

Rationale and Design of a Novel Method to Assess the Usability of Body-Worn Absorbent Incontinence Care Products by Caregivers

Sophie Kanerva Rice ◆ Leslie Pendrill ◆ Niclas Petersson ◆ Jesper Nordlinder ◆ Anne Farbrot

ABSTRACT

PURPOSE: The purpose of this study was to develop and test a new method to measure the usability of absorbent incontinence care products from the caregivers' perspective and to investigate if the method can be used to differentiate between product types in a product change.

DESIGN: Process evaluation and validation study.

SUBJECTS AND SETTING: Product developers and end users participated in designing the new method. Thereafter, professional caregivers acted as testers of the new method, ranking usability when performing absorbent product changes on patients in a simulated nursing home care environment, assisted by third-party research institute moderators.

METHODS: Design and evaluation of a new method designed to assess the usability of body-worn absorbent incontinence care products for lay caregivers were completed. The evaluation included formative and summative evaluations of effectiveness (product fit), efficiency (time and physical workload), and satisfaction. A person-centered approach aimed at including all subjects and settings to generate a single usability score for decision making and product benchmarking. Experienced caregivers changed 4 types of products: (1) disposable body-worn pads with mesh briefs (2-piece system); (2) disposable all-in-one briefs; (3) disposable, T-shaped, and belted brief; and (4) disposable pull-up pants on simulated patients in standing or lying position. Each product change was performed by 1 unassisted experienced caregiver. The probability of success as a score for each product type was calculated across the 4 metrics and reported with 95% confidence intervals (CIs). Descriptive and inferential statistics were developed assuming a binary statistical model, using the weighted scores from each of the factors. An overall usability score was calculated.

RESULTS: The method we developed discriminated between usability of different product types. The overall score for the disposable pull-up product (90%; CI: 83%-97%) was better ($P < .05$) than for the disposable T-shaped brief (83%; CI: 77%-89%), the disposable brief (53%; CI: 45%-61%), and the disposable body-worn pad with mesh pant (61%; CI: 56%-66%) in standing patients. For lying patients, the overall score for the disposable T-shaped brief product (81%; CI: 73%-89%) was better ($P < .05$) than the disposable brief (65%; CI: 45%-61%) and the disposable body-worn pads with mesh brief (62%; CI: 55%-69%). Reliability was evaluated quantitatively in terms of measurement uncertainties in the results.

CONCLUSION: The method we described demonstrated differentiation of usability based on product type indicating concurrent validity. Further testing in diverse real-world care environments is needed to evaluate and confirm the validity and to assess reliability of this method in the research setting.

KEY WORDS: Absorbent products, Benchmarking, Caregiving, Incontinence, Incontinence products, Nursing, Usability.

INTRODUCTION

Urinary incontinence is a globally recognized issue with a significant social and economic cost; prevalence estimates vary according to the definition of incontinence and the population studied.¹ There is a need to match patients and caregivers

with appropriate, effective absorbent incontinence products that they can change easily; this need is particularly acute for disabled patients in a home care environment. An objective method is needed to assess the ease of changing absorbent incontinence products in different care environments.

Sophie Kanerva Rice, MSc, RISE Research Institutes of Sweden AB, Gothenburg, Sweden.

Leslie Pendrill, PhD, RISE Research Institutes of Sweden AB, Gothenburg, Sweden.

Niclas Petersson, PhD, RISE Research Institutes of Sweden AB, Gothenburg, Sweden.

Jesper Nordlinder, MSc, SCA Hygiene Products AB (since June 2017, Essity AB), Gothenburg, Sweden.

Anne Farbrot, PhD, SCA Hygiene Products AB (since June 2017, Essity AB), Gothenburg, Sweden.

All authors confirm that they had full access to data and contributed to conception and design, acquisition, analysis, and interpretation of data, and drafting of the paper and of revising it critically. The article is

the authors' original work and is not under consideration for publication elsewhere.

Jesper Nordlinder is an employee of SCA Hygiene Products AB, and Anne Farbrot was a previous employee of SCA Hygiene Products AB. Sophie Kanerva Rice has worked as a consultant for SCA Hygiene Products AB and is a previous employee of SCA Hygiene Products AB. Leslie Pendrill and Niclas Petersson have no conflicts of interest to declare.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (JWOCNOnline.com).

Correspondence: Leslie Pendrill, PhD, RISE Research Institutes of Sweden AB, Ekländagatan 86, 41261 Gothenburg, Sweden (leslie.pendrill@ri.se).

DOI: 10.1097/WON.0000000000000462

Historically, evaluation of incontinence care products has focused on absorption capacity, with little emphasis on user-related and environmental factors. Recent guidelines have highlighted the needs of users and caregivers in evaluating absorbent incontinence care products.² However, quantitative methods for evaluating these products are lacking.^{2,3}

A human factor study in continence care described a novel system approach to define the comfort of incontinent patients. This approach outlined a holistic view of comfort as the sum of interactions between a product user, the incontinence product, and the task, compared with the incontinence standard ISO 15621.³ This standard describes a number of different constructs to consider when selecting the best absorbent product for individuals with urinary incontinence; they are divided into 3 categories: user, usage, and product. Several guidelines and standards describe the level of complexity required to identify the right product; they emphasize the importance of individual assessment, especially since evaluative factors will differ between product users and their caregivers.^{5,6} There are many absorbent incontinence products in the market, but it is not always clear which is the most suitable for a specific change situation. Therefore, a method is needed to match each carer and user with the most appropriate absorbent incontinence product type to facilitate the best care.

Usability is defined as the extent that a product can be used to achieve the user’s goals effectively and efficiently.⁷ In a more recent document, the ISO defined usability as the outcome of an interaction between a user and a product, service, or system.⁸ This concept of usability was introduced into medical device and user interface regulations^{9,10} and was promoted by the US Food and Drug Administration.¹¹ This study describes development and evaluation of a method to determine the comparative usability of different absorbing incontinence care product types. Our principal aims were to develop a valid and reliable method to measure usability of absorbent incontinence care product types, to compare product types, and to benchmark their suitability.

METHODS

Research and development of the method evaluated in this study were based on knowledge of documentary standards and product experience related to absorbent incontinence products. We then combined this knowledge with analogous methods used to measure usability in other product categories.

Measurement of usability includes interaction goals, elements of product use that influence usability, and target values for the chosen metrics of usability.⁷

In order to define interaction goals, we considered the type of product we wish to study and its intended use. Absorbent incontinence care products are designed to absorb and contain urine away from the skin. All of the products considered in this study were designed for heavy urinary incontinence. We focused on the interaction between an experienced caregiver and the product when product is changed; therefore, the outcome we measured is a successfully changed product. This usability benchmark study was limited to the task of product handling (removal and application) by caregiver participants in a home care environment. We evaluated changes by the caregiver with the patient in supine and standing postures. The tasks included taking off and putting on the product. Cleansing of the perineal area was not incorporated into this study. In order to identify context attributes essential to usability in absorbent incontinence products, we incorporated recommendations of various standards (Table 1).

Based on input from a cross-disciplinary group recruited by the product manufacturer (SCA Hygiene Products AB) and the testing institute (RISE Research Institutes of Sweden AB), we identified usability quality characteristics related to incontinence product changes in a nursing home (Table 2). We operationally defined usability based on 3 domains: (1) effectiveness: product fit after a change; (2) efficiency: the workload and time required to remove and put on the product; and (3) satisfaction: measured as a cumulative score on an instrument completed by the caregiver after changing the product.⁷

Our operational definition was guided by the standard definition of usability and, as proposed by ISO 9241-11:1998, metrics were chosen so that the data would reflect the result of interacting with the incontinence care product during a product change. The metrics and their methods of measurement were designed to develop a robust, streamlined, and reproducible method to benchmark product usability in incontinence care from the perspective of the caregiver. Inspired by a preexisting Single Usability Metric (SUM) method, the cumulative score we developed combines usability metrics that reflect the ease of interaction between the caregiver and test product in a test environment rather than an overall product score applicable to all contexts.¹²⁻¹⁴ Context attributes to be considered together with the selected attributes of the study are described

TABLE 1. Context Attributes Important to Usability in Incontinence Care Products—Based on ISO 15621:2011³ and ISO/IEC 62366:2015¹¹

End User	Caregiver	Equipment	Task	Environment
Intellectual ability	Intellectual ability	Product identification	Task procedure	Job function
Attitude	Attitude	Product description	Frequency of use	Interruptions
Motivation	Motivation	eg, Procedure of application	Physical and mental demands	Stress
Physical limitations	Physical limitations	Application areas	Output	Technical environment
Physical capabilities	Physical capabilities	Major functions	Risk from error	Workplace conditions
Body shape	Gender	Services	Safety demands	Atmospheric, auditory, thermal, visual conditions
Gender	Skills/experience			Space and furniture
Nature of incontinence	Training			Location
Skills/experience				Health hazards
Activities				

TABLE 2.
Selected Usability Quality Characteristics

	Usability Quality Characteristics ⁷	Metrics Chosen	Incontinence Standard ISO 15621:2011
Overall usability	Effectiveness	Product fit: Estimate of success rate of achieving a good enough fit according to predefined parameters	"The correct fixation and product fit to the body are very important and influence the leakage properties of any product." (§5.2 Product-related factors: Freedom from leakage)
	Efficiency	Time on task: Estimate of success rate of finishing task within a specified time limit	"The ease with which a urine-absorbing aid can be put on or taken off is important to all end-users." (§4.8 User-related factors: Handling product)
		Workload: Estimate of success rate of finishing task within a workload limit	"When helping a person with incontinence with their personal hygiene and change of incontinence products, the ergonomics has to be considered." (§6.1 Usage-related factors: Ergonomics)
	Satisfaction	User experience: Estimate of success rate of scoring product higher than a specification limit	"If a carer is required to apply or change the product, then it may be important to involve him or her in the selection of the product and to establish his or her willingness and ability to use it." (§6.2 Usage-related factors: Needs of carer)

in Supplemental Digital Content Tables S1-S5 and Figure S1 (available at: <http://links.lww.com/JWOCN/A44>).

Weighting (w) reflecting the relative importance of each product factor was determined by questioning a group of 115 experienced caregivers. The metrics were rated independently along a 9-point Likert scale, from "equal" (scored as 1) to "extreme" (scored as 9). As anticipated, the distribution of ratings for each metric was positively skewed because all of the metrics being considered were perceived as important for caregivers. The weights were then generated to match the relative importance using a Rasch analysis.¹⁵ Finally, target values were determined for each of the chosen usability metrics.

Method of Measurement

The method for assessment of usability of body-worn absorbent incontinence care products for lay caregivers was tested in a laboratory environment simulating a nursing home. Professional caregivers acting as "testers" changed a simulated patient's absorbent product and ranked usability based on the following criteria: effectiveness (product fit), efficiency (time and physical workload), and satisfaction, as described later.

Each tester was assisted in test performance and data evaluation by an evaluator and a moderator from the testing institute. We also employed simulated patients who played a more passive role; specifically, they were instructed not to assist the caregiver during the changing procedure.

Effectiveness

Effectiveness was assessed as product fit based on visual inspection by a tester after each product change; the inspection was guided by a structured protocol. Specifically, the tester assessed leakage security (Was the product applied with absorbing material facing the skin and backing away from skin?) and securement (Did the caregiver use the intended fastening system and was the product deemed snug as measured at the waist using a ruler?). Finally, coverage was measured by determining whether the tester visually inspected and measured the distance between the anterior superior iliac spine and the core, followed by inspection of the absorbing core height at the back. Coverage on the back was deemed incorrect when the absorbing core did not cover the groove between the buttocks. The tester also assessed contact at the crotch by determining whether the elasticized leg bands were

in contact with the crotch. Each product was photographed prior to visual inspection. The detailed procedures and product fit scoring and protocols are described in detail in Supplemental Digital Content Figures S2-S5 (available at: <http://links.lww.com/JWOCN/A44>). Photographs were stored for future reference of the fit of the different products as an aspect of effectiveness. For the evaluation, the photographs were only used as support for reviewing the scoring of use of fastening on comfort.

Variations of the elements were measured through the visual inspection sheet graded according to a color system with assigned penalties. Dark green indicated the intended fit, light green indicated a slight deviation with a penalty of 0.25, yellow indicated a larger deviation with a penalty of 0.5, orange indicated a larger deviation with a penalty of 0.75, and red indicated critical failure.

Individual product fit received a binary score of 1 (indicating success) when a change resulted in a fit that did not trigger a critical failure or multiple penalty failure with a cumulative score of 0.75 or more. Otherwise, the product fit was scored as a 0 (failure); critical failure of product occurred when 1 aspect of the fit rendered the product ineffective. Examples of such critical failures were as follows: (1) product being applied inside out, (2) intended fastening system not engaged, and (3) absorbent core too low at the front (see Supplemental Digital Content Figures S2-S5, available at: <http://links.lww.com/JWOCN/A44>, for further details).

Efficiency

Efficiency was measured by the time required to complete product change and the task workload (see Supplemental Digital Content Tables S6-S10, available at: <http://links.lww.com/JWOCN/A44>). Time (recorded in seconds) was measured from the moment when the caregiver first received the product until the task was completed. Time spent on actions other than person/product handling, such as putting on or taking off gloves, talking to the moderator, and adjusting the bed, was subtracted from the overall time. No maximum time limit was specified. For purpose of scoring, time on task was scored as a binary measure, defined by limits of a standing change taking more than 90 seconds or a lying change taking more than 120 seconds to distinguish incontinence products. These choices of cut point for time on task were based mostly on end user requirements and values in the literature.¹³

Task workload was defined as a function of time and load of the inherent postures for the caregiver in a predefined set of actions and the weight factor resulting on the spine. For standing changes, back flexions were timed. For lying changes, 90° turnovers, lifts, repositioning, and the action of reaching over (defined as the caregiver reaching over to work on the other side of the patient) were timed (see Supplemental Digital Content Tables S7–S10, available at: <http://links.lww.com/JWOCN/A44>). A total time for each action was calculated.

Each action type was associated with a load on the spine based on the postures that the legs, arms, and trunk are in and the weight the person has in his or her hands.¹⁶ For each action, we set up a typical posture. For example, we set that during a turnover, the trunk is moderately flexed, the arms are over 60°, and the weight in hand is more than 22.68 kg.

We also determined a “physical workload index” of the time spent on the different actions multiplied by their weighting factor, accounting for compressive forces on the caregiver’s spine associated with the action.¹⁶ For physical workload index scores exceeding 100, the change was considered a failure and was attributed a score of 0. Otherwise, the change was considered a success and was given a score of 1.

Satisfaction

The caregiver’s satisfaction was measured by a self-reported questionnaire; it was also scored as a binary outcome, where 1 indicated satisfied and 0 indicated dissatisfied (see Supplemental Digital Content Table S11, available at: <http://links.lww.com/JWOCN/A44>). The questionnaire was based on the Post-Study System Usability Questionnaire¹⁷ and adapted to changing incontinence care products. Our aim was to differentiate the products, assuming they would all meet a lower standard. Statements were positively phrased, comparing the present product with each carer’s past experience, and responses were rated according to the 5-point Likert scale from “disagree completely” (score as 1) to “agree completely” (scored as 5).¹⁸

Cumulative Score

A cumulative score was calculated based on weighted averages from each of the domains. The weights were as follows: $w_1 = 0.300$, fit; $w_2 = 0.215$, time; $w_3 = 0.255$, workload; $w_4 = 0.230$, satisfaction. A higher cumulative score indicates a higher usability; no cut points have yet been defined.

Study Procedures

The study was conducted in facilities provided by Scandinfo Marketing Research AB, based in Malmö and Gothenburg, Sweden. Study rooms were configured to resemble an end user’s bedroom. An adjustable bed, including sheets and pillow,

was placed with the headboard against a wall and positioned to allow access from both sides. A bedside table was placed on the right-hand side of the bed. Gloves were placed on the bedside table, with a rubbish bin to dispose of used products. Other materials included vinyl gloves (all sizes); incontinence care products; and a simulated patient with black protective wear (over which the products were placed): tight, covering, long-armed top and long-legged trousers.

In order to create a test setting as close to reality as possible, simulated patients were recruited to participate in the study. The simulated patients were adult females who fit the medium size of the incontinence products in the study (hip measurement 85-97 cm). The simulated patients were instructed to be cooperative; they were permitted to stand steadily with support and could lift their legs in a standing position if the lift was initiated by the caregiver, could be rolled over if the caregiver initiated the process, but were not permitted to assist in any other way.

Two cameras were installed to monitor the task, and a moderator was present to provide information on the task and to hand out the post-task questionnaire. An evaluator in an adjacent room viewed the proceedings on a monitor to determine if the change was performed according to the product’s intended use. The evaluator entered the room after each change and rated the product change. A table and a chair were set up behind a dividing panel for the caregiver to complete the satisfaction questionnaire. Four incontinence care product types were investigated (Table 3, see also Supplemental Digital Content Table S4, available at: <http://links.lww.com/JWOCN/A44>). For each change of the disposable body-worn pads with mesh brief, both the pad and the fixation underwear were changed. No other manufacturer’s products were tested.

Study Participants

The testers were practicing Swedish caregivers currently working as professional caregivers in a nursing home environment for 12 months or more. Their regular duties included regular changes of at least 2 of the incontinence products in the tested change situations, that is, changing on their own on a patient in standing and lying positions (see Supplemental Digital Content Table S2, available at: <http://links.lww.com/JWOCN/A44>). Each caregiver changed 2 different products on patients in standing and lying positions, except the protective underwear, which was only changed in the standing position. Participation required understanding Swedish written and oral as well as willingness to be filmed and sign consent and confidentiality forms. Individual consent was collected from all participants. Exclusion criteria were not putting on product correctly under observation and no experience with a pertinent product or changing the product when patients were in a particular position.

TABLE 3. Incontinence Care Product Studied	
Single-use pads for urine and feces , also called and herein referred to as disposable body-worn pads with mesh brief	Absorbent incontinence product for urine and feces that is held in place by fixation underwear
Single-use absorbent products for urine and feces with built-in fastener system , also called AiO or adult brief, and herein referred to as disposable brief	Absorbent incontinence product in which the absorbent core is mounted within a chassis, equipped with readjustable fastener system
Single-use protective or disposable underwear , also called protective underwear or pants, and herein referred to as disposable pull-up product	Absorbent incontinence products shaped and designed to resemble normal underwear designed especially for male or female users or as unisex products
Single-use absorbent products with belted fastener system , also called and herein referred to as disposable T-shaped brief	Absorbent incontinence product in which the absorbent core is mounted within a chassis and equipped with readjustable waist belt that is first fastened around the user’s waist before the front part of the chassis is fixed on the belt

The different product and position combinations involved testing (changing) by between 25 and 34 caregivers. The range in the number of caregivers arose because only changes performed in line with the intended use were included in the final analysis, while aiming for the most economical, balanced, and representative study design. Each test, in which each tester performed task (product changes), included 2 products and 2 patient postures (standing/lying). All products except those that were pulled up were changed standing and lying. Pull-up products were not tested in the lying position since they are not usually used or recommended for that posture. There were a total of 24 different orders of balanced combinations of product and patient posture. Product changes were 28 for standing position and 28 for lying position for disposable brief; 34 for position standing and 32 for lying position for the body-worn pads + mesh briefs; 29 for standing position and 31 for lying position for the disposable T-shaped brief; and 25 for standing position for products that were pulled up. Caregivers were asked to watch an instruction video of the intended change procedures; 43 of 61 (70%) watched the video.

Caregivers and simulated patients were filmed during the 3 to 4 product changes per test session, depending on the product. The completed change was photographed from 4 angles (front, back, right, and left) and marked with the participant number. The caregivers then completed a satisfaction questionnaire. Product fit, time on task, task workload, and satisfaction were scored separately. All caregivers were scored according to the same criteria (except product-specific criteria that did not apply to all products).

Data Analysis

The probability of success for each product type was calculated across the 4 metrics of usability and reported with 95% confidence intervals (CIs). Descriptive and inferential statistics were developed by a binary statistical model, using the scores from each of the 4 factors. An overall score similar to SUM but adapted to the metrics collected herein was calculated using the agreed weightings.¹²⁻¹⁴

Based on the usability metrics defined, every caregiver received a binary score (0 = fail; 1 = success) on each separate factor, that is, for all 4 metrics (product fit, task time, task workload, and caregiver satisfaction). The choice of binary results for all metrics allowed for the same type of statistical analysis for the “usability score,” regardless of metric, to be applied. The probability of succeeding was denoted P_{success} and was estimated together with a 95% CI based on the result of the test of each metric.

Additionally, all metrics had a complementary analysis to the usability score, which focused on the aspects of each specific metric. A metric such as *Time on task* was analyzed with respect to the distribution of different times and not only the (crude) score of success versus failure. For scoring satisfaction, success was achieved if and only if the “Rasch parameter” (from a Rasch psychometric invariant measurement analysis of the satisfaction questionnaire¹⁵) of the test caregiver was greater or equal to the lower specification limit. The chosen specification limits were $L_{\text{SL,Standing}} = 1.0$ and $L_{\text{SL,Lying}} = 1.0$, respectively, which correspond to a success probability of about 74% for an average user.

Estimation of Metric-Specific Scores

We assumed that a binary response was achieved with probability P_{success} for a randomly selected caregiver, as described in the previous section. Furthermore, it was assumed that all care-

givers succeeded independently of each other and that there was no “memory” between tests of different product types. As economical, balanced, and representative study design as possible was aimed at, as described in the section “Study Participants.” Any uncompensated correlation would add to the measurement uncertainties.

As evident from Figure 1, caregivers had some previous experience of using the tested product types. This was assumed not to bias the current evaluations but indeed was used in a positive sense in the satisfaction questionnaire, where the caregiver was asked to rank in terms of “my experience of changing pads on a care recipient...” as a means of enhancing sensitivity.

Given a test population of N_{TP} test caregivers, the total number of successes X was then a binomial distributed random variable:

$$X \in \text{Bin}(N_{\text{TP}}, P_{\text{success}}).$$

The target was to estimate P_{success} together with a suitable CI. There was no simple way to construct an exact CI $I_{P_{\text{success}}}$, α for P_{success} with CI $1-\alpha$. Instead, we aimed for an approximate CI via a suitable normal approximation of the binomial distribution. Instead of using the most common point estimate of P_{success} , that is, the relative frequency:

$$\hat{p} = d / N_{\text{TP}},$$

where d is the number of successes, we used the Wilson point estimate:

$$\tilde{p} = \frac{d + \frac{1}{2} z_{\alpha/2}^2}{N_{\text{TP}} + z_{\alpha/2}^2},$$

where $z_{\alpha/2}$ is the $\alpha/2$ quantile of the $N(0,1)$ distribution. The estimated standard deviation was given by:

$$\tilde{\sigma} = \frac{1}{N_{\text{TP}} + z_{\alpha/2}^2} \sqrt{N_{\text{TP}} \tilde{p}(1-\tilde{p}) + \frac{1}{4} z_{\alpha/2}^2}$$

and the corresponding Wilson CI (which has \tilde{p} as midpoint) was written as:

$$\tilde{I}_{P_{\text{success}}, \alpha} = \frac{1}{N_{\text{TP}} + z_{\alpha/2}^2} \left[\left(N_{\text{TP}} \tilde{p} + \frac{1}{2} z_{\alpha/2}^2 \right) \pm z_{\alpha/2} \sqrt{N_{\text{TP}} \tilde{p}(1-\tilde{p}) + \frac{1}{4} z_{\alpha/2}^2} \right].$$

We thus rejected the standard normal approximation for an exact CI in favor of the Wilson CI (also called the Wilson score CI). Note that the CI for P_{success} was only expressed in terms of \hat{p} (and not \tilde{p}) to avoid confusion in calculations.

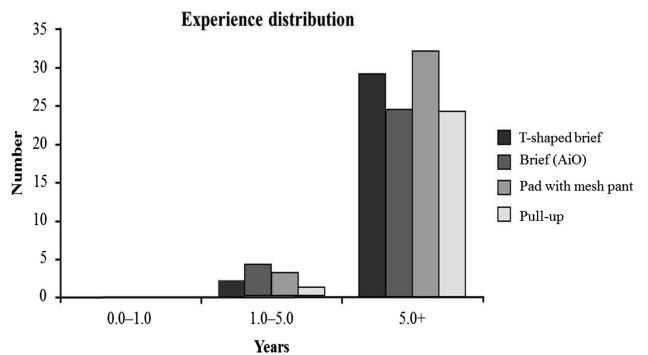


Figure 1. Experience distribution of test participants for the 4 product categories.

The advantage of the Wilson CI compared with the “standard normal approximation CI” is that it provides better coverage and is better suited to handle completion rates approaching 0 or 1. Data for each combination of product and patient posture were analyzed separately.

Estimation of Usability Score

The usability score for a product $P_{usability}$ was a weighted measure of the Wilson point estimates for the 4 metrics, with the weights determined by caregiver input as previously described, yielding:

$$\tilde{p}_{usability} = 0.300\tilde{p}_{fit} + 0.215\tilde{p}_{time} + 0.255\tilde{p}_{workload} + 0.230\tilde{p}_{satisfaction}$$

We used \tilde{p}_{fit} , \tilde{p}_{time} , etc, when we wanted to specify the metric of the estimated $P_{success}$. The 4 metrics were assumed to be independent, and the standard deviation was estimated as:

$$\tilde{\sigma}_{usability} = \sqrt{(0.300\tilde{\sigma}_{fit})^2 + (0.215\tilde{\sigma}_{time})^2 + (0.255\tilde{\sigma}_{workload})^2 + (0.230\tilde{\sigma}_{satisfaction})^2}$$

Discussions of possible correlations among the various metrics can be found in the literature,¹² since these need to be considered when attempting to combine the metrics into a single usability score.

The approximate CI for $P_{usability}$ was given by:

$$\tilde{I}_{p_{usability}, \alpha} = \tilde{p}_{usability} \pm z_{\alpha/2} \tilde{\sigma}_{usability}$$

To compare different products, we chose in this first study of the new method to make pairwise hypothesis tests of the scores; for example, comparing disposable T-shaped brief (TENA Flex) Standing [FS] with (body-worn pads + mesh briefs [TENA Comfort/TENA Fixation Pant] Standing [CS]) with respect to workload, as opposed to extended group tests. The null hypothesis was:

$$H_0 : P_{workload}^{FS} = P_{workload}^{CS}$$

The estimate of the common value $P_{work}^{FS} = P_{work}^{CS}$ was given by:

$$\tilde{p}_{workload}^{FS,CS} = \frac{\tilde{p}_{workload}^{FS} N_{TP,workload}^{FS} + \tilde{p}_{workload}^{CS} N_{TP,workload}^{CS}}{N_{TP,workload}^{FS} + N_{TP,workload}^{CS}}$$

using the Wilson point estimates. The estimated standard deviation was given by:

$$\tilde{\sigma}_{workload}^{FS,CS} = \sqrt{\tilde{p}_{workload}^{FS,CS} (1 - \tilde{p}_{workload}^{FS,CS}) \left(\frac{1}{N_{TP,workload}^{FS}} + \frac{1}{N_{TP,workload}^{CS}} \right)}$$

Now, assuming that $\tilde{p}_{workload}^{FS} > \tilde{p}_{workload}^{CS}$, we then wanted to formulate the alternative hypothesis:

$$H_1 : P_{workload}^{FS} > P_{workload}^{CS}$$

(one-sided—as recommended in ISO/TS 20282-2¹⁹).

The null hypothesis was rejected if:

$$\frac{\tilde{p}_{workload}^{FS} - \tilde{p}_{workload}^{CS}}{\tilde{\sigma}_{workload}^{FS,CS}} \geq 1.6449$$

with significance level $\alpha = .05$. This was the standard approximate hypothesis test for comparing results from 2 binomial distributions¹⁸ but with Wilson point estimates (instead of frequency point estimates).

Validity and Reliability

In line with a patient-centric approach, our study comprised a quality-assured treatment of data around caregivers’ interactions with incontinence products.⁴ The sample size was based on current recommendations for comparative tests (at least 30) deemed sufficient to compare the different product types regarding the factors identified and to produce a total usability score for each product.²⁰ Reliability of the method was indicated quantitatively in terms of measurement uncertainties in the results. For the purposes of this study, we measured reliability based on reproducibility of findings if measurements were repeated (test-retest reliability) promulgated in §5.4 Validity and reliability, ISO/TS 20282.¹⁹ Concerning validity, as this is, to our knowledge, the first study of its kind, the validity of the results in different environments is yet to be tested.

RESULTS

The method we developed was able to discriminate between the usability of different product types (Figure 2). The overall usability score for the disposable pull-up product (90%; CI: 83%-97%) was better ($P < .05$) than for the disposable T-shaped brief (83%; CI: 77%-89%), the disposable brief (53%; CI: 45%-61%), and the pad with body-worn pads + mesh briefs (61%; CI: 56%-66%) in standing patients. The overall score for the disposable T-shaped brief was better ($P < .05$) than the disposable brief and the body-worn pads + mesh briefs. In lying patients, the disposable T-shaped brief scored better ($P < .05$) (score 81%; CI: 73%-89%) than the disposable brief (65%; CI: 59%-71%) and the body-worn pads + mesh briefs (62%; CI : 55%-69%). No other comparisons were statistically significant (Figure 2).

Comparison of Metric-Specific Scores

In the standing position, the disposable pull-up product showed a better product fit score (93%; CI: 86%-100%) than the disposable T-shaped brief (79%; CI: 73%-85%), the disposable brief (72%; CI: 64%-80%), and the body-worn pads + mesh briefs (79%; CI: 74%-84%) ($P < .05$) product types. No other comparisons were statistically significant in the standing or lying position (Figure 3). In addition, the disposable T-shaped brief (82%; CI: 76%-88%) and the disposable pull-up product (90%; CI: 83%-97%) product types showed a better time on task score than the disposable brief (34%; CI: 26%-42%) and the body-worn pads + mesh briefs (58%; CI: 53%-63%) ($P < .05$). The body-worn pads + mesh briefs showed a better time on task score than the disposable brief. In the lying position, the disposable T-shaped brief (74%; CI: 66%-82%) scored better than the body-worn pads + mesh briefs (53%; CI: 46%-60%) ($P < .05$). No other comparisons were statistically significant (Figure 4).

In the standing position, the disposable T-shaped brief (91%; CI: 85%-97%) and the disposable pull-up product (90%; CI: 83%-97%) product showed a better workload score than the disposable brief (50%; CI: 42%-58%) and the body-worn pads + mesh briefs (45%; CI: 40%-50%) ($P < .05$). In the lying position, the disposable T-shaped brief (86%; CI: 78%-94%) scored better than the disposable brief (59%; CI: 53%-65%) and the body-worn pads + mesh briefs (53%; CI: 46%-60%) ($P < .05$). No other comparisons were statistically significant (Figure 5).

In the standing position, the disposable T-shaped brief (82%; CI: 76%-88%) and the disposable pull-up product (86%; CI:

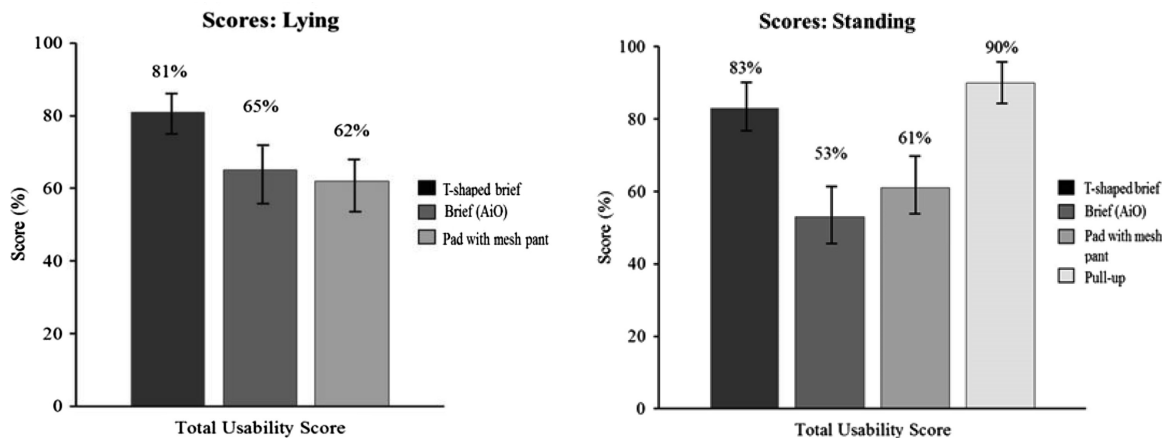


Figure 2. Total usability scores with 95% confidence interval.

79%-93%) showed a better caregiver satisfaction score than the disposable brief (50%; CI: 42%-58%) and the body-worn pads + mesh briefs (58%; CI: 53%-63%) ($P < .05$). In the lying position, the disposable T-shaped brief (83%; CI: 75%-91%) scored better than the disposable brief (66%; CI: 60%-72%) and the body-worn pads + mesh briefs (47%; CI: 40%-54%) ($P < .05$) product types, while the disposable brief scored better than the body-worn pads + mesh briefs. No other comparisons were statistically significant (Figure 6).

DISCUSSION

This study described the development of a method designed to measure the usability of different incontinence absorbent care product categories. Our method enables a comparison of the usability of changes of the tested incontinence care products performed by caregivers. The different product attributes can be measured, including product fit, handling effort (time and workload), and satisfaction. We recommend combining this method with existing instruments designed to identify the

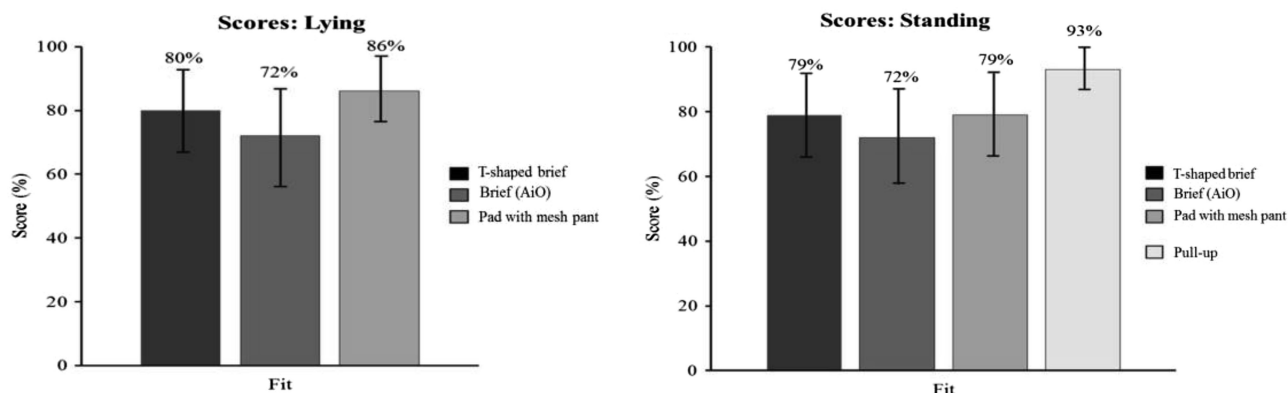


Figure 3. Product fit scores with 95% confidence interval.

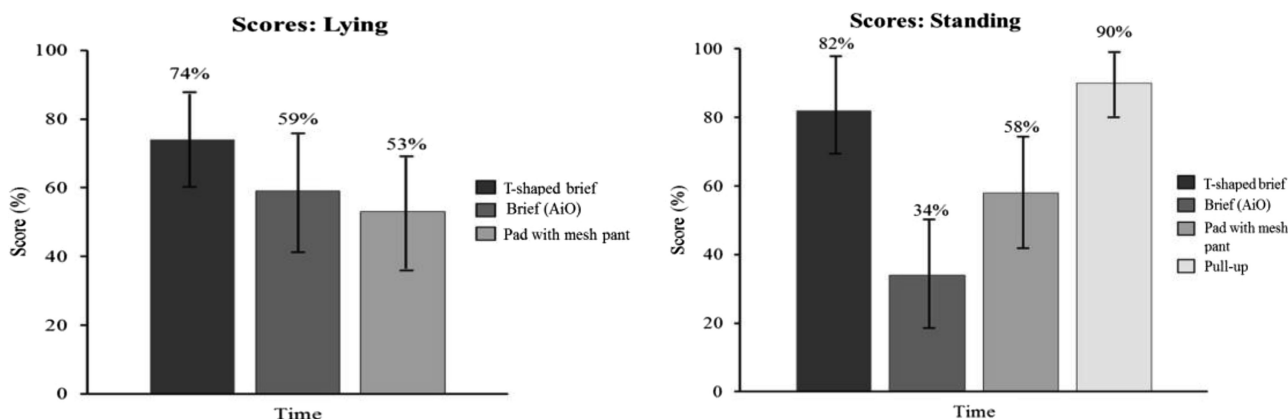


Figure 4. Time on task scores with 95% confidence interval.

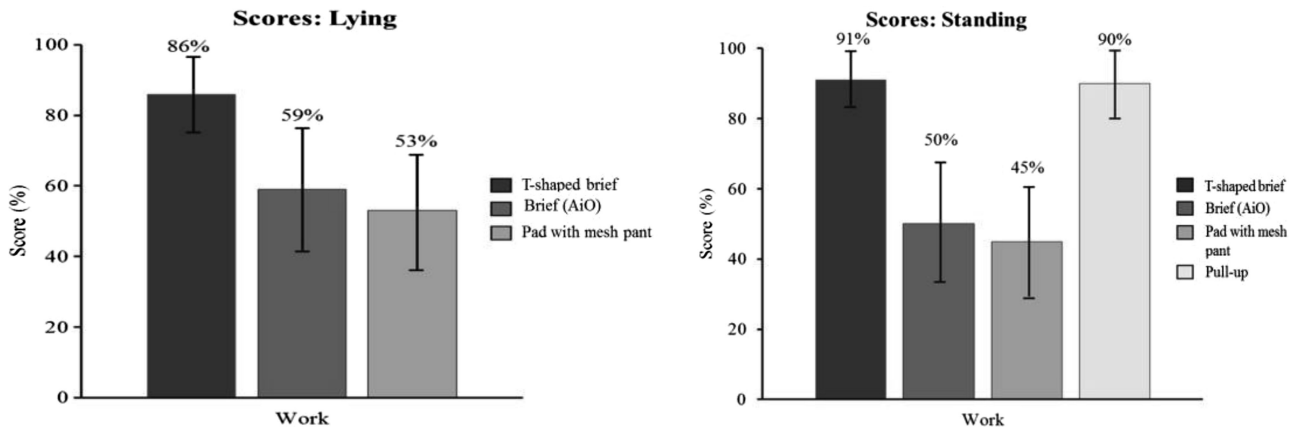


Figure 5. Workload on task scores with 95% confidence interval.

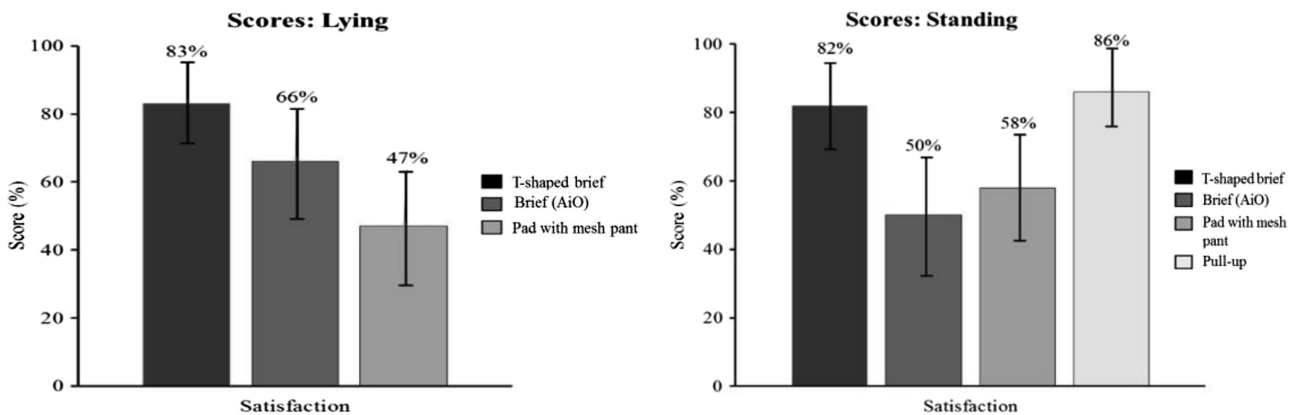


Figure 6. Satisfaction scores with 95% confidence interval.

most appropriate products for caregivers and users alike. The study encompasses a quality-assured analysis of caregivers’ interactions with incontinence products.

To our knowledge, this is the first application of a Rasch psychometric invariant measurement analysis of usability measures.¹⁵ Tezza and colleagues²⁰ previously reported a Rasch-type analysis of Web usability. The Rasch approach²² has been proposed as a viable alternative to methods described in ISO/TS 20282-2:2013, stating that although satisfaction questionnaires produce ordinal data, parametric statistics are more meaningful when analyzing satisfaction questionnaires.¹⁹

Research suggests that patients have higher health-related quality of life if they are informed and supported in their choice of product.²³ Usability is an important concept in helping caregivers select the best product for their patients. The adoption of usability as a factor in selecting incontinence care products is expected, considering the increasing use of usability in other industries and the need for usable products in the home care environment.

The choice of a binary result for each metric allowed similar statistical analyses to be performed. Reliability was indicated quantitatively in terms of measurement uncertainties in the results obtained with the method (Figures 2-6). Nevertheless, this study is the first using the novel method we developed, and it must be validated in different care environments before it can be adopted for widespread use in the research or clinical setting.

Limitations

Additional testing is needed to demonstrate validity and reliability of the method we developed. We collected data in a simulated, general care environment, using experienced caregivers familiar with the different product types. While this test environment provided suitable proof of concept, additional testing is needed to fully understand how robust and useful the method is in various real-world care environments. A number of assumptions of independence, such as between satisfaction and product fit, will need to be investigated in future work determining the degree of correlation between the various metrics. Extended group tests in the future may reveal limitations, such as the effects of multiple testing not so far revealed in the pairwise product comparisons currently made in this work.

CONCLUSIONS

We developed a method to assess the usability of body-worn absorbent incontinence care in a simulated long-term care environment. The method has similarities to ISO/TS 20282-2:2013, which was developed in parallel with our work. Findings indicate that our method can differentiate between product types and combines the results into a single overall usability score.

ACKNOWLEDGMENTS

The authors thank Ragne Emardson, Petra Sommarlund, and Anders Trana at RISE Research Institutes of Sweden AB, for their

assistance, and Scandinfo Marketing Research AB, for provision of the trial facilities, recruitment of participants, and performing of tests. The authors also thank Cilla Persson for the illustrations in the fit protocol. Editorial support and writing assistance were provided by Jaya Shumoogam of Ketchum, funded by SCA Hygiene Products AB. The project was funded by SCA Hygiene Products AB, which also provided the products for the study.

REFERENCES

1. Lucas MG, Bedretidnova D, Berghmans LC, et al. Guidelines on urinary incontinence. http://uroweb.org/wp-content/uploads/20-Urinary-Incontinence_LR1.pdf. Accessed May 18, 2017.
2. International Organization for Standardization. ISO 16021:2000. Urine-absorbing aids—basic principles for evaluation of single-use adult incontinence absorbing aids from the perspective of users and caregivers. <https://www.iso.org/obp/ui/#iso:std:29832:en>. Published 2000. Accessed May 18, 2017.
3. International Organization for Standardization. ISO 15621:2011. Urine-absorbing aids—general principles of evaluation. <https://www.iso.org/obp/ui/#iso:std:50762:en>. Published 2011. Accessed May 18, 2017.
4. Farbroth A, Abbas S, Nihlstrand A, et al. Defining comfort for heavily-incontinent patients assisted by health care products in several contexts. Paper presented at: The Simon Foundation for Continence's Innovating for Continence Conference Series; April 2013; Chicago, IL.
5. NHS Choices. Urinary incontinence. <http://www.nhs.uk/conditions/incontinence-urinary/pages/introduction.aspx>. Published 2017. Accessed May 18, 2017.
6. Gray M, Kent D, Ermer-Seltun K, McNichil L. Assessment, use, and evaluation of body worn absorbent products for adults with incontinence: a WOCN Society Consensus Conference. *J Wound Ostomy Continence Nurs*. 2018;45(3):243-264.
7. International Organization for Standardization. ISO 9241-11:1998. Ergonomic requirements for office work with visual display terminals. www.iso.org/iso/catalogue_detail.htm?csnumber=16883. Published 1998. Accessed May 18, 2017.
8. International Organization for Standardization. ISO/DIS 9241-11.2:2016. Ergonomics of human-system interaction—part 11: usability: definitions and concepts. www.iso.org/iso/catalogue_detail.htm?csnumber=63500. Published 2016. Accessed May 18, 2017.
9. Swedish Ministry of Health and Social Affairs. National eHealth—the strategy for accessible and secure information in health and social care. <http://www.government.se/reports/2011/05/national-ehealth—the-strategy-for-accessible-and-secure-information-in-health-and-social-care>. Published 2011. Accessed May 18, 2017.
10. US Food and Drug Administration. Applying human factors and usability engineering to medical devices. Guidance for industry and Food and Drug Administration staff. <http://www.fda.gov/downloads/MedicalDevices/.../UCM259760.pdf>. Accessed May 18, 2017.
11. International Organization for Standardization. IEC 62366-1:2015 Medical devices—part 1: application of usability engineering to medical devices. <https://www.iso.org/standard/63179.html>. Accessed May 18, 2017.
12. Sauro J, Kindlund E. Making sense of usability metrics: usability and Six Sigma. Paper presented at: Proceedings of the Conference in Human Factors in Computing Systems (CHI 2005); April 22-27, 2006; Montreal, Canada.
13. Sauro J, Kindlund E. A method to standardize usability metrics into a single score. Paper presented at: Proceedings of the Conference in Human Factors in Computing Systems (CHI 2005); April 2-7, 2005; Portland, OR.
14. Sauro J, Kindlund E. Using a single usability metric (SUM) to compare the usability of competing products. Paper presented at: the Human Computer Interaction International Conference; July 22-27, 2005; Las Vegas, NV.
15. Pendrill LR, Petersson N. Metrology of human-based and other qualitative measurements. *Meas Sci Technol*. 2016;27:094003.
16. Kurowski A, Buchholz B, Punnett L; ProCare Research Team. A physical workload index to evaluate a safe resident handling program for nursing home personnel. *Hum Factors*. 2014;56:669-683.
17. Sauro J, Lewis JR. *Quantifying the User Experience*. 2nd ed. Cambridge, MA: Morgan-Kaufmann; 2016.
18. Likert R. A technique for the measurement of attitudes. *Arch Psychol*. 1932;140:1-55.
19. International Organization for Standardization. ISO/TS 20282-2:2013. IDT usability of consumer products for public use—part 2: summative test method. www.iso.org/iso/catalogue_detail.htm?csnumber=62733. Published 2013. Accessed May 18, 2017.
20. Wiklund M, Kendler J, Strohlic AY. *Usability Testing of Medical Devices*. Boca Raton, FL: CRC Press Taylor & Francis Group; 2011.
21. Tezza R, Bornia AC, de Andrade SF. Measuring Web usability using item response theory: principles, features and opportunities. *Interact Comp*. 2011;23:67-175.
22. Rasch G. On general laws and the meaning of measurement in psychology. *Proc Fourth Berkeley Symp Math Stat Probability*. 1961;4:321-333.
23. Parent AS. Management for containment: A review of current continence care provisions. Paper presented at: 6th Global Forum on Incontinence; March 30, 2016; Berlin, Germany.