	0.00 to 14.12	85.88 to 100.00	46.6	0.8
DASH, Work	44	1.2	0.00 to 14.08	85.92 to 100.00	58.3	2.4
DASH, Sport	20	8.6	0.00 to 30.99	69.01 to 100.00	65.7	18.6

PROM, patient-reported outcome measure; SDC, smallest detectable change

higher SDC value of 20.5, indicating that a patient must improve their total score by at least 20% to ensure an improvement beyond measurement error. This high difference can partly be explained by the fact that the previous authors used a CI of 90% instead of 95%. In their study, outcomes were more homogeneous than in our study, which normally led to a lower ICC. However, surprisingly, the authors calculated a higher ICC than we did in our study. Based on their data, we could not find an explanation for this difference. As a result, their SDC was correspondingly very low.

Walenkamp et al¹⁶ found a SDC of 11. This difference can also partly be explained by the fact that the authors used a CI of 90% instead of 95%. However, the main

reason for this difference is probably that the authors used Cronbach's α instead of test-retest parameters (e.g. ICC) to calculate the SDC. Cronbach's α is assessed at a single point in time, and it does not reflect the variation in scores when the measurement is assessed at different timepoints. Therefore, it is not sufficient to base the SDC on Cronbach's α .²⁴

John et al³¹ reported a SDC of 22.5 in 51 patients with resection interposition arthroplasty for carpometacarpal osteoarthritis. Although this is a different patient population, their methodology, and therefore their results, are more comparable to ours.

For both the PRWE and the total DASH, no (substantial) floor or ceiling effects were found. However, when

the SDC is taken into account, 45.4% and 46.6% of the patients were within the SDC range of a floor effect, respectively. No real health improvement beyond measurement error could be detected in this group of patients. Besides the aforementioned high SDC, this floor effect makes the PRWE and DASH less useful for individual patients with distal radial fracture in clinical practice. However, when measuring groups of patients (e.g. a randomized controlled trial), the SDC is reduced by a factor \sqrt{n} , where n patients is studied.²⁴

A strength of this study is that we used the COSMIN standards for adequate study design and statistical analysis. Of our large population of patients with a distal radial fracture, only ten patients (8%) were lost to follow-up for the test-retest reliability. Furthermore, we had no unanswered items in the assessment of the reliability for the PRWE and only two unanswered items for the total DASH. This was an advantage of the online questionnaires with required responses.

A limitation of this study is that we could not determine the responsiveness and minimal important change (MIC) of the PRWE and DASH in patients with a distal radial fracture. Responsiveness is defined as the ability to detect clinically important changes over time;³ MIC is part of the responsiveness measurement property. The MIC of PRWE and DASH was determined in patients who were treated non-operatively for isolated tendinitis, arthritis, or nerve compression syndromes from forearm to hand.³² These data cannot be generalized to patients with a distal radial fracture. Walenkamp et al¹⁶ determined the MIC for the PRWE in patients with a distal radial fracture. However, based on the COSMIN guidelines, this study lacks high methodological quality.⁴ Determining the MIC of the PRWE and DASH in patients with a distal radial fracture should be an important part of future research.

In conclusion, the Dutch versions of the PRWE and the DASH are valid and reliable PROMs in assessing function and disability in patients with a displaced distal radial fracture and are therefore useful in clinical research. Due to the high SDC, the PRWE and DASH are less useful for individual patients with a distal radial fracture in clinical practice.

References

1. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2:i-iv,1-74.
2. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19: 539-549.
3. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63: 737-745.
4. Kleinlugtenbelt YV, Nienhuis RW, Bhandari M, et al. Are validated outcome measures used in distal radial fractures truly valid? A critical assessment using the
- COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. *Bone Joint Res* 2016;5:153-161.
5. Gabl M, Krappinger D, Arora R, et al. Acceptance of patient-related evaluation of wrist function following distal radius fracture (DRF). *Handchir Mikrochir Plast Chir* 2007;39:68-72. (Article in German)
6. Hemelaers L, Angst F, Drerup S, Simmen BR, Wood-Dauphinee S. Reliability and validity of the German version of "the Patient-rated Wrist Evaluation (PRWE)" as an outcome measure of wrist pain and disability in patients with acute distal radius fractures. *J Hand Ther* 2008;21:366-376.
7. MacDermid JC, Turgeon T, Richards RS, Beadle M, Roth JH. Patient rating of wrist pain and disability: a reliable and valid measurement tool. *J Orthop Trauma* 1998;12:577-586.
8. MacDermid JC, Richards RS, Donner A, Bellamy N, Roth JH. Responsiveness of the short form-36, disability of the arm, shoulder, and hand questionnaire, patient-rated wrist evaluation, and physical impairment measurements in evaluating recovery after a distal radius fracture. *J Hand Surg Am* 2000;25:330-340.
9. Wilcke MT, Abbaszadegan H, Adolphson PY. Evaluation of a Swedish version of the patient-rated wrist evaluation outcome questionnaire: good responsiveness, validity, and reliability, in 99 patients recovering from a fracture of the distal radius. *Scand J Plast Reconstr Surg Hand Surg* 2009;43:94-101.
10. Lövgren A, Hellström K. Reliability and validity of measurement and associations between disability and behavioural factors in patients with Colles' fracture. *Physiother Theory Pract* 2012;28:188-197.
11. Mehta SP, Mhatre B, MacDermid JC, Mehta A. Cross-cultural adaptation and psychometric testing of the Hindi version of the patient-rated wrist evaluation. *J Hand Ther* 2012;25:65-77.
12. Kim JK, Kang JS. Evaluation of the Korean version of the patient-rated wrist evaluation. *J Hand Ther* 2013;26:238-243.
13. Schönnemann JO, Hansen TB, Søballe K. Translation and validation of the Danish version of the Patient Rated Wrist Evaluation questionnaire. *J Plast Surg Hand Surg* 2013;47:489-492.
14. Westphal T, Piatek S, Schubert S, Schuschke T, Winckler S. Reliability and validity of the upper limb DASH questionnaire in patients with distal radius fractures. *Z Orthop Ihre Grenzgeb* 2002;140:447-451. (Article in German)
15. Westphal T. Reliability and responsiveness of the German version of the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH). *Unfallchirurg* 2007;110:548-552. (Article in German)
16. Walenkamp MM, de Muinck Keizer RJ, Goslings JC, et al. The minimum clinically important difference of the Patient-rated Wrist Evaluation Score for patients with distal radius fractures. *Clin Orthop Relat Res* 2015;473:3235-3241.
17. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651-657.
18. Marsh JL, Slongo TF, Agel J, et al. Fracture and dislocation classification compendium - 2007: Orthopaedic Trauma Association classification, database and outcomes committee. *J Orthop Trauma* 2007;21(10 Suppl):S1-S133.
19. MacDermid JC. Development of a scale for patient rating of wrist pain and disability. *J Hand Ther* 1996;9:178-183.
20. MacDermid JC, Tottenham V. Responsiveness of the disability of the arm, shoulder, and hand (DASH) and patient-rated wrist/hand evaluation (PRWHE) in evaluating change after hand therapy. *J Hand Ther* 2004;17:18-23.
21. Videler AJ, Schreuders T. De PRWHE-DLV: De Nederlandse versie van de Patient rated Wrist/Hand Evaluation. *Nederlands tijdschrift voor Handtherapie* 2008;17:8-11. (In Dutch)
22. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med* 1996;29:602-608.
23. Veehof MM, Slegers EJ, van Veldhoven NH, et al. Psychometric qualities of the Dutch language version of the Disabilities of the Arm, Shoulder, and Hand questionnaire (DASH-DLV). *J Hand Ther* 2002;15:347-354.
24. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine*. Cambridge: Cambridge University Press, 2011.
25. Karanicolas PJ, Bhandari M, Kreder H, et al. Evaluating agreement: conducting a reliability study. *J Bone Joint Surg [Am]* 2009;91-A(Suppl 3):99-106.
26. Nunnally JC, Bernstein IH. *Psychometric Theory*. Third ed. New York: McGraw-Hill, 1994.
27. Lohr KN, Aaronson NK, Alonso J, et al. Evaluating quality-of-life and health status instruments: development of scientific review criteria. *Clin Ther* 1996;18:979-992.
28. Pynsent PB. Choosing an outcome measure. *J Bone Joint Surg [Br]* 2001;83-B:792-794.

29. **Scholtes VA, Terwee CB, Poolman RW.** What makes a measurement instrument valid and reliable? *Injury* 2011;42:236-240.
30. **McHorney CA, Tarlov AR.** Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995;4:293-307.
31. **John M, Angst F, Awiszus F, et al.** The patient-rated wrist evaluation (PRWE): cross-cultural adaptation into German and evaluation of its psychometric properties. *Clin Exp Rheumatol* 2008;26:1047-1058.
32. **Sorensen AA, Howard D, Tan WH, Ketchersid J, Calfee RP.** Minimal clinically important differences of 3 patient-rated outcomes instruments. *J Hand Surg Am* 2013;38:641-649.

Funding Statement

- This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Author Contribution

- Y. V. Kleinlugtenbelt: Design of study, Acquisition, Analysis and interpretation of data; Writing and revision of manuscript.
- R. G. Krol: Design of study, Acquisition, Analysis and interpretation of data, Revision of manuscript.
- M. Bhandari: Analysis and interpretation of data, Revision of manuscript.
- J. C. Goslings: Analysis and interpretation of data, Revision of manuscript.
- R. W. Poolman: Analysis and interpretation of data, Revision of manuscript.
- V. A. B. Scholtes: Design of study, Analysis and interpretation of data, Revision of manuscript.

Conflicts of Interest Statement

- The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

© 2018 Kleinlugtenbelt et al. This is an open-access article distributed under the terms of the Creative Commons Attributions licence (CC-BY-NC), which permits unrestricted use, distribution, and reproduction in any medium, but not for commercial gain, provided the original author and source are credited.