ORIGINAL ARTICLE

WILEY

Further psychometric validation and test–retest reproducibility of the WOUND-Q

Nina Vestergaard Simonsen[1] | Anne F. Klassen[2] | Charlene Rae[2] | Farima Dalaei[1] | Stefan Cano[3] | Lotte Poulsen[4] | Andrea L. Pusic[5] | Jens Ahm Sørensen[1]

[1]Department of Plastic Surgery, Odense University Hospital, Odense, Denmark, University of Southern Denmark, Odense, Denmark

[2]Department of Pediatrics, McMaster University, Hamilton, Canada

[3]Modus Outcomes, Letchworth Garden City, UK

[4]Research Unit for Plastic Surgery, Odense University Hospital, Løntoft, Nyhøj and Poulsen Plastic Surgery, Odense, Denmark

[5]Patient-Reported Outcomes, Value & Experience (PROVE) Center, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

Correspondence
Nina Vestergaard Simonsen, Department of Plastic Surgery, Odense University Hospital, University of Southern Denmark, Kløvervænget 26b, 1. sal to J.B. Winsløws Vej 4, Entrance 20 Penthouse 2. floor, 5000 Odense C, Denmark.
Email: nina.vestergaard.simonsen2@rsyd.dk; nina6533@hotmail.com

Abstract

WOUND-Q is a condition-specific patient-reported outcome measure developed for all types of chronic wounds, located anywhere on the body. To establish reliability and validity of a patient-reported outcome measure, multiple pieces of evidence are required. The purpose of this study was to examine the measurement properties of 9 of the 13 WOUND-Q scales and perform a test–retest reproducibility study in an international sample. In August 2022, we invited members of an international online community (Prolific.com) with any type of chronic wound to complete a survey containing the WOUND-Q scales, the Wound-QoL and EQ-5D. A test–retest survey was performed 7 days after the first survey. It was possible to examine the reliability and validity of eight of the nine WOUND-Q scales by Rasch Measurement Theory (RMT). To examine test–retest reproducibility intraclass correlation coefficients (ICCs), the standard error of the measurement and the smallest detectable change were calculated. In total, 421 patients from 22 different countries with 11 different types of chronic wounds took part in this study. Our analyses provided further evidence of the reliability and validity of the scales measuring wound characteristics (assessment, drainage, smell), health-related quality of life (life impact, psychological, sleep, social) and wound treatment (dressing).

KEYWORDS
health-related quality of life, patient-reported outcome measure, psychometrics, questionnaire, Rasch, reliability and validity, test–retest reliability, wounds

Key Messages
- WOUND-Q scales measuring health-related quality of life, wound characteristics and treatment were tested in 421 chronic wound patients from 22 different countries.
- We examined the psychometric performance of eight WOUND-Q scales using a modern psychometric approach (Rasch Measurement Theory) and performed a test–retest study.

- Psychometric analysis supported the reliability and validity of eight WOUND-Q scales.
- The smallest detectable change was also calculated for eight WOUND-Q scales.

1 | INTRODUCTION

Patient-reported outcomes measures (PROMs) are instruments that measure outcomes important to patients. The WOUND-Q is a condition-specific PROM designed to measure outcomes important to patients with any type of chronic wound(s) (≥3 months duration), located anywhere on the body.[1–4] This PROM was developed through in-depth interviews with 60 patients and input from 26 experts. The WOUND-Q was field-tested (FT) in a sample of 881 patients with 12 different types of wounds in Canada, Denmark, the Netherlands and the United States. After psychometric analysis, the final version of the WOUND-Q consisted of 13 scales and 111 items. The scales measure four domains (Figure 1), each scale ranges from 5 to 11 items, has 4 response options, and is scored separately on a scale of 0 (worst) to 100 (best).[2,3]

When evaluating the psychometric properties (i.e., reliability, validity, responsiveness) of a PROM, multiple pieces of evidence are required about its development and validation to determine whether results obtained when using the PROM can be trusted. To help researchers, clinicians and other stakeholders in this judgement, the COSMIN (Consensus-based Standard for the selection of health Measurement Instruments) initiative has published a list of criteria and tests for evaluating the quality of PROMs.[5] In a recent literature review[6] that applied the COSMIN guidelines of wound-specific PROMs, 33 PROMs were identified. Of these, 9 of 33 were developed for any type of chronic wounds and 17 of 33 were developed for specific wound types (e.g., venous leg ulcers). The WOUND-Q and the SCI-QOL were determined to be the highest quality PROMs in the review. However, as opposed to the SCI-QOL, the WOUND-Q can be used for all types of chronic wounds with any anatomic location, a reason why the authors recommend using the WOUND-Q in future research and clinical care.[6]

Psychometric validation of a new PROM is an ongoing process. The WOUND-Q field-test study included a clinical sample and focused on Rasch Measurement Theory (RMT) analysis. Some COSMIN criteria and tests were not examined in the original publication, including test–retest reliability and convergent validity. Therefore, the aim of this study was to further examine the reliability and validity of the WOUND-Q in a large international sample. Specifically, we evaluated its psychometric performance and performed a test–retest reproducibility study in an online international community-based sample.

2 | METHODS

Research ethic approval was obtained from McMaster University (Hamilton Integrated Research Ethics Board #14946) prior to the start of the survey licence to use the
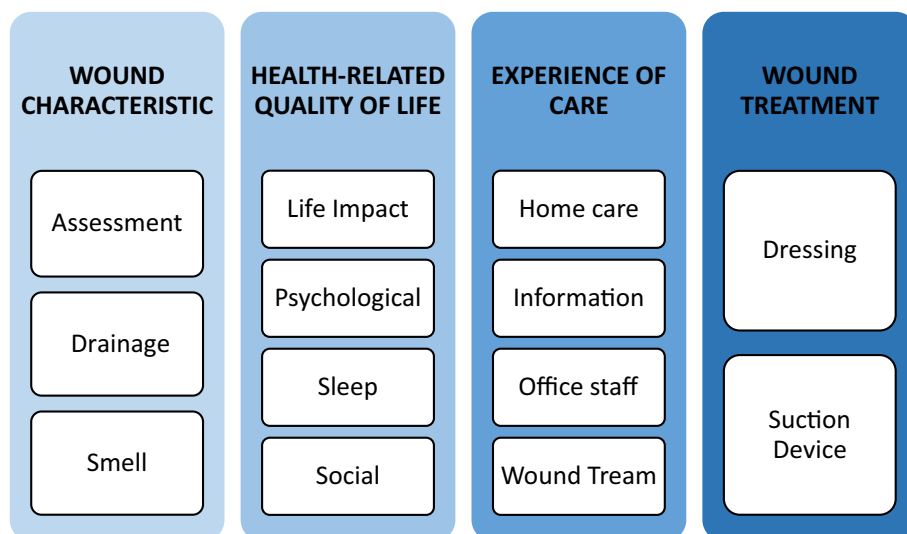

FIGURE 1 WOUND-Q conceptual framework.

EQ-5D-5L[7] and Wound-QoL[8] was obtained from licence holders.

3 | SAMPLE AND RECRUITMENT

Persons aged 18 years or older, with a chronic wound (lasted at least 3 months) of any type, anywhere on their body, and who were able to read, write and speak English were eligible for this study. Participants were recruited online through Prolific Academic (www.prolific.co). To identify a cohort for our study, prolific members were invited to complete a short screening survey in REDCap (Research Electronic Data Capture). The screening surveys were performed between August 22 and 29, 2022. Those who met the inclusion criteria and gave consent to participate were invited to the main survey, which ran from September 5 to 7, 2022. Participants were at minimum paid 10 £/h.

The survey included questions about demographics, wound characteristics, wound symptoms and wound treatment, 9 of the 13 WOUND-Q scales, the generic EQ-5D-5L (Canadian version),[9] and the condition-specific Wound-QoL-17 (US version).[10,11] Questions about wound symptoms were used to branch to specific WOUND-Q scales. Nine WOUND-Q scales were included: assessment, drainage, smell, psychological, sleep, dressing, vacuum, social and life impact, measuring the three domains: health-related quality of life (HRQL), wound characteristic and wound treatment. The 4 WOUND-Q scales covering experience of care were outside the scope of this study and therefore not included. Each WOUND-Q scale was scored using the scoring key to convert the raw score into a score that ranged from 0 to 100 (worst to best). To produce a raw score, at least 50% of the items in a scale must be completed. In the case of missing data (<50%), the within-person mean for the completed items was imputed for the missing items prior to computing the total raw score.[12] The EQ-5D-5L consists of two parts. The first part is a descriptive system for HRQL states, containing five items, measuring five domains (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). Each item is rated by five levels of severity. Using UK normative values,[13] the HRQL states were summarized into a single summary number (index value). The second part is the EQ-VAS where participants rate their overall current HRQL from 0 to 100 (worst to best).[9] The Wound-QoL-17 consists of 17 items and 3 subscales. (Body, Psyche, Everyday life).[14] A global score of disease-specific HRQL (range 0–4) was computed by summation of all items (range 0–68) divided by the number of completed items. At least 13 of the 17 items should be completed to be able to compute the global score. Lower scores indicate a better outcome.[14] At the end of the survey, participants were asked to indicate (yes/no) if they would be willing to complete the WOUND-Q again in 7 days as part of a test–retest reproducibility (TRT) study.

4 | ANALYSIS

The psychometric properties including reliability, internal consistency, measurement invariance, measurement error and construct validity as described as part of the COSMIN standards were examined.[5] RMT analysis[15] was performed in RUMM2030 software with the unrestricted Rasch model for polytomous data (RUMM version 2030, RUMM Laboratory Pty Ltd, Duncraig, Western Australia, Australia, 1998–2021). Descriptive analyses, test–retest and construct validity analyses were completed in IBM SPSS Statistics version 28 (IBM Corporation, Armonk NY, USA). The following tests were performed.

4.1 | Fit statistics

This test tells whether items work together to define a single variable (internal consistency). Three indicators of item fit to the Rasch model were examined: (1) item characteristic curves (IC); (2) item fit residuals (item-person interaction); and (3) χ^2 values (item-trait interaction). IC was inspected graphically. A fit residual within ±2.5 is considered ideal. Non-significant χ^2 values after Bonferroni adjustment support the items that fit the Rasch model.

4.2 | Category threshold order

Threshold maps were examined, to determine whether the response options worked as intended meaning that '1' on a four-point scale must sit lower in the continuum than '2', and so forth.

4.3 | Targeting

To investigate whether the items within each scale measure the construct as experienced by the sample, person-item threshold distribution and the proportion of the participants scoring outside the scale were examined graphically and statistically. A scale that is targeted to the construct it is intended to measure should have persons and item locations that mirror each other. Item locations were also inspected for gaps in measurement.[16]

The proportion of participants scoring the highest (ceiling) or lowest (floor) was calculated for each scale. Floor/ceiling effect is significant if ≥15%.[17] Participants were not forced to answer the WOUND-Q scales or items. To examine acceptability, the proportion of participants skipping items was calculated for each scale (i.e., missing data). Item-level missing data should be less than 10%.[18]

4.4 | Differential item functioning

To examine whether items in the scales were responded to differently by participants from the original field-test study[3] and the Prolific sample, we examined differential item functioning (DIF). For each scale, we selected random samples with a maximum combined sample size of 500, with each group balanced in size. The DIF analysis was performed three times for each scale to determine whether the results were stable. Items with significant χ^2 values after Bonferroni adjustment in the DIF analysis were split by sample characteristic. We then conducted Pearson correlations between the original and the new person locations to determine the impact of DIF on scoring.

4.5 | Local dependency

This test examines whether a person' response to an item on a scale influences their response to another item on the same scale. Item pairs with correlations >0.20 were included in subtest analysis, to determine their impact on scale reliability.

4.6 | Reliability

The scales' ability to discriminate between people in the sample was examined by the Person Separation Index (PSI). Cronbach alpha was examined to judge internal consistency, and intraclass correlation coefficients (ICCs) to examine test–retest reliability. All reliability coefficients were considered acceptable if ≥0.70.[5,17]

For the TRT analysis, participants who reported important change in the scale construct were excluded. Normal distribution of the score difference between the first (T1) and second (T2) assessment was visually inspected by Q–Q plots. Eventual outliers were visualized on boxplots and removed from the TRT analysis. The two-way random effect model evaluating for consistency was used to calculate ICC.

4.7 | Measurement error

To determine the amount of measurement error in the score of individuals who reported no important change between the test (T1) and retest (T2), we computed the standard error of measurement (SEM). We used the formula $SEM_{consistency} = (T1(SD) + T2(SD)/2) * \sqrt{(1 - ICC_{consistency})}$. From the SEM, we determined the smallest detectable change (SDC) which is the minimal change that must occur in a score, to be real and not due to measurement error. The SDC was computed at an individual ($SDC_{ind} = 1.96 * \sqrt{2} * SEM$) and group level ($SDC_{group} = SDCind/\sqrt{n}$).[16,17]

[Correction added on 21 December 2023 after first online publication: In the preceding sentence, SDC_{ind} has been corrected to (SDC) and $\sqrt{(2 * SEM)}$ has been corrected to ($SDC_{ind} = 1.96 * \sqrt{2} * SEM$) in this version of the article.]

4.8 | Construct validity

The scores of EQ-5D, Wound-QoL and WOUND-Q were used to test hypotheses to establish construct validity. Normality of data was examined using kurtosis, skewness and visual inspection of normal Q–Q plots. Kurtosis and skewness should be within ±2 for data to be normally distributed.[19] Independent *t*-tests were used to explore differences between groups. Non-parametric analysis was used if data were not normally distributed.

First, we hypothesized that the strongest correlation with WOUND-Q scale scores would be with scales within the same domain, and weaker correlations with scales measuring other domains.

Second, to test convergent validity (i.e., correlation to PROMs measuring a similar construct), correlations of WOUND-Q to Wound-QoL and EQ-5D were examined. COSMIN criteria state that correlations between scales measuring similar constructs should be ≥0.50 and <0.3 between scales measuring dissimilar constructs.[5] Therefore, we hypothesized that correlations between the HRQL scales of WOUND-Q and EQ-5D and Wound-QoL would be ≥0.50. Furthermore, we expected EQ-5D correlations of 0.30–0.50 with the wound characteristic scales, and <0.3 with the treatment scales. Third, we tested four clinical hypotheses available in Table 1.

5 | RESULTS

A total of 1248 prolific members were screened, and of these, 671 were invited to complete the survey. The response rate was 76.5%, resulting in 421 study participants. Figure 2 shows a flow chart of participant

TABLE 1 Hypotheses tested known group differences.

Hypothesis (First group will score lower)	Assessment	Life Impact	Psych	Social	Sleep[a]
1. Participants with >1 wound will have lower scale scores compared to those who had 1 wound (>1 wound vs. 1 wound)	Worse*	Worse*	Worse*	Worse*	Worse*
2. Participants who report that their wound has smell will have lower scale scores compared with those who have no smell (smell vs. no smell)	Worse*	Worse*	Worse*	Worse*	N/A
3. Participants who report drainage of their wound will have lower scale scores than participants who report no drainage (drainage vs. no drainage)	Worse*	Worse*	Worse*	Worse*	N/A
4. Higher occurrence of sleep disturbances[b] will be moderately (0.3–0.5) negatively correlated with scale scores.	−0.46*	−0.5*	−0.48*	−0.5*	−0.56*

[a]Only for participants reporting sleep disturbances in the past week.
[b]Spearman's correlation due to categorical variable.
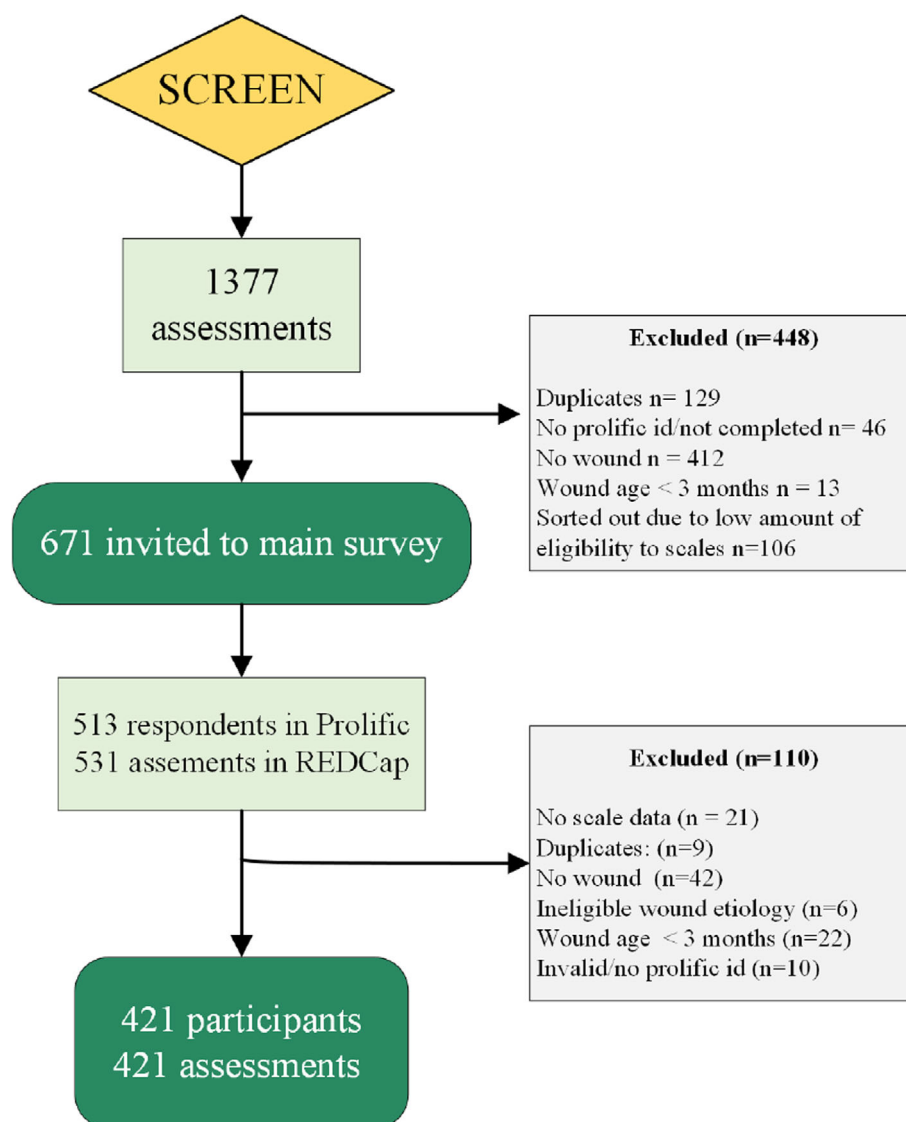*$p < 0.001$.

FIGURE 2 Participant flowchart.

TABLE 2 Demographic and clinical characteristics of the sample.

	n	*%*
Age (mean; SD)	37.9 years; 13.6 years	Range 18–84
Gender		
Man	212	50.4
Woman	204	48.5
Other	5	1.2
BMI		
Underweight (<18.5)	33	7.8
Normal weight (18.5–24.9)	135	32.1
Overweight (25–29.9)	120	28.5
Obese (≥30)	128	30.4
Missing	5	1.2
Country		
Australia	3	0.7
Canada	13	3.1
Chile	2	0.5
Czech Republic	2	0.5
Denmark	1	0.2
Estonia	3	0.7
France	3	0.7
Germany	2	0.5
Greece	6	1.4
Hungary	6	1.4
Ireland	7	1.7
Italy	9	2.1
Latvia	2	0.5
Netherlands	1	0.2
New Zealand	3	0.7
Poland	18	4.3
Portugal	18	4.3
Slovenia	1	0.2
South Africa	81	19.2
Spain	3	0.7
UK	145	34.4
USA	89	21.1
Prefer not to answer	3	0.7
Ethnicity		
White	278	66.0
Black	87	20.7
Other	55	13.1
Missing	1	0.2

(Continues)

TABLE 2 (Continued)

	n	*%*
English as the first language		
Yes	329	78.1
Missing	2	0.5
Smoking/vaping		
Yes	132	31.4
Marital status		
Never married	193	45.8
Separated	10	2.4
Divorced	22	5.2
Widowed	4	1.0
Living common-law	42	10.0
Married	142	33.7
Other	5	1.2
Prefer not to answer	3	0.7
Highest level of education		
Some high school	6	1.4
Completed high school	52	12.4
Some college or trade school or university	83	19.7
Completed college or trade school or university degree	179	42.5
Some Masters or Doctoral degree	35	8.3
Completed Masters or Doctoral degree	65	15.4
Prefer not to answer	1	0.2
Work (able to choose multiple)		
Full time (≥37 h)	237	56.3
Part time (<37 h)	70	16.6
Unemployed	38	9.0
Student	30	7.1
Retired	17	4.0
On leave	2	0.4
Sick leave	23	5.5
Other	9	2.1
Prefer not to answer	2	0.5
Comorbidities (able to choose multiple)		
Amputated	80	19.0
Diabetes	91	21.6
Peripheral Artery Disease (PAD)	22	5.2
Peripheral Venous Disease (PVD)	31	7.4

(Continues)

TABLE 2 (Continued)

	n	%
Paralysed	5	1.2
Hypertension	98	23.3
Ischaemic heart disease	22	5.2
Heart failure	8	1.9
Arthritis	51	12.1
Cancer	8	1.9
Chronic obstructive pulmonary disease (COPD)	14	3.3
Kidney failure	5	1.2
Neuropathy	51	12.1
Stroke	8	1.9
Inflammatory bowel disease	11	2.6
Skin disease (inflammatory and autoimmune)	11	2.6
Multiple sclerosis	4	1.0
Psychological disorder	3	0.7
Other	67	15.9
Number of chronic wounds		
1	298	70.8
2	80	19.0
3	27	6.4
4	5	1.2
5	5	1.2
6	3	0.7
7	1	0.2
8	1	0.2
10	1	0.2
Wound location (able to choose multiple)		
Face or neck	13	3.1
Hand	8	1.9
Arm	34	8.1
Shoulder	10	2.4
Chest	14	3.3
Abdomen	36	8.6
Back	27	6.4
Buttocks	24	5.7
Genitals	10	2.4
Leg	104	24.7
Foot	87	20.7
Toe(s)	41	9.7
Other	8	1.9

(Continues)

TABLE 2 (Continued)

	n	%
Wound type		
Diabetic foot ulcer	37	8.8
Venous ulcer	8	1.9
Arterial ulcer	9	2.1
Pressure ulcer	17	4.0
Surgery	68	16.2
Radiation	3	0.7
Trauma/injury	129	30.6
Hidradenitis suppurativa	21	5.0
Pilonidal abscess	12	2.9
Multiple	53	12.6
Not sure	30	7.1
Other	34	8.1
Wound age		
3–6 months	215	51.1
7–12 months	47	11.2
1–2 years	66	15.7
3–4 years	39	9.3
5–10 years	41	9.7
11–30 years	10	2.4
>30 years	2	0.5
Prefer not to answer	1	0.2
Wound size (width × length) (cm^2)		
Median; range	2.04 cm^2	Range 0.01–1282.1 cm^2
Wound size (cm^2)		
<1	125	29.7
1–2.4	75	17.8
2.5–4.9	53	12.6
5–9.9	41	9.7
10–24.9	41	9.7
>24.9	49	11.6
Missing	37	8.8

selection. The study sample was resident in 22 countries. Most participants ($n = 298$, 70.8%) had one wound, and half of the sample ($n = 212$, 50.4%) had a history of recurrent chronic wounds. In the past 12 months, most of the sample ($n = 343$, 81.5%) had seen a doctor or nurse about their wound. The chronic wounds were primarily located on the lower extremity ($n = 233$, 55.3%) and the most common cause of the wound was an injury/trauma ($n = 129$, 30.6%). Sample characteristics are shown in Table 2. Based on the past week, 205 (48.7%) reported having drainage, 171 (40.8%) reported smell from their

TABLE 3 WOUND-Q scales—RMT scale level statistics.

Scale	No. completed scale	No. included in RMT	Scored on scale %	χ^2	DF	p-value	PSI + extr[a]	PSI − extr[a]	α + extr[a]	α − extr[a]	DT	±2.5	Chi	Residuals >0.20[a]	Floor %	Ceiling %	Missing %
Assessment	421	407	96.7	76.01	55	0.03	0.85	0.84	0.89	0.88	0	1	0	3	0.5	2.9	11.2
Drainage	206	193	93.7	22.33	16	0.13	0.86	0.85	0.90	0.88	0	0	0	0	0.5	5.8	3.9
Smell	171	159	93	19.85	16	0.23	0.87	0.86	0.92	0.91	0	0	0	0	1.2	5.8	7.6
Life Impact	420	385	89.7	85.56	48	<0.001	0.87	0.86	0.93	0.91	0	2	1	2	1.2	7.1	5.2
Psychological	419	390	93.1	91.28	50	<0.001	0.90	0.90	0.94	0.92	0	3	0	2	0.7	6.4	4.2
Social	416	307	73.8	33.32	20	0.03	0.85	0.82	0.93	0.86	0	0	0	0	3.4	22.8	4.0
Sleep	294	274	93.2	21.08	15	0.13	0.84	0.79	0.88	0.83	0	1	0	0	5.4	1.4	2
Dressing	207	190	91.8%	10.58	18	0.91	0.83	0.81	0.87	0.84	1	0	0	2	0	8.2	8.1
Suction	14	N/A															

[a]Number of item pairs.

Abbreviations: DF, degrees of freedom; DT, disordered threshold; extr, extremes; missing data, proportion of eligible participants who skipped at least 1 item in the scale; no = number; PSI, Person Separation Index; RMT, Rasch Measurement Theory.

wound, and 294 (69.8%) had sleep interference due to their wound. In the past 3 months, 209 (49.6%) had used a dressing on their wound. Only 14 (3.3%) participants reported using a suction device, therefore the vacuum scale was excluded from further analyses.

5.1 | Psychometric findings

The RMT analysis was conducted for eight WOUND-Q scales with a total of 64 items. Results of RMT analysis for each scale are described below and are shown in Table 3.

5.1.1 | Fit statistics

The 64 items in the eight WOUND-Q scales examined had a good fit to the Rasch model. Overall, 57 were inside the item fit criteria of ±2.5, and 63 had non-significant χ^2 p-value after Bonferroni adjustment. The item with a significant p-value evidenced reasonable adherence between observed and predicted values on graphical inspected on the IC for 1 scale, resulting in less concern about misfit (see Appendix A). In addition to item fit, 63 items had ordered thresholds. The item with disordered thresholds was in the dressing scale. Overall, the data for four scales fit the Rasch model, with marginal misfit for two scales. The two scales with some misfit of data to the Rasch model were HRQL scales (i.e., life impact and psychological).

5.1.2 | Targeting

The sample was well targeted to the scales. For seven scales, close to 90% of respondents scored on the scale (see Table 3). The maximum floor and ceiling value for seven of the scales was 5.4% and 8.2%, respectively. The social scale had 22.8% of participants scoring at the ceiling. The assessment scale was the only scale with >10% of at least one missing item in the scale (missing data), no item was systematically skipped.

Appendix B shows the person-item threshold distribution for all scales. Participants who scored outside the scale range primarily scored to the right in each figure (better outcome). Most scales had good coverage for the concept. Only the sleep scale evidenced a notable measurement gap, which was between item logits 0.25–2.

5.1.3 | Differential item functioning

For the DIF analysis, the sample was below 500 for the scales smell, drainage and dressing, because these scales

TABLE 4 Test–retest statistics, SEM and SDC.

SCALE	Excluded due to change (n)	Included ± extr (n)		ICC	95% CI		SEM	SDC$_{ind}$*	SDC$_{group}$
Assessment	16	−	84	0.83	0.74	0.89	5.93	16.45	1.79
		+	86	0.77	0.65	0.85	7.11	19.72	2.13
Drainage	16	−	79	0.88	0.81	0.92	6.48	17.96	2.02
		+	81	0.85	0.76	0.92	7.23	20.05	2.23
Smell	15	−	79	0.83	0.73	0.89	7.08	19.62	2.21
		+	87	0.68	0.51	0.79	10.75	29.81	3.20
Life impact	23	−	74	0.92	0.87	0.95	6.35	17.60	2.05
		+	79	0.85	0.76	0.90	9.06	25.10	2.82
Psychological	31	−	64	0.93	0.89	0.96	4.60	12.75	1.59
		+	71	0.84	0.74	0.90	7.45	20.66	2.45
Sleep	35	−	63	0.94	0.90	0.96	5.44	15.08	1.90
		+	67	0.89	0.82	0.93	7.31	20.27	2.48
Social	21	−	73	0.96	0.94	0.98	5.86	16.23	1.90
		+	81	0.92	0.87	0.95	8.71	24.13	2.68
Dressing	13	−	81	0.83	0.73	0.89	6.86	19.02	2.11
		+	88	0.61	0.40	0.74	10.94	30.33	3.23

Abbreviations: ICC, intraclass correlation coefficient; SDC, smallest detectable change; SEM, standard error of the measurement.

*SDC$_{group}$ should be applied when considering group-level statistics. The SDC$_{ind}$ is applied at the patient level and should be interpreted with caution. Since SDC is based on classical test theory (CTT) it assumes the SEM around an individual score is constant regardless of a person's location on the scale. However, in Rasch analysis the precision of the scale is greatest at the centre and lowest at the floor and ceiling, therefore error is dependent upon the location of a person on the scale rather than a constant like in CTT.

[Correction added on 21 December 2023 after first online publication: In table 4, a footnote was added, and the data in column SDC$_{ind}$ and SDC$_{group}$ have been corrected in this version.]

were completed by fewer participants. Significant DIF was found for 11 of 64 items, and of these, DIF occurred in all three random samples for three items, Appendix A. When the 11 items were split by sample (FT or Prolific), the Pearson correlations between the original and split person locations were ≥0.999, showing an insignificant impact on the scoring.

5.1.4 | Dependency

Four scales (assessment, life impact, psychological and dressing) had 1–3 item pairs with item residual correlations >0.20. The subtest analysis for the correlated items showed a maximum drop in the PSI value at 0.03, showing minor influence on scale reliability, Table 3.

5.1.5 | Reliability

All scales had PSI values with and without extremes at >0.83 and >0.79, respectively. Cronbach alpha values

were all >0.82. Table 4 shows the results for the TRT. In total, 209 participants completed at least one scale in the TRT. Depending on the scale, between 67 and 88 reported no important change and were eligible for TRT analysis. ICC values with outliers ranged from 0.61 to 0.92 and without outliers from 0.83 to 0.96.

5.1.6 | Measurement error

The SEMs ranged from 4.6 to 7.1, which resulted in SDC$_{ind}$ of 12.8–30.3 and SDC$_{group}$ of 1.6–3.2.

5.1.7 | Construct validity

All hypotheses tested were confirmed. Correlations for the construct validity analysis are shown in Tables 1 and 5 and subgroup characteristics for one of the clinical hypotheses are available in Appendix C. Scale scores tended to correlate most strongly with scales in their top-level domain, with some exceptions (sleep and social).

TABLE 5 Pearson correlations between WOUND-Q scales, EQ-5D-5L and Wound-QoL.

Domain	Scale	Assessment	Drainage	Smell	Life impact	Psychological	Sleep	Social	Dressing
Wound characteristics	Drainage	0.685**							
	Smell	0.394**	0.538**						
HRQL	Life impact	0.558**	0.645**	0.549**					
	Psychological	0.586**	0.565**	0.430**	0.741**				
	Sleep	0.387**	0.323**	0.308**	0.509**	0.587**			
	Social	0.491**	0.526**	0.441**	0.757**	0.685**	0.520**		
Wound treatment	Dressing	0.443**	0.495**	0.306**	0.400**	0.366**	0.214**	0.437**	
	EQ-5D[a]	0.322**	0.181**	0.280**	0.525**	0.552**	0.512**	0.513**	0.165*
	Wound-Qol	−0.512**	−0.510**	−0.433**	−0.671**	−0.604**	−0.545**	−0.673**	−0.322**

[a]UK index score.

**Correlation is significant at the 0.01 level (two-tailed).

*Correlation is significant at the 0.05 level (two-tailed).

Mean scores on the assessment and HRQL scales were significantly ($p < 0.001$) lower for the group with multiple wounds and greater incidence of sleep interference (see Figure 3a,b). In terms of convergent validity, the four HRQL scales in the WOUND-Q correlated most strongly with the Wound-QoL score, as expected. All WOUND-Q scales (except dressing and smell) had strong correlations (>0.50) with the Wound-QoL score. Strong correlations were also found between the EQ-5D index score and the WOUND-Q scales in the HRQL domain.

6 | DISCUSSION

The WOUND-Q is a wound-specific PROM developed by following a multi-phased mixed methods approach that adhered to international guidelines of PROM development.[20–24] Compared with other wound-specific PROMs, the WOUND-Q is unique as it is the only carefully developed PROM that met COSMIN criteria for PROM quality[6] that measures outcomes important to patients with any type of chronic wounds, located anywhere on the body. The use of RMT analysis in its development means that each of the 13 scales is unidimensional with interval-level measurement properties (i.e., each scale function like a 'ruler'), and that only items with the best psychometric performance were retained in the final WOUND-Q.[15,16]

In this study, we examined the psychometric performance and performed a TRT of 8 WOUND-Q scales in their final form in an international sample of people with chronic wounds resident from 22 countries. RMT analysis was performed to investigate how the data collected performed in regard to the expectations of the Rasch model, which predicts how items must perform to create a reliable and valid total score for each scale.[16] It should be noted that the Prolific sample differed from the FT sample for several patient characteristics (Appendix D). Prolific participants were considerably younger with a mean age of 37.9 years, conditions associated with chronic wounds, such as diabetes, venous or arterial insufficiency were less represented and the number, size and duration of wounds were fewer, smaller and less compared with the FT sample. These differences were expected since the FT study recruited from hospitals, where many wound patients are older age and have more complex wounds that can be slow to heal.[25]

The RMT and TRT analysis provided further evidence supporting the reliability and validity of the eight WOUND-Q scales, which met the COSMIN criteria of good measurement properties, in this younger and more international community-based chronic wound sample.[5] All scales evidenced reliability with PSI, Cronbach alpha
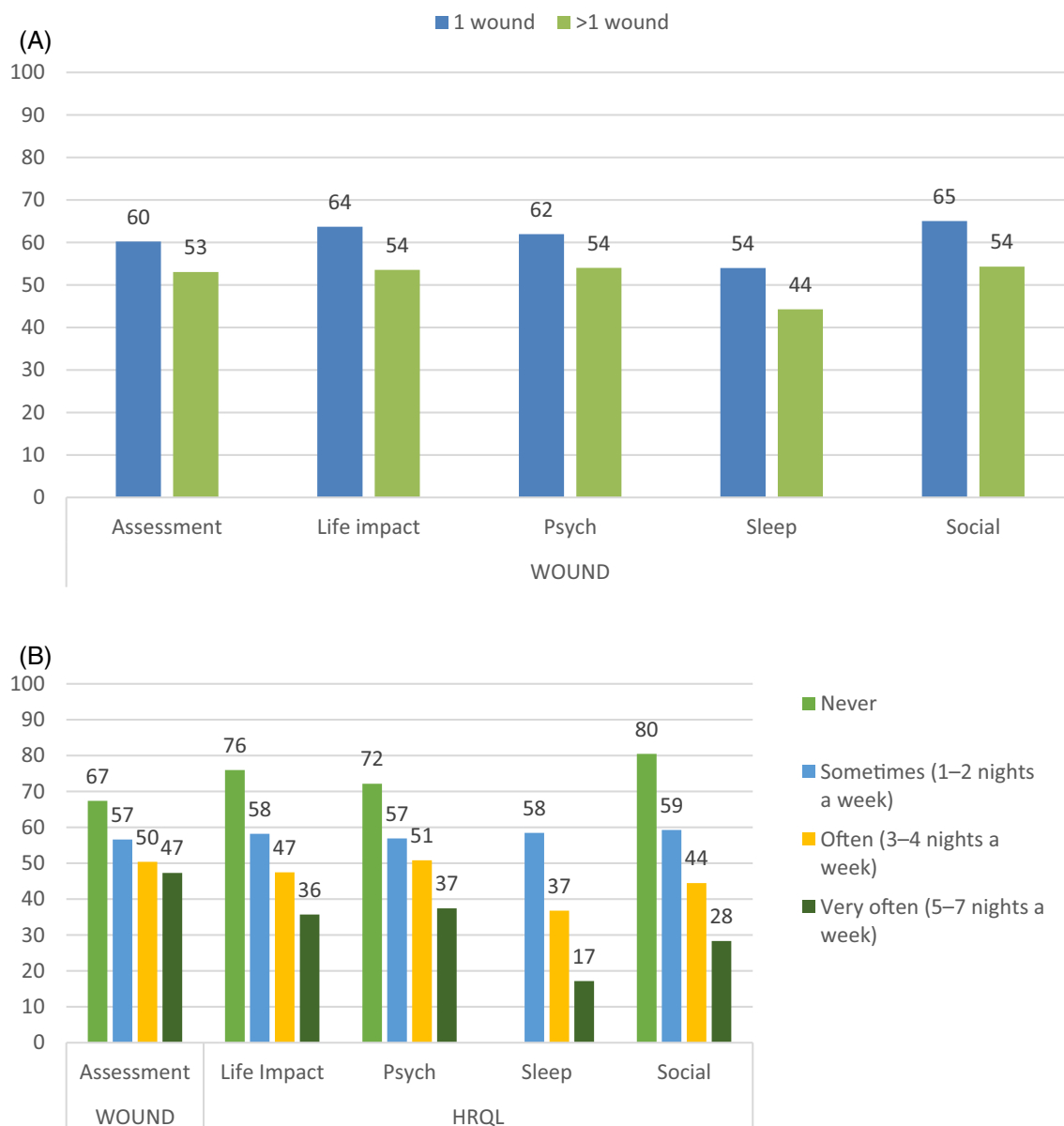
FIGURE 3 (A) Mean scores of participants with one or multiple wounds. (B) Mean scores of participants by sleep disturbances in the past week.

and ICC values >0.70 exceeding the COSMIN criteria. Furthermore, the SDC is now available for these eight scales, stating the size of change in an score that is true, and not due to measurement error, which are important to statistically interpret change in scores.[26] Construct validity was demonstrated by hypothesis testing, and convergent validity was demonstrated for the HRQL scales in comparison to the Wound-QoL and EQ-5D.

Some limitations in the psychometric performance of the WOUND-Q scales were found. While data for most scales either fit or had marginal misfit to the Rasch model, there was more misfit for two of the HRQL scales. In the life impact scale, the item 'your close relationships' was the only item tested that evidenced misfit to

the Rasch model, which likely accounts for the overall misfit for this scale. Unlike the FT sample that used translations for people who did not speak English, the Prolific sample was conducted entirely in English and included 90 (21.4%) people for whom English was not their first language. Therefore, it is possible that this item was misunderstood, resulting in error in the data and misfit. This could have been examined in DIF analysis; however, due to subgroup sample of less than 150, we were unable to do so. The other item that underperformed was in the dressing scale. Specifically, item 3 (absorb) had a disordered threshold, meaning that the response options did not function as intended. As for the item with misfit, it is possible this finding could be

due to a misunderstanding of the item or that the item was less relevant in the non-clinical sample, as 70 of the 207 patients that completed the dressing scale had no drainage from their wound.[16]

Despite differences between the international sample of Prolific participants and the clinic-based FT sample in terms of DIF, correlations between the person locations before and after adjusting for DIF provided evidence to support the use of the original WOUND-Q scoring key.

In the social scale, significant ceiling effect (22.8%) was noted. This finding was also the case in the original WOUND-Q paper, with 27.3%.[3] Together, these findings show that social life is not affected for some people with chronic wounds. This finding is in line with other research, showing that chronic wounds in general impact social life, however to different degree and for different reasons.[3,27] Those who scored at the top of the social scale tended to have smaller wounds (median 1.8 cm² vs. 2.35 cm²), less wound symptoms (sleep interference, smell, exudate) and less contact with health care professionals. However, ceiling effects should be acknowledged if the goal is to measure change.[17] The sleep scale had some limitation in measuring the entire level of sleep disturbances, illustrated by a gap on the person-item map, Appendix B. The shortness of the scale, with only five items and the younger population compared with the FT sample[3] may account for this finding. Studies have shown that wound-related pain impairs sleep[28,29] and that arterial and venous ulcers have more frequent pain.[30,31] As our sample consisted of a different distribution of wound types with fewer people with venous and arterial ulcers, this may affect the range of experienced sleep disturbances.

Our study had several limitations. First, our sample was obtained from the online platform Prolific. Using Prolific makes it possible to recruit a large sample of participants in a short amount of time. Disadvantages are that payment may influence the participants,[32,33] people self-select for participation and that all data are self-reported. However, data from online platforms have been found to be reliable and valid.[34–36] Second, only 14 participants had used a suction device and we were unable to perform the RMT analysis for this scale. Third, while we were able to calculate the SDC for each of the eight scales, it is important to note that these changes are from a much younger population than the development sample and that the SDC is solely a statistical expression, not stating whether this change is important for the patients.[26] To be able to interpret meaningful change in scores, longitudinal studies are needed to examine the ability of the WOUND-Q to measure change and to determine the minimal important difference.

7 | CONCLUSION

In an international community-based sample of people with chronic wound who were younger and less ill compared with the FT sample, the WOUND-Q evidenced strong psychometric performance. The RMT analysis provided further evidence that 8 of the 13 WOUND-Q scales were reliable and valid. Furthermore, this study provided evidence of the test–retest reliability and values for SDC for eight of the WOUND-Q scales, both of which have not been previously reported in the literature. More information on the WOUND-Q can be found at https://qportfolio.org/wound-q/.

CONFLICT OF INTEREST STATEMENT
Klassen and Pusic are co-developers of the WOUND-Q and receive a share of any licence revenue on the inventor sharing policies from the institutions that own the WOUND-Q. Anne Klassen is the owner of EVENTUM Research, which provides consulting services to the pharmaceutical industry. None of the other authors have a conflict of interest.

DATA AVAILABILITY STATEMENT
The corresponding author can provide the supporting data for this study upon request. However, please note that the data cannot be made publicly available due to privacy or ethical restrictions.

ORCID
Nina Vestergaard Simonsen https://orcid.org/0000-0001-6006-9753

REFERENCES

1. Klassen A, van Haren ELWG, Cross K, et al. International mixed methods study protocol to develop a patient-reported outcome measure for all types of chronic wounds (the WOUND-Q). *BMJ Open*. 2020;10(3):e032332.
2. Squitieri L, Tsangaris E, Klassen AF, et al. Patient-reported experience measures are essential to improving quality of care for chronic wounds: an international qualitative study. *Int Wound J*. 2020;17:1052-1061.
3. Klassen AF, Haren ELWG, Alphen TC, et al. International study to develop the WOUND-Q patient-reported outcome measure for all types of chronic wounds. *Int Wound J*. 2021;18(4):487-509.
4. van Alphen TC, Poulsen L, van Haren ELWG, et al. Danish and Dutch linguistic validation and cultural adaptation of the WOUND-Q, a PROM for chronic wounds. *Eur J Plast Surg*. 2019;42:495-504.

5. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-1157.

6. van Alphen TC, ter Brugge F, van Haren ELWG, Hoogbergen MM, Rakhorst H. SCI-QOL and WOUND-Q have the best patient-reported outcome measure design: a systematic literature review of PROMs used in chronic wounds. *Plast Reconstr Surg Glob Open*. 2023;11(1):e4723.

7. EQ-5D: EuroQoL Research Foundation; 2023. Available from: https://euroqol.org/

8. Wound-QoL licensing 2023. Available from: https://www.wound-qol.com/licensing/

9. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337-343.

10. Blome C, Baade K, Sebastian Debus E, Price P, Augustin M. The "wound-QoL": a short questionnaire measuring quality of life in patients with chronic wounds based on three established disease-specific instruments. *Wound Repair Regen*. 2014;22(4):504-514.

11. Augustin M, Conde Montero E, Zander N, et al. Validity and feasibility of the wound-QoL questionnaire on health-related quality of life in chronic wounds. *Wound Repair Regen*. 2017;25(5):852-857.

12. WOUND-Q A User's Guide for Researchers and Clinicians 2023. Available from: https://qportfolio.org/wp-content/uploads/2023/01/WOUND-Q-USERS-GUIDE.pdf

13. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. *Health Econ*. 2018;27(1):7-22.

14. Wound-QoL User Manual. ENGLISH version https://www.wound-qol.com/download/2023. Available from: https://www.wound-qol.com/wp-content/uploads/User-manualWound-QoL2023-02-10.pdf

15. Rasch G. *Studies in Mathematical Psychology: 1*. Danmarks pædagogiske Institut; 1960.

16. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess*. 2009;13(12) iii, ix-x:1-177.

17. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.

18. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. *Soc Sci Med*. 1998;46(12):1569-1585.

19. Kim HY. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restor Dent Endod*. 2013;38(1):52-54.

20. Aaronson N, Alonso J, Burnam A, et al. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res*. 2002;11(3):193-205.

21. Administration USFD. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims 2009. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims

22. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 2—assessing respondent understanding. *Value in Health*. 2011;14(8):978-988.

23. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value in Health*. 2005;8(2):94-104.

24. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. *Value in Health*. 2011;14(8):967-977.

25. Gould L, Abadir P, Brem H, et al. Chronic wound repair and healing in older adults: current status and future research. *Wound Repair Regen*. 2015;23(1):1-13.

26. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes*. 2006;4:54.

27. Klein TM, Andrees V, Kirsten N, Protz K, Augustin M, Blome C. Social participation of people with chronic wounds: a systematic review. *Int Wound J*. 2021;18(3):287-311.

28. Gonzalez-Consuegra RV, Verdu J. Quality of life in people with venous leg ulcers: an integrative review. *J Adv Nurs*. 2011;67(5):926-944.

29. Hellström A, Nilsson C, Nilsson A, Fagerström C. Leg ulcers in older people: a national study addressing variation in diagnosis, pain and sleep disturbance. *BMC Geriatr*. 2016;16:25.

30. Paul J. A cross-sectional study of chronic wound-related pain and itching. *Ostomy Wound Manage*. 2013;59(7):28-34.

31. Price PE, Fagervik-Morton H, Mudge EJ, et al. Dressing-related pain in patients with chronic wounds: an international patient perspective. *Int Wound J*. 2008;5(2):159-171.

32. Brühlmann F, Petralito S, Aeschbach LF, Opwis K. The quality of data collected online: an investigation of careless responding in a crowdsourced sample. *Methods Psychol*. 2020;2:100022.

33. Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull*. 1999;125(6):627-668; discussion 92–700.

34. Chandler J, Shapiro D. Conducting clinical research using crowdsourced convenience samples. *Annu Rev Clin Psychol*. 2016;12:53-81.

35. Kelly JD, Comstock B, Kowalewski TM, Smartt JM. Crowdsourced reliability of an Assessment of lower facial aging using a validated visual scale. *Plast Reconstr Surg Glob Open*. 2021;9(1):e3315.

36. Strickland JC, Stoops WW. The use of crowdsourcing in addiction science research: Amazon Mechanical Turk. *Exp Clin Psychopharmacol*. 2019;27(1):1-18.

APPENDIX A: RMT ITEM FIT STATISTICS AND DIFFERENTIAL ITEM FUNCTION (DIF) RESULTS

Scales	Item	Item Fit Statistics							DIF[a]
		Location	SE	Fit residual	DF	χ^2	DF	Probability	
Assessment									
	Bleeding	−0.83	0.08	2.22	365.68	4.89	5	0.43	
	Smell	−0.75	0.08	0.02	360.27	1.79	5	0.88	
	Burning	−0.14	0.07	2.78	358.47	15.28	5	0.01	
	Holes	−0.46	0.07	0.25	360.27	3.90	5	0.56	
	Swelling	0.28	0.07	−0.39	364.77	3.11	5	0.68	
	Edges	0.16	0.07	−0.59	362.07	5.11	5	0.40	
	Colour	0.15	0.07	−0.59	360.27	6.92	5	0.23	
	Drainage	−0.08	0.07	−0.53	361.17	2.39	5	0.79	
	Pain	0.95	0.07	0.41	365.68	4.85	5	0.43	1, 3
	Deep	0.15	0.07	−1.82	360.27	12.26	5	0.03	
	Size	0.57	0.07	−1.75	362.07	15.51	5	0.01	3
Drainage									
	Colour	−0.24	0.12	0.21	165.84	1.47	2	0.48	
	Thick	−0.50	0.11	0.31	164.98	2.12	2	0.35	
	Smell	−0.13	0.10	1.30	163.26	3.83	2	0.15	
	Noticing	0.11	0.10	1.15	164.98	2.36	2	0.31	
	Clothes	0.00	0.10	−0.75	165.84	0.28	2	0.87	
	Enjoy life	0.47	0.11	−0.63	164.98	1.76	2	0.41	
	Amount	0.03	0.12	−1.99	164.12	8.56	2	0.01	
	Dressing	0.26	0.11	0.96	164.98	1.94	2	0.38	1, 2
Smell									
	Relationships	−0.12	0.12	1.27	135.16	4.28	2	0.12	
	Comments	−0.63	0.13	−0.54	136.02	1.29	2	0.53	
	Social life	−0.16	0.13	−1.28	133.45	3.14	2	0.21	
	Dressing on	−0.66	0.14	0.29	134.30	1.14	2	0.57	
	Noticing	−0.17	0.13	−2.21	133.45	5.45	2	0.07	
	Stopping	0.45	0.12	0.01	135.16	0.29	2	0.87	
	Unpleasant	0.71	0.13	0.22	134.30	3.85	2	0.15	
	Dressing off	0.56	0.13	1.07	135.16	0.41	2	0.82	
Life Impact									
	Relationships	−0.67	0.08	3.46	332.78	24.33	6	0.00**	
	Relax	0.46	0.08	1.05	331.05	8.27	6	0.22	1, 2, 3
	Emotional	0.50	0.08	−0.35	329.32	7.70	6	0.26	3
	Social life	−0.08	0.08	−1.19	329.32	7.86	6	0.25	
	Independence	−0.77	0.08	−1.59	330.18	9.49	6	0.15	1, 2, 3
	Move around	−0.39	0.08	0.32	333.65	1.70	6	0.95	2
	Activities enjoy	0.36	0.08	−2.85	331.92	16.53	6	0.01	
	Physically active	0.60	0.08	−1.73	332.78	9.68	6	0.14	

| Scales	Item	Item Fit Statistics							DIF[a]
		Location	SE	Fit residual	DF	χ^2	DF	Probability	
Psychological									
	Hopeless	−0.83	0.09	−0.98	347.03	12.46	5	0.03	
	Desperate	−0.85	0.09	−2.29	345.24	12.53	5	0.03	
	Overwhelmed	−0.17	0.08	−3.69	344.35	17.66	5	0.00	3
	Sorry for self	−0.13	0.08	2.10	345.24	3.16	5	0.68	
	Depressed	−0.14	0.08	−0.84	345.24	3.70	5	0.59	
	Self-conscious	0.63	0.08	3.09	346.14	18.85	5	0.00	1, 2, 3
	Anxious	0.08	0.08	−2.34	347.03	6.91	5	0.23	
	Irritated	0.19	0.08	2.69	343.46	11.79	5	0.04	
	Frustrated	0.52	0.08	−0.27	345.24	0.94	5	0.97	
	Worried	0.70	0.08	−1.47	347.03	3.27	5	0.66	2, 3
Sleep									
	Falling asleep	−0.32	0.12	1.43	211.60	2.12	3	0.55	
	Enough sleep	−0.36	0.11	−2.72	211.60	6.86	3	0.08	
	Staying asleep	−0.31	0.11	−2.06	211.60	4.21	3	0.24	
	Position	0.79	0.10	2.40	211.60	6.32	3	0.10	
	Woken up	0.21	0.10	−1.02	211.60	2.43	3	0.49	
Social									
	Isolated	−1.00	0.09	1.66	242.42	5.89	4	0.21	1
	Meet people	−0.12	0.09	−1.71	240.05	9.41	4	0.05	
	Missed out	−0.06	0.09	−1.66	240.05	6.98	4	0.14	
	Cut down	0.49	0.09	−0.56	240.05	2.59	4	0.63	
	Enjoy life	0.69	0.10	1.55	242.42	8.45	4	0.08	
Dressing									
	Put on	−0.11	0.12	0.54	165.77	3.23	2	0.20	
	Look	−0.03	0.11	1.05	163.16	0.62	2	0.73	
	Absorb	−0.26	0.11	−0.12	163.16	2.60	2	0.27	
	Smell	−0.65	0.11	−0.90	164.90	1.18	2	0.55	
	Remove—easy	0.17	0.10	0.03	164.90	0.75	2	0.69	
	Change	0.25	0.11	−0.38	164.03	0.99	2	0.61	
	Comfortable	0.11	0.11	−0.18	164.90	0.99	2	0.61	
	Remove—felt	0.26	0.10	1.19	163.16	0.09	2	0.95	
	Active	0.26	0.11	0.03	164.03	0.13	2	0.94	

[a] 1 = Significant DIF after Bonferroni adjustment in first random sample, 2 = second random sample, 3 = third random sample)

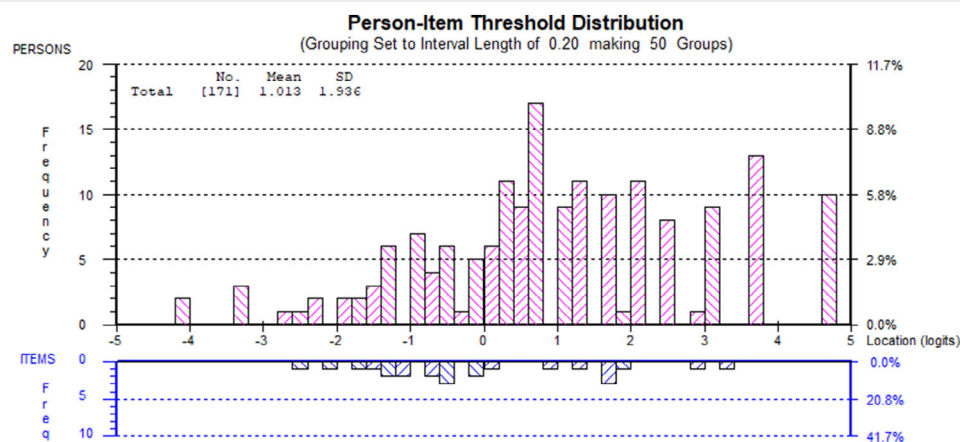** p-values significant after Bonferroni adjustment.

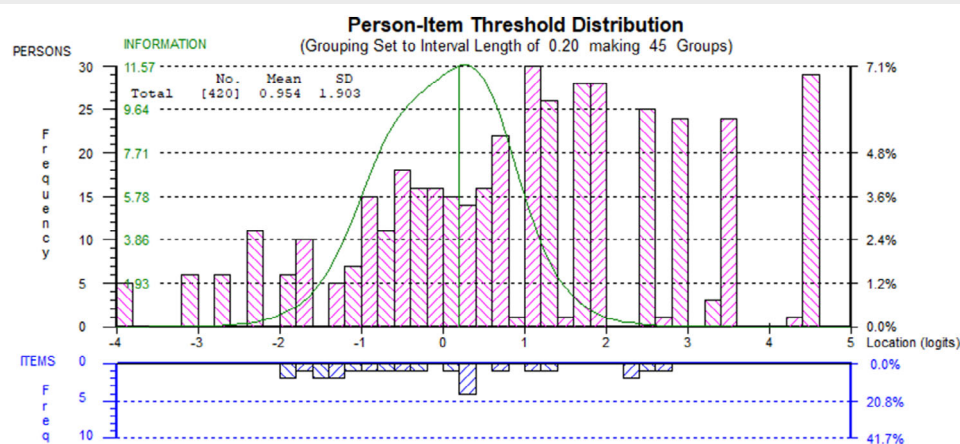**Item characteristic curve**: Item 1 in life impact scale.
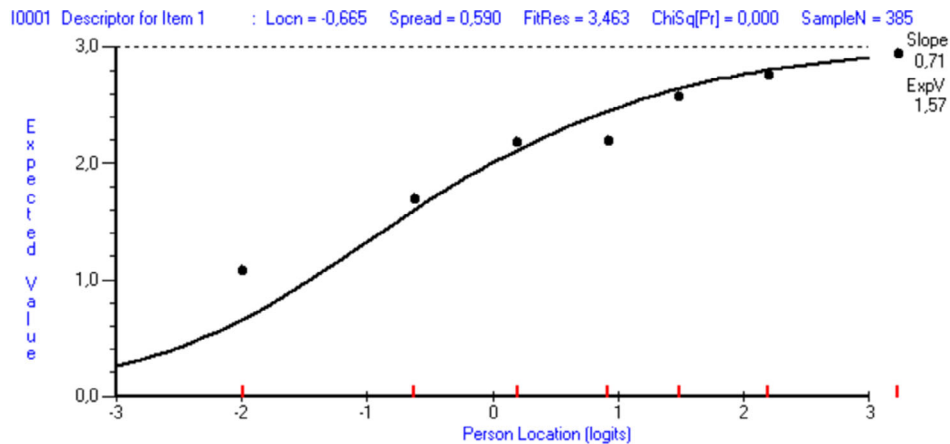
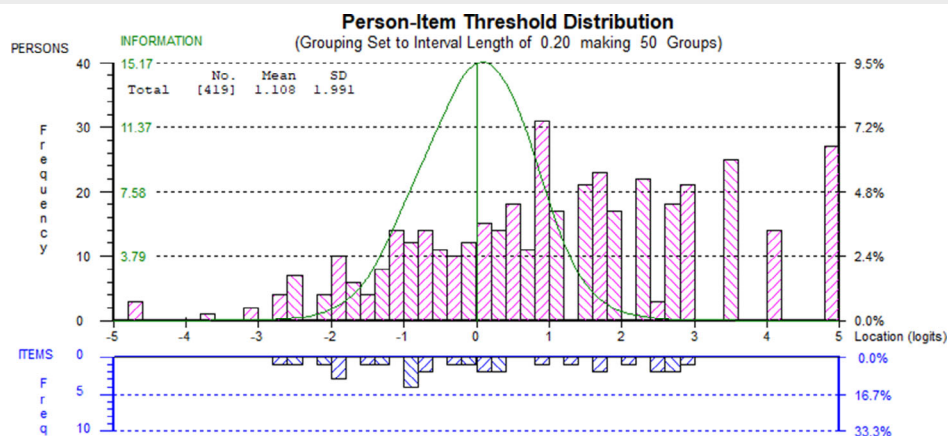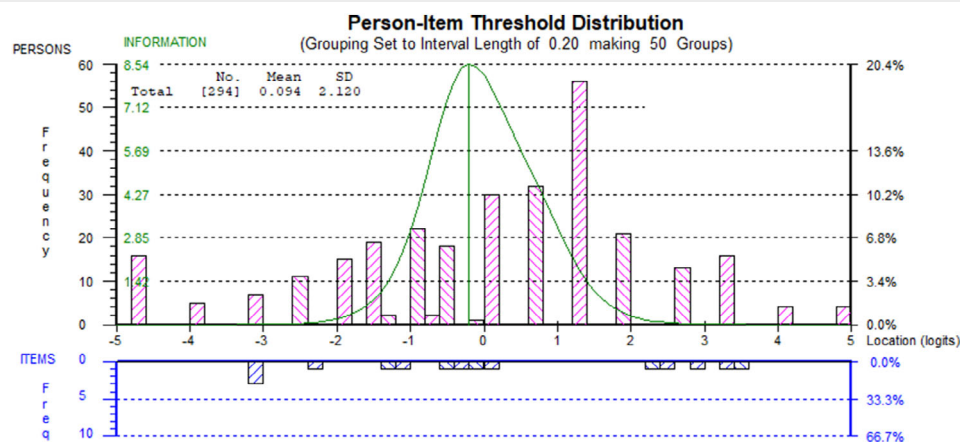## APPENDIX B: PERSON-ITEM THRESHOLD DISTRIBUTIONS FOR EACH WOUND-Q SCALE FROM RMT ANALYSIS

Assessment

Drainage
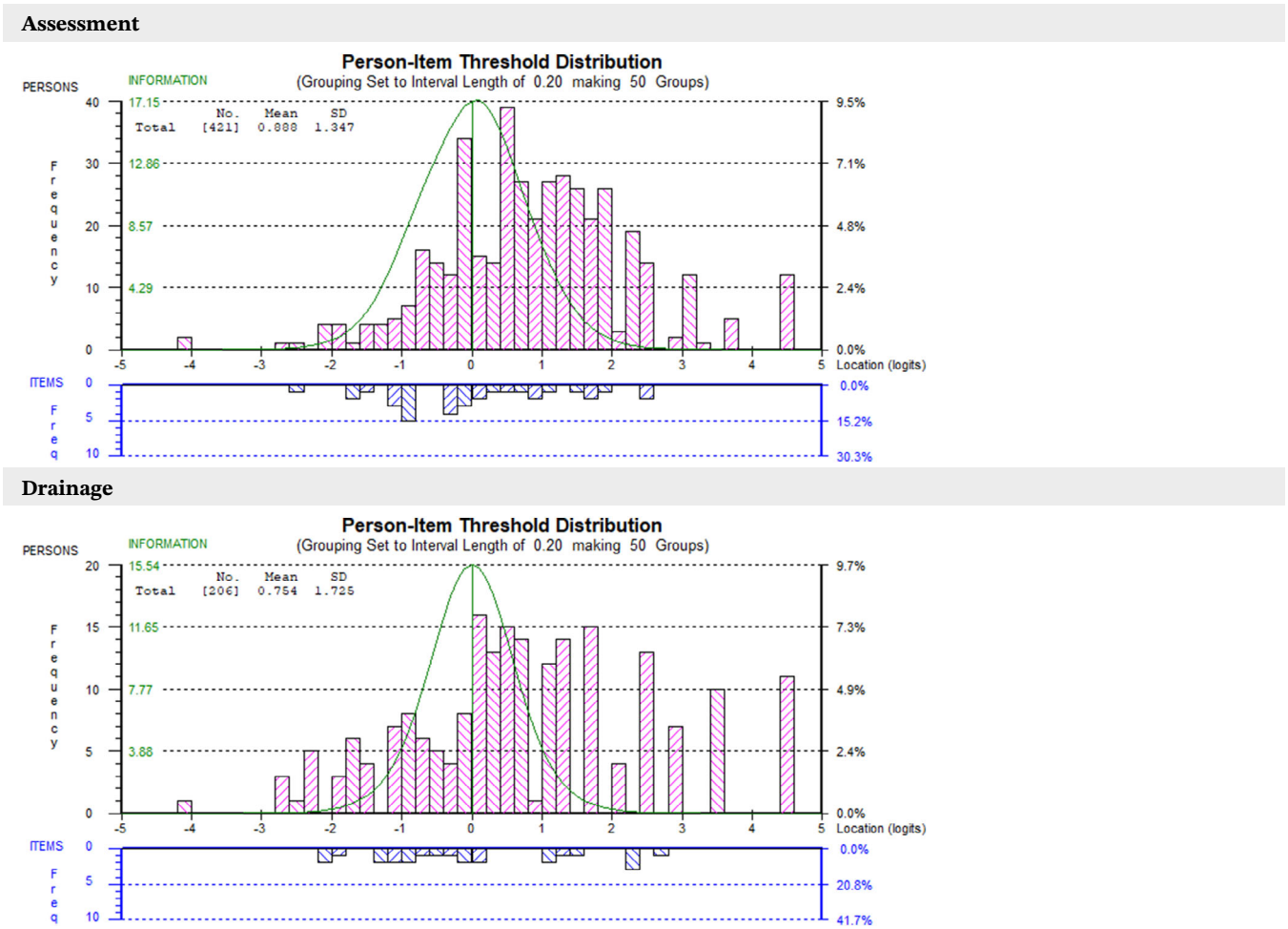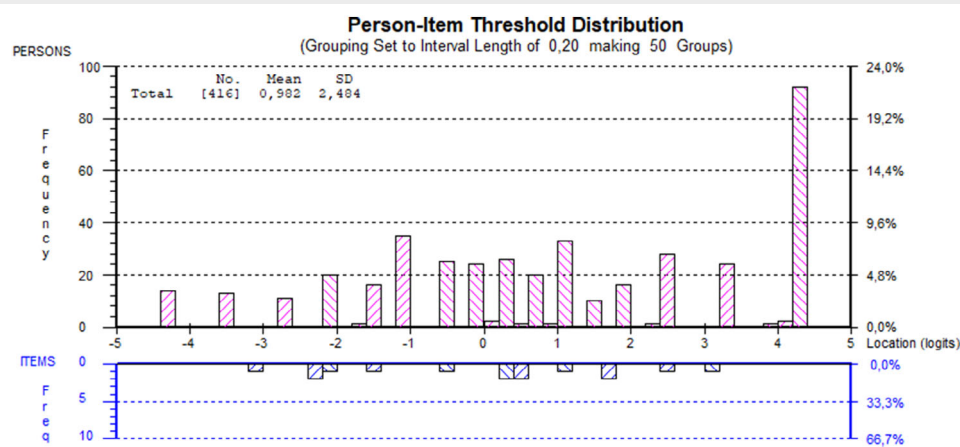
Smell

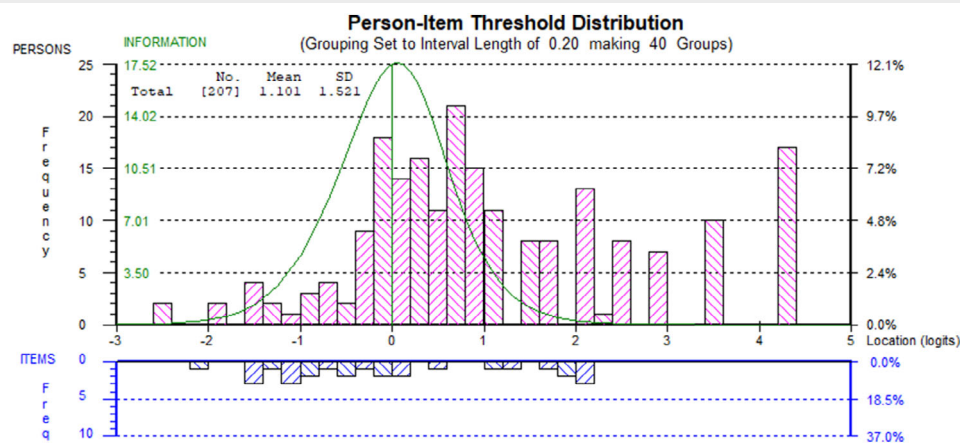

Life Impact

Psychological

(Continues)

Sleep

Social

Dressing

APPENDIX C: CHARACTERISTICS OF THE SUBGROUPS FOR HYPOTHESIS TESTING

	1 wound		>1 wound	
	n	*%*	*n*	*%*
Age (mean; SD)	38.16 years; 13.74 years		37.19 years; 13.26 years	
Gender				
Male	144	48.3	60	48.8
Female	152	51.0	60	48.8
Other	2	0.7	3	2.4
BMI (mean; SD)	27.38; 7.41		28.01; 13.26	
Country				
Australia	3	1.0	0	
Canada	10	3.4	3	2.4
Chile	1	0.3	1	0.8
Czech Republic	1	0.3	1	0.8
Denmark	0	1.0	1	0.8
Estonia	3	1.0	0	
France	3	0.7	0	
Germany	2		0	
Greece	6	2.0	0	
Hungary	4	1.3	2	1.6
Ireland	4	1.3	3	2.4
Italy	8	2.7	1	0.8
Latvia	0		1	0.8
Netherlands	1	0.3	1	0.8
New Zealand	3	1.0	0	
Poland	15	5.0	3	2.4
Portugal	0		2	1.6
Slovenia	16	5.4	1	0.8
South Africa	55	18.5	26	21.1
Spain	3	1.0	0	
UK	101	33.9	44	35.8
USA	57	19.1	32	26.0
Prefer not to answer	2	0.7	1	0.8
Ethnicity				
White	203	68.1	75	61.0
Black	57	19.1	30	24.4
Other	37	7.4	18	14.6
Missing	1	0.2	0	
Smoking/vaping				
Yes	89	29.9	43	35%
Marital status				
Never married	134	45.0	59	48.0

(Continues)

	1 wound		>1 wound	
	n	%	*n*	%
Separated	5	1.7	5	4.1
Divorced	15	5.0	7	5.7
Widowed	4	1.3	0	
Living common-law	28	9.4	14	11.4
Married	105	35.2	37	30.1
Other	5	1.7	0	
Prefer not to answer	2	0.7	1	0.8
Highest level of education				
Some high school	4	1.3	2	1.6
Completed high school	39	13.1	13	10.6
Some college or trade school or university	58	19.5	25	20.3
Completed college or trade school or university degree	125	41.9	54	43.9
Some Masters or Doctoral degree	22	7.4	13	10.6
Completed Masters or Doctoral degree	49	16.4	16	13
Prefer not to answer	1	0.3	0	
Work (able to choose multiple)				
Full time (≥37 h)	171	57.4	66	53.7
Part time (<37 h)	49	16.4	21	17.1
Unemployed	39	13.1	14	11.4
Student	21	7.0	9	7.3
Retired	14	4.7	3	2.4
On leave	1	0.3	1	0.8
Sick leave	12	4.0	11	8.9
Other	9	3.0	6	4.9
Comorbidities (able to choose multiple)				
Amputated	49	16.4	30	24.4
Diabetes	59	19.8	32	26.0
Peripheral Artery Disease (PAD)	16	5.4	6	4.9
Peripheral Venous Disease (PVD)	17	5.7	14	11.4
Paralysed	2	0.7	3	2.4
Hypertension	63	21.1	35	28.5
Ischaemic heart disease	13	4.4	9	7.3
Heart failure	7	2.3	1	0.8
Arthritis	30	10.1	21	17.1
Cancer	3	1.0	5	4.1
Chronic obstructive pulmonary disease (COPD)	5	1.7	9	7.3
Kidney failure	5	1.7	0	
Neuropathy	32	10.7	19	15.4
Stroke	4	1.3	4	3.3
Inflammatory bowel disease	7	2.3	4	3.3
Skin disease (inflammatory and autoimmune)	5	1.7	6	4.9
Multiple sclerosis	2	0.7	2	1.6

	1 wound		>1 wound	
	n	*%*	*n*	*%*
Psychological disorder	2	0.7	1	0.8
Other	45	15.1	22	17.9
Wound location (able to choose multiple)				
Face or neck	13	4.4	0	
Hand	7	2.3	1	0.8
Arm	27	9.1	7	5.7
Shoulder	8	2.7	2	1.6
Chest	10	3.4	4	3.3
Abdomen	27	9.1	9	7.3
Back	22	7.4	5	4.1
Buttocks	16	5.4	8	6.5
Genitals	7	2.3	3	2.4
Leg	75	25.2	29	23.6
Foot	58	19.5	29	23.6
Toe(s)	31	10.4	10	8.1
Other	5	1.7	3	2.4
Wound type				
Diabetic foot ulcer	27	9.1	10	8.1
Venous ulcer	6	2.0	2	1.6
Arterial ulcer	8	2.7	1	0.8
Pressure ulcer	12	4.0	5	4.1
Surgery	59	19.8	9	7.3
Radiation	1	0.3	2	1.6
Trauma/injury	106	35.6	23	18.7
Hidradenitis suppurativa	8	2.7	13	10.6
Pilonidal abscess	11	3.7	1	0.8
Multiple	11	3.7	42	34.1
Not sure	23	7.7	7	5.7
Other	26	8.7	8	6.5
Wound age				
3–6 months	160	53.7	55	44.7
7–12 months	29	9.7	18	14.6
1–2 years	42	14.1	24	19.5
3–4 years	27	9.1	12	9.8
5–10 years	32	10.7	9	7.3
11–30 years	6	2	4	3.3
>30 years	1	0.3	1	0.8
Prefer not to answer	1	0.3	0	
Wound size (width × length) (cm^2)				
Median, range	1.9 cm^2; 0.01–900 cm^2		3 cm^2; 0.03–1282.0 cm^2	
Wound size				
<1	99	33.2	26	21.1
1–2.4	51	17.1	24	19.5

(Continues)

	1 wound		>1 wound	
	n	**%**	**n**	**%**
2.5–4.9	40	13.4	13	10.6
5–9.9	30	10.1	11	8.9
10–24.9	21	7	20	16.3
>24.9	29	9.7	20	16.3
Missing	28	9.4	9	7.3

Note: 1 wound = 298 participants, >1 wound = 123 participants.

APPENDIX D: COMPARISON BETWEEN PROLIFIC AND THE FIELD-TEST SAMPLE

	Prolific (*n* = 421)		Field test (*n* = 881)	
	n	**%**	**n**	**%**
Age (mean; SD)	37.9 years; 13.6 years	Range 18–84	62.8 years; 14.5 years	Range 18–95
Gender				
Male	212	50.4	519	58.9
Female	204	48.5	357	40.5
Other	5	1.2	2	0.2
Missing	0		3	0.3
BMI				
Underweight	33	7.8	29	3.3
Normal weight	135	32.1	235	26.7
Overweight	120	28.5	258	29.3
Obese	128	30.4	328	37.2
Missing	5	1.2	32	3.6
Country	22 different countries, 3 main countries (34.4% UK, 21.1% USA, 19.2% South Africa)		4 different countries, 3 main countries (33.9% DK, 26.4% USA, 25% Netherland)	
Smoking/vaping (Yes)	132	31.4	114	12.9
Comorbidities	72.0% one or more comorbidities Most common hypertension (23.3%)		81.6% one or more comorbidities, most common DM (39.5%)	
Number of chronic wounds				
1	298	70.8	553	62.8
2	80	19.0	162	18.4
3	27	6.4	62	7.0
4	5	1.2	34	3.9
5+	11	2.5	59	6.7

	Prolific ($n = 421$)		Field test ($n = 881$)	
	n	%	*n*	%
Primary wound aetiology	Trauma	30.6	Diabetic foot ulcer	17.2
Wound age				
3–6 months	215	51.1	318	36.1
7–12 months	47	11.2	166	18.8
1–2 years	66	15.7	181	20.5
3–4 years	39	9.3	88	9.9
5–10 years	41	9.7	49	5.6
>10 years	12	12.1	36	4.1
Prefer not to answer/missing	1	0.2	43	5.0
Wound size (width $\times$ length) (cm^2)				
Median; range	2.04 cm^2	Range 0.01–1282.1 cm^2	4.0 cm^2	Range 0.0001–961.0
Wound size (cm^2)				
<1	125	29.7	206	23.4
1–2.4	75	17.8	146	16.6
2.5–4.9	53	12.6	98	11.1
5–9.9	41	9.7	94	10.7
10–24.9	41	9.7	136	15.4
>24.9	49	11.6	138	15.7
Missing	37	8.8	63	7.2
Wound symptoms (in the past week)				
Drainage	205	48.7	632	71.7
Smell	171	40.8	258	29.3
Sleep interference	294	69.8	445	50.5