

Older adults' use of an on-line decision support system: Usability and stability of assistive technology recommendations

Claudine Auger^{a,b*}, W.Ben Mortenson^c, Jeffrey W. Jutai^{d,e}, Noemie Seguin-Tremblay^f, Vanessa Chenel^{a,b}, and Manon Guay^f

^aDepartment, University, City, Country; ^bDepartment, University, City, Country

^aCenter for Interdisciplinary Research in Rehabilitation of Greater Montreal, Montréal, Canada; ^bSchool of Rehabilitation, Université de Montréal, Montréal, Canada;

^cDepartment of Occupational Science & Occupational Therapy, University of British Columbia, Vancouver, Canada, ^dInterdisciplinary School of Health Sciences, University of Ottawa, Ottawa, Canada, ^eLIFE Research Institute, Ottawa, Canada; ^fSchool of Rehabilitation, Université de Sherbrooke, Sherbrooke, Canada,

**Claudine Auger, École de réadaptation, Faculté de médecine, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, QC, Canada, H3C 3J7, claudine.auger@umontreal.ca*

Abstract

Online decision support systems (DSS) may help older adults self-select assistive technology (AT) by offering recommendations. User interactions with DSSs may change the recommendations they receive. Objective: We evaluated recommendations stability and usability of an online DSS. Methods: Middle aged and older adults (n=43) were observed while using the DSS. The stability of DSS recommendations (ATs and advice) was compared between two time points, using a three-point scale: no, partial, or full agreement. Usability was coded, referencing ISO standards. Results: Half (51%) of participants received AT recommendations from the DSS in both sessions, with full (14%) or partial (12%) agreement. All but one participant received advice, and almost all of them had full (40%) or partial (56%) agreement between sessions. Many of the usability issues appear to be the result of the users inaccurately measuring their

environment, challenges in understanding the questions being asked, and improperly making selections from the system. Discussion: Strict AT matching rules versus generic advice, and usability issues, likely reduced the matching rate and stability of AT recommendations. Conclusion: It appears that some users may require assistance with the system, and we suggest changes to the DSS format and content to improve stability and usability. (199 words)

Keywords: activities of daily living; aging; assistive technology; decision support system; validation studies; information and communication technologies

Introduction

Among Canadians aged 65 and older, 33% (1.4 million) report experiencing disability and 76% of these indicate they need help with activities of daily living (ADLs) (Statistics Canada, 2006; 2012a; 2012b). Bathing is consistently identified as being the most challenging ADL (Guay, Dubois, Corrada, Garant, & Kawas, 2014). It has great personal significance for older persons and self-bathing can help forestall future functional decline and disability (Ahluwalia, Gill, Baker, & Fried, 2010; Zingmark, Nilsson, Norström, Sahlén, & Lindholm, 2016). Assistive technology (AT) may be beneficial to promote safety, increase independence and maintain quality of life (World Health Organization [WHO] & The World Bank [TWB], 2011). Unfortunately, individuals may encounter a variety of challenges with AT procurement (Parant, Schiano-Lomoriello, & Marchan, 2017; WHO & TWB, 2011). Some areas of Canada

report wait lists up to 800 days to be assessed by home care professional due to limited public funding of community rehabilitation services (Cott, Devitt, Falter, Soever, & Passalent, 2007; McColl, Aiken, Birtwhistle, Corbett, Schroder, & Schaub, 2009; Raymond, Feldman, Prud'homme, & Demers, 2013). If individuals make AT related decisions without guidance from health care professionals, they will likely be confronted with an overwhelming number of products and may be concerned with the impartiality of equipment vendors.

To help older adults and their caregivers make informed decisions about AT, decision support systems (DSSs) were recently developed (Tucker et al., 2011). Online DSSs typically have three components: a) a knowledge base, b) a program for combining that knowledge with user-specific information, and c) an interface to collect data about the user and to provide relevant information (Berner, 2009). To have confidence in such DSSs, system error and user error have to be evaluated. System errors occurs if the same sequence of responses is inputted on multiple occasions, and the computer algorithm does not provide the same recommendations. User error occurs when the user does not provide the system with the same information during repeated assessments under similar circumstances. One would expect that if the same user navigated the system over a relatively short time frame, they would receive similar recommendations, unless a life change occurred (e.g., health status, new residence; Elwyn et al., 2009; Frost et al., 2007). Surprisingly, a recent scoping review (Chenel, Mortenson, Guay, Jutai, & Auger, 2018), revealed that only two studies, among 12 reviewed, measured test-retest reliability of DSSs (Lawrence, Streiner, Hazuda, Naylor, Levine, & Gafni, 2000; Sebban et al., 1995). None of these publications pertained to DSSs that guide AT self-selection. Another important aspect to investigate is the DSS's usability, which considers how a user interacts with a system in a specific context of use

to achieve a goal (International Organization for Standardization, 2018). For example, if older adults do not understand intuitively how to click on buttons in the DSSs interface, they may not enter their responses correctly and this would affect the recommendations sent by the DSS. Usability can therefore influence stability and has to be considered concomitantly.

This study focuses on a particular DSS designed in the UK by *ADL Smartcare* (ADL Smartcare, 2019) to guide AT selection by older adults and their caregivers. This system has essential features in common with other DSSs. Further details about the system are presented in Methods below. The purpose of the study was to examine the stability of the Canadian English version of the *ADL Smartcare* DSS over a one-week period among a sample of older adults who had experienced problems related to bathroom ADLs. A secondary objective was to explore how the usability of the system, since this may influence its stability.

Methods

A test-retest design was used to determine the stability of recommendations (i.e., AT and advice) offered to Canadian older adults who used the system on two occasions, one week apart. Ethical approval was granted by the local university ethics board and relevant health authorities (approval number: MP-22-2016-564).

Participants

Participants were 50 years old or over and lived in the community (i.e., they were not hospitalized or in a long-term care facility). To be included in the study, participants had to self-report experiencing, or having experienced, difficulties with at least one activity usually performed in the bathroom (e.g., bathing, toileting, grooming). Prior to conducting the second session, the interviewer explored whether the participant's

situation (e.g. health, social or living situation) had changed in the last week, since only stable participants were eligible. Those who were unable to read and understand English or provide their own consent were excluded.

Decision support system

Our study is the first to report on the measurement properties of the *ADL Smartcare* system's (ADL Smartcare, 2019). The version under study was addressing over 175 difficulties older adults might face in their day-to-day activities. The knowledge base of this DSS is supported by the clinical reasoning of expert panels, and it uses a hierarchical ordering of loss of capacity in a wide range of basic and instrumental activities of daily living among middle- and old-aged adults (Kingston, Collerton, Davies, Bond, Robinson, & Jagger, 2012; Wloch, Kuh, & Cooper, 2016). For example, in the bathroom area, users identify a problem they would like to resolve such as "*I am finding it quite hard to get in and out of the tub to soak in the bath*". These problems allow the identification of specific areas of need (AON). For each older adult, the DSS uses algorithms to match potential recommendations in any given AON (e.g., descriptions and pictures of AT products, as well as advice on how to install and use those products or perform ADLs) with the needs of the respondent. A series of sub-questions relating to health condition, capacities, and environment guides the final AT(s) and/or advice offered to the individual. Building on work done in the United Kingdom (UK), the *ADL Smartcare system* has been adapted to the English Canadian context (Chenel et al., 2016).

The DSS is accessible from a computer, tablet, or smart phone. The bathroom module of the DSS makes seven AONs available to each participant: 1-*I have difficulty getting on and off the toilet* (named 'getting on/off toilet' in the manuscript), 2-*I have difficulty*

cleaning myself after using the toilet ('cleaning myself'), 3-*I have difficulty cutting my toe nails* ('cutting toe nails'), 4-*As a man, when sitting on the toilet I have difficulty directing my urine into the bowl* ('directing urine'), 5-*Can you help me find the right Bath Lift* ('finding bath lift'), 6-*I am finding it quite hard to get in and out of the tub to soak in the bath* ('soaking in tub'), and 7-*I am finding it quite hard to step into the bathtub and (or) stand to take a shower* ('stepping into tub/standing in shower'. Each AON has a different set of questions (some questions appear in more than one AON) and different sets of possible recommendations. Figure 1 shows the three question formats. Numerical entries (weight, height) use a circular widget (Fig 1a) that requires clicking to increase or decrease the number. Multiple response options use buttons to select a statement (Fig 1b) or an illustration and statement (Fig 1c). At the end of each session, the participant viewed a list of AT recommendations and advice suggested by the DSS.

<Insert Figure 1 about here>

Procedure

Participants were interviewed twice 7 to 10 days apart. The first session (T1) began with a discussion on the participant's experience with the provision process of AT using response options and semi-structured questions. Then, a 3-minute introduction video to the DSS was presented, after which the participant's first impressions were gathered. The participant was then asked to choose the AON that best represented their past or present difficulties in the bathroom, and complete the questions using the study laptop computer. Notes on the usability of the DSS were taken as participants were asked to "think out loud" (Nielsen Norman Group, 2018), express feelings and opinions related to the DSS, and were prompted regarding their reactions (e.g., You seem to be hesitating, what is causing this?). The interviewer also took field notes of the

participant's behavior related to usability with the corresponding questions or sections of the DSS (e.g., did not scroll down the screen to see all the response options before answering; clicked multiple times; clicked on the wrong button). The interview was concluded with a sociodemographic questionnaire. By design, a test-retest study has to be conducted under the same conditions. Therefore, each participant was re-assessed with the same AON and procedure approximately one week later during the second session (T2).

All responses given by each participant during each session (T1 and T2) were audio recorded. All the data entered in the DSS and recommendations given to each participant were automatically exported in an Excel spreadsheet database.

Data analysis

Descriptive statistics were computed for all quantitative data using counts, range and percentages. Stability was assessed by comparing the recommendations (AT and advice separately) at T1 and T2. Cases were classified in four agreement levels: no agreement, partial agreement, full agreement, no product/no advice as described in Table 1.

<Insert Table 1 about here>

To compare AT recommended at both time points two approaches were used. First, specific brands and models of AT recommendations (e.g., AT ID; Savanah 2 inches raised toilet seat) during each session were compared. Second, a classification system (Centre régional d'information de démonstration et d'évaluation des aides techniques, 2002) was used to regroup specific AT products into generic categories of AT (e.g., Category: Transfers, Sub-category: Toilet transfers, Group: Raised toilet seats, with a

total of 26 distinct groups of bathroom ATs). Because the results of both approaches were comparable in terms of proportions of full/partial/no agreement, only the analysis by AT ID is presented. Statistical analyses were performed using Excel.

The content analysis of the usability field notes was conducted using a mix of a priori codes based on the usability framework of ISO 9241-11 (International Organization for Standardization, 2018) and emerging codes. The usability issues observed or reported by participants for each question and/or screen view were transcribed in a table by the third author and a research assistant, and they were regrouped into themes and sub-themes. Triangulation was enhanced by discussing iterations of the coding guide between the fourth and fifth author, having different backgrounds (i.e., rehabilitation science and web interface ergonomics), until a consensual final version was obtained through discussion. Logbooks and team meeting notes were used to support reflexivity.

Results

Forty-three (43) participants were recruited (Table 2). The majority were women (65%), retired (81%) and living in an urban area (95%). Almost all participants had experienced using the Internet (98%) and owned an information and communication technology device (93%), navigating it daily (77%). Half of the participants (51%) were receiving human assistance for their care in the bathroom. Almost all participants were currently using ATs in their bathrooms (98%). ATs included grab bars (93%), bathtub/shower seats (67%), raised toilet seats (49%), and commodes (19%). Use of mobility aids was common: cane (33%), wheelchair (33%), regular walker (16%), rollator walker (12%), or mobility scooter (5%).

<Insert Table 2 about here>

Agreement on ATs and on advice

Our analyses did not identify any system error that could affect the stability of the underlying algorithm; the following results thus report about user error, that is the information entered by the older adult into the system. Figure 2 displays the stability of AT and advice recommendations between both sessions. Almost half of the sample (n=21; 49%) received no AT recommendations on both sessions, while only one participant received no advice recommendation from the DSS at either T1 or T2. Agreement results (full, partial, no agreement) are presented below.

Regarding agreement on AT, one quarter of the participants had full (n=6; 14%) or partial (n=5; 12%) agreement. Another quarter of the sample (n=11; 26%) were classified as having no agreement on AT. The reason for this no agreement result was the same for all these participants: they received AT recommendations at one session, but no match was possible for the other session, since no ATs were recommended.

<Insert Figure 2 about here>

Regarding agreement on advice (Figure 2), 96% had full (n=17; 40%) or partial (n=24; 56%) agreement. Only 1 (2%) out of 43 participants had no agreement since he received no advice at T1.

Analysis regrouped per area of need (AON)

Erreur! Source du renvoi introuvable.3 illustrates the differences in AT and advice between the different AONs. Participants received up to 18 AT recommendations and up to 17 pieces of advice per session. These lists of recommendations varied widely. They could pertain to different brands of the same categories (e.g. various brands of bathing seats), as well as to a mix of categories (e.g. advice to install grab bars, step-by-

step instructions to enter in the tub, warnings).

<Insert Table 3 about here>

For AONs *getting on/off toilet, soaking in tub, stepping into tub/standing in shower* participants were recommended a greater number of ATs (up to 14-18 per session) in comparison to participants who chose AONs *cleaning myself, cutting toe nails* or *finding bath lift* (up to 3 per session). Similarly, the most advice was given to participants who had chosen AON *soaking in tub* or *stepping into tub/standing in shower* (up to 15-17 pieces of advice per session). AON 4 (“*As a man, when sitting on the toilet I have difficulty directing my urine into the bowl*”) was not chosen by any participant.

Agreement on AT regrouped per AON

Figure 3 represents the proportion and number of participants per level of agreement on AT for each AON. The distributions show that, except for *directing urine*, 2 to 18 participants selected each area of need. The three most frequently selected were AON *cutting toenails* (n=18), followed by *stepping into tub/standing in shower* (n=10), and *getting on/off toilet* (n=7).

<Insert Figure 3 about here>

Concerning agreement on ATs, four AONs had some agreement (full or partial) on AT between the two sessions (*getting on/off toilet, cleaning myself, soaking in tub, stepping into tub/standing in shower*) while the remaining two illustrated on Figure 3 did not. The AONs with the highest rate of full agreement on AT were *getting on/off toilet* (n=3), followed by *stepping into tub/standing in shower* (n=2), and *cleaning myself* (n=1) accounting for all 6 study participants with full agreement. The remaining

5 participants with partial agreement were distributed between AONs *getting on/off toilet* (n=3), *soaking in tub* (n=1) and *stepping into tub/standing in shower* (n=2).

Two AONs (*cutting toenails*, *finding bath lift*) were characterized by the highest proportions of participants without a single match of AT recommendations. For *cutting toenails*, three quarters (n=13) of participants had no AT recommended and the remaining participants had no agreement. In addition, neither of the two participants who chose *finding bath lift* had any AT recommended.

Agreement on advice regrouped per AON

The proportion of participants per agreement on advice for each AON is presented at Figure 4.

<Insert Figure 4 about here>

All AONs except *cleaning myself* had full or partial agreement on all advice. In the *cleaning myself* AON, one participant received no advice on both sessions, which explains the single no advice result for this AON in Figure 4. In this same AON, the no agreement comes from the fact that one participant did not receive any advice at T1. We noted that some of the advice provided was generic and not necessarily as specific as the ATs recommended. For example, AON 3 gave the advice “Hints and Tips for cutting nails” on both sessions to 17 out of 18 participants and AON 1 gave the advice “Hints and Tips for difficulties getting on and off the toilet” on both sessions to 6 out of 7 participants although not all of them received the same AT recommendations on both sessions.

Usability

Table 4 shows examples of discrepancy between the responses given at T1 and T2 by different participants for the three types of response options shown on Figure 1 (two types of multiple response options and the numerical entry). The last column on the right presents the field notes taken at T1 and T2. For Q5, at T1, one participant did not fully relate to the response options given since his/her condition varied from day to day. In addition, he entered an answer into the system that was contrary to intent he indicated verbally at T2. For question Q9, one participant did not understand the question or the illustration at T1, which led to her entering a different response from the one given at T2. For question Q8, one participant had measured in inches and since the display in feet and inches did not give the option to enter only inches, an error was made at T1. At T2, he made the conversion from inches to cm.

<Insert Table 4 about here>

From a total of 1374 usability notes from all sessions (T1 and T2), we identified four general themes (i.e., navigation, user behavior or system error, information content, and information presentation) and 31 subthemes. These themes as well as subthemes and examples are presented in Table 5.

<Insert Table 5>

Seven main usability challenges were identified regarding how to take and enter the physical bathroom measurements into the DSS. This is important since all AON required taking measurements, except AON 2 and AON 3. First, participants indicated having difficulty understanding some images that illustrated how measures should be taken. Second, mentally converting measurements taken in inches with the tape to metrics (e.g., feet and inches), as requested by the DSS was challenging. Third, physical

demands when taking measurements, including getting back and forth to the computer to answer questions about bathroom features, represented a burden. Indeed, different questions required measurements and they were not presented successively. Fourth, inadequate measurement procedures were observed (e.g., measuring tape was held at an angle, making the measure inaccurate). Fifth, three participants estimated the measurements instead of taking accurate measures. Sixth, there were issues with the use of the online measurement entry tool (e.g., difficulty entering a measurement with precision with the circular widget provided in the DSS). Finally, it was noted that some participants did not see or read the last part of questions, causing omissions or errors entering measures according to instructions.

Beside the measurement's usability issues, some questions were judged unclear (e.g., "Can you squeeze someone's hand tightly?" Tightly relative to what?) or not relevant to the participant's situation (e.g., participant in a wheelchair). Moreover, participants indicated that some response options were unclear or missing. For example, response options did not enable participants to indicate when their capacity varied from day-to-day. Some participants expressed being confused when images were provided as response options (e.g., the participant had to select the appropriate image following the question: "What is the shape of your most used toilet bowl?"). It was also observed that participants sometimes misread the question or did not scroll down the screen to look at all response options before answering. Finally, we sometimes observed a mismatch between what participants had said aloud and the selected button, resulting in a wrong entry or entirely skipping the question.

Discussion

We examined the stability and observed the usability of a DSS with 43 older adults who

used this online tool in their homes to address difficulties in performing ADLs in their bathroom. The one-week stability of recommendations for AT was relatively low, whereas the stability of advice recommendations was much higher. The higher percentage of participants with partial or full agreement on advice is explained by the fact that advice recommendations were not as specific as the AT recommendations, i.e. the system considered less of the user-entered data when providing some of the advice. Usability issues, many in relation with environmental measurements by older adults, have been noted and partly explain the lack of stability of the AT recommendations. We believe that there are important lessons to be learned that are not specific to the particular DSS in our study but which may be relevant for the research and development of similar systems.

There are several possible explanations for the relatively low stability of AT recommendations between the two time points. The framework of Collins (2003) identifies four cognitive processes (i.e., comprehension, retrieval, judgment and response formatting) required to enable completion of self-report measures that are relevant here. First, as noted above, we observed that respondents did not read all the words of the question or the available instructions. Therefore, comprehension might have been affected, leading to inconsistent responses between the two sessions. DSS systems should provide information and instructions that are understandable to the intended users, and responses should be captured in ways that minimize errors, and item irrelevance (Messick, 1995). Moreover, the system could provide a summary of all answers prior to the final submission to allow double-checking.

Second, while navigating, older adults had to retrieve information from their past history regarding rare activities (e.g. “Q: Which of the following do you have when

sitting on a toilet? R: Discomfort or pain at the base of your spine/ Discomfort in the back of your thighs/ Discomfort or pain at the base of your spine and in the back of your thighs/ None of these”), potentially leading to guessing. The effect of guessing on item reliability has been documented as an important source of inconsistency (Zimmerman & Williams, 2003). We would thus recommend that healthcare teams consider evidence of test-retest reliability as a basic selection criterion for any DSS, especially if data were collected through self-report.

Third, respondents’ judgment could be an issue, especially when they were asked to extrapolate their capacity to do unusual tasks they may have never performed (e.g., “Sitting on the bottom step of some stairs, would you be able to put your hands on the step behind you and push your bottom up onto it?”). Moreover, documenting environmental details requiring abstract judgement and observational skills might be challenging for untrained people [e.g., “When facing your most used toilet, are there any pipes or obstacles where the toilet paper rolls are in the pictures? (In other words, there is not enough clear space to put 1 toilet paper roll as shown.)”]. Although the DSS algorithm attempts to reproduce the expert clinical reasoning by breaking down activities into observable behaviors, some questions might benefit from being simplified and broken down into smaller elements that older adults can reliably report.

Finally, the design of graphical elements (e.g., widgets and images) affected the strategies used by older adults when responding. The concept of perceived affordance, referring to the design elements suggesting to the person how an object should be used (Norman, 1999; Zhao, Liu, Tang, & Zhu, 2013), is a key aspect to improve system usability. Affordance dictates that the user should understand intuitively the functions of icons and symbols in a system. As older adults are not all experienced DSS users, affordance of the interface might need to be improved to enhance the consistency of the

human-computer interaction during navigation. Addressing usability challenges captured in this study might ultimately improve the stability of the DSS.

It would be interesting to explore further if the nature of the questions plays a role in the failure to match. The usability investigation pinpointed a few specific questions that are problematic, particularly those about the environment. Bathroom measurements, selecting the shape of features, and identifying the presence of obstacles were complex questions (e.g., locating walls and pipes). This may mean that some older adults cannot reliably report environmental information and alternative ways to capture this information are needed. Indeed, some usability comments apply to many environmental questions that are presented in a similar format; addressing these issues could improve the reliability of future respondents. It is also possible that a guided-navigation by trained personnel rather than self-assessment would lead to higher matching rates and better reliability.

The fact that a large proportion of older adults did not receive any AT recommendation after navigating requires consideration. If the system was eventually commercialized as a pay-for-service model it would be important to make consumers aware of the fact that they might not receive ATs or advice to avoid disappointment if they expect to have AT recommendations *per se*. A possible “free access” to a minimal set of questions to determine eligibility might be an option worth exploring in the business model of such a DSS. This model would potentially avoid disappointment due to registration cost as well as unnecessary effort in answering a long series of questions by non-eligible persons. A subset of pre-requisite questions might clarify if the person is a good candidate for the system or direct them towards some kind of support (e.g., occupational therapist).

Due to the complexity of aspects that may affect the quality of the data entered by older adults in a DSS dedicated to the selection of ATs, we advise organizations to work with experts in the field of psychometrics, gerontology and rehabilitation to determine the level of support needed to obtain reliable recommendations. The stability of the recommendations is a first step that must be completed to provide evidence of their clinical validity. Indeed, users may provide the same response twice (thus be reliable) but if these responses are not accurate recommendations will be invalid. Therefore, a key piece of future work is to explore the validity of participants' responses against the judgment of a blinded health professional (e.g., occupational therapists). This process would allow an in-depth discussion about the pattern of responses and the logical links that lead to recommendations.

Study limitations

The study had two main limitations. First, there was an uneven distribution of AONs selected by study participants. No participant selected the AON *As a man, when sitting on the toilet I have difficulty directing my urine into the bowl*, a foreseeable result given that a majority of participants were women. Moreover, almost half the sample opted for *cutting toenails*, an AON with fewer questions and fewer available ATs, which in turn could have lowered the matching rates. Still, the aim was not to obtain an even distribution of participants among the seven AONs. Imposing a problem (to get data about each AON) about hypothetical situations could have affected the validity of our results.

Second, small non-random sample combined with a high number of potential recommendations (up to 18 ATs and 17 pieces of advice for a single participant) meant that we were underpowered to quantify the reliability of recommendations using

inferential statistics. Examining different sources of errors with a larger sample may help to guide the refinement of the DSS. For example, if specific areas of need are unreliable, it may mean that assistance in answering the DSS's questions is a better option for this particular problem or it may indicate that the inclusion/exclusion criteria for self-assessment have to be narrowed.

Conclusion

In a context of limited professional resources, recent initiatives have sought to support self-evaluation through DSSs for users of assistive technologies. We evaluated a Canadian adaptation of a commercially available DSS as a potential tool to be used by older adults in the absence of professional assistance. On one hand, we found that the stability of the device recommendations was low, despite including long lists of device recommendations. On the other hand, the advice provided was consistent, but quite broad and generic. Many of the instabilities of the device recommendations appear to be the result of the users inaccurately measuring their environment, challenges in understanding the questions being asked, and improperly making selections from the system. In light of these findings, it appears many potential users require assistance to successfully use the system. Although our results do not necessarily apply to all DSSs, they highlight the importance of studying how user error affects the results of such systems as well as their psychometric properties. To increase confidence in such DSSs, our results suggest changes to improve system usability, additional supports to guide older adults, and refinement of the eligibility criteria. The latter is particularly important since our participants did not necessarily represent the typical users of the original system in the UK.

Acknowledgements

This work was supported by the AGE-WELL Network of Centres of Excellence (Grant # AW CRP 2015-WP8.1). Drs Auger and Guay received salary support from the *Fonds de Recherche en Santé du Québec (FRQS)* and Dr Mortenson from the Canadian Institutes of Health Research (CIHR). The 2016 version of the ADL Smartcare^{LTD} system was shown during the interviews. It was modified since then to address some of the limitations identified.

Conflict of interest

All authors declare they have no conflicts of interest.

Data availability

The data that support the findings of this study are available on request from the corresponding author, CA.

References

- ADL Smartcare, (2019). About. Retrieved from <https://www.adlsmartcare.com/#aboutus>
- Ahluwalia, S. C., Gill, T. M., Baker, D. I., & Fried, T. R. (2010). Perspectives of older persons on bathing and bathing disability: a qualitative study. *Journal of the American Geriatrics Society*, 58(3), 450–456.
- Berner, E. S. (2009). *Clinical decision support systems: State of the art* (Publication No. 09-0069-EF). Rockville, Maryland: Agency for Healthcare Research and Quality.
- Centre régional d'information de démonstration et d'évaluation des aides techniques. (2002). *Classification des aides techniques* [Assistive device classification]. Retrieved from http://www.irglm.qc.ca/files/Classification_Aides_techniques_CRIDEAT.pdf
- Chenel, V., Auger, C., Mortenson, B. W., Jutai, J. W., Gore, P., Johnson, G., & Guay, M. (2016, May). Reliability and acceptability of an online decision support system for the self-selection of assistive technologies by older Canadians: A

- research protocol. *2016 IEEE International Symposium on Ethics in Engineering, Science and Technology* (pp. 1-7), Vancouver, BC: IEEE ETHICS.
- Chenel, V., Mortenson, W. B., Guay, M., Jutai, J. W., & Auger, C. (2018). Cultural adaptation and validation of patient decision aids: a scoping review. *Patient Preference and Adherence*, *12*, 321–332.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, *12*(3), 229–238
- Cott, C., Devitt, R., Falter, L., Soever, L., & Passalent, L. (2007). Barriers to Rehabilitation in Primary Health Care in Ontario: Funding and Wait Times for Physical Therapy Services. *Physiotherapy Canada*, *59*(3), 173–183.
- Disability Living Foundation (2019). *Looking for equipment to live independently?* Retrieved from <https://asksara.dlf.org.uk/?auth=sara5>
- Elwyn G., O'Connor A. M., Bennett C., Newcombe, R. G., Politi, M., Durand M.-A., ... Edwards, A. (2009). Assessing the quality of decision support technologies using the International Patient Decision Aid Standards instrument (IPDASi). *PLoS One*, *4*(3), e4705.
- Frost, M. H., Reeve, B. B., Liepa, A. M., Stauffer, J. W., Hays R. D., & Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group (2007). What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health*, *10*(Suppl 2): S94–S105.
- Guay, M., Dubois, M. F., Corrada, M., Garant, M. P., & Kawas, C. (2014). Exponential Increases in the Prevalence of Disability in the Oldest Old: A Canadian National Survey. *Gerontology*, *60*(5), 395–401.
- International Organization for Standardization (2018). ISO 9241-11:2018, Ergonomics of human-system interaction - Part 11: Usability: Definitions and concepts, 2nd edition. *ISO Standards catalogue*. Geneva, Switzerland: International Organization for Standardization.
- Kingston A., Collerton J., Davies K., Bond J., Robinson L., & Jagger C. (2012). Losing the Ability in Activities of Daily Living in the Oldest Old: A Hierarchic Disability Scale from the Newcastle 85+ Study. *PLoS ONE*, *7*(2): e31665.
- Lawrence, V. A., Streiner, D., Hazuda, H. P., Naylor, R., Levine, M., & Gafni, A. (2000). A cross-cultural consumer-based decision aid for screening mammography. *Preventive medicine*, *30*(3), 200–208.

- McCull, M., Aiken, A., Birtwhistle, R., Corbett, S., Schroder, C., & Schaub, M. (2009). *Why Are There No Rehabilitation Professionals in Family Health Teams? Final report submitted to the Ontario Neurotrauma Foundation and Ontario Rehabilitation Research Advisory Network*. Kingston, ON: Queen's University.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741–749.
- Nielsen Norman Group (2018). *Thinking aloud: The #1 usability tool*. Retrieved from <https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/>.
- Norman, D. A. (1999). Affordance, conventions, and design. *Interaction*, *6*(3), 38–42.
- Parant, A., Schiano-Lomoriello, S., & Marchan, F. (2017). How would I live with a disability? Expectations of bio-psychosocial consequences and assistive technology use. *Disability and Rehabilitation: Assistive Technology*, *12*(7), 681–685.
- Raymond, M.-H., Feldman, D., Prud'homme, M.-P., & Demers, L. (2013). Who's Next? Referral Prioritization Criteria for Occupational Therapy in Home Care. *International Journal of Therapy and Rehabilitation*, *20*(12), 580–589.
- Sebban, C., Browman, G., Gafni, A., et al. Design and validation of a bedside decision instrument to elicit a patient's preference concerning allogeneic bone marrow transplantation in chronic myeloid leukemia. *American Journal of Hematology*. *1995*, *48*(4):221–227
- Statistics Canada (2006). *Adults with disabilities that need help with everyday activities, by sex and age groups, Canada, 2001 and 2006: Table 1*. Retrieved from: <http://www.statcan.gc.ca/pub/89-628-x/2010015/tbl/tbl1-eng.htm>
- Statistics Canada (2012a). *Disability in Canada: Initial findings from the Canadian Survey on Disability*. Retrieved from <http://www.statcan.gc.ca/pub/89-654-x/89-654-x2013002-eng.htm>.
- Statistics Canada (2012b). *Prevalence of disability for adults by sex and age group, Canada, 2012: Table 1.1*. Retrieved from: <http://www.statcan.gc.ca/pub/89-654-x/2013001/tbl/tbl1.1-eng.htm>.
- Tucker, S., Brand, C., O'Shea, S., Abendstern, M., Clarkson, P., Hughes, J., ... Challis, D. (2011). An evaluation of the use of self-assessment for the provision of community equipment and adaptations in English local authorities. *British Journal of Occupational Therapy*, *74*, 119–28.

- World Health Organization & The World Bank (2011). *World report on disability*. Geneva, Switzerland: World Health Organization.
- Wloch E. G., Kuh, D., & Cooper, R. (2016). Is the Hierarchy of Loss in Functional Ability Evident in Midlife? Findings from a British Birth Cohort. *PLoS ONE* 11(5), e0155815.
- Zhao, Y., Liu, J., Tang, J., & Zhu, Q. (2013). Conceptualizing perceived affordances in social media interaction design. *Aslib Proceedings*, 65(3), 289–303.
- Zimmerman, D. W., & Williams, R. H. (2003). A New Look at the Influence of Guessing on the Reliability of Multiple-Choice Tests. *Applied Psychological Measurement*, 27(5), 357–371.
- Zingmark, M., Nilsson, I., Norström, F., Sahlén, K. G., & Lindholm, L. (2016). Cost effectiveness of an intervention focused on reducing bathing disability. *European Journal of Ageing*, 14(3), 233–241.

Table 1: Definitions and examples of agreement levels

Agreement level	Definition	Criteria	Examples	
			T1	T2
No agreement	No identical AT/advice recommendations between T1 and T2	At least one AT/advice at T1 OR at T2. Otherwise, the number of AT/advice has no importance as long as all are different.	e.g., Participant PH1 No AT	QC_05 Swedish bath rails - Bath safety bar; Swedish bath rails - Tub clamp rail; Swedish bath rails - Bathtub shower bar
Partial agreement	Some overlap between AT/advice recommendations at T1 and T2	At least two AT/advice at T1 AND at T2. Otherwise, the number of AT/advice at T1 and T2 has no importance, as long as at least one is identical and at least one is different.	e.g., Participant PH1 Hints and Tips - Hints and Tips: How to set the height of adjustable toilet frames; Hints and Tips - Hints and Tips for how to use grab rails around your toilet; Take Note - Caution when fixing products to bathroom floors; Take Note - Common solutions for using more than one toileting product; Warning - For safety, please ensure that all metal grab rails are grounded.	BC_01 Hints and Tips - Hints and Tips: How to set the height of adjustable toilet frames; Hints and Tips - Hints and Tips for how to use grab rails around your toilet; Take Note - Caution when fixing products to bathroom floors; Take Note - Common solutions for using more than one toileting product
Full agreement	Identical AT/advice at T1 and T2	The number of AT/advice at T1 AND T2 are the same and all are identical.	e.g., PH2 Bath or shower boards and benches - Savannah Slatted Bathboard (WITH handle) 76 cm long (30 in); Bath or shower boards and benches - Aquasense bathboard; Swedish bath rails - Bath safety bar; Swedish bath rails - Tub clamp rail; Swedish bath rails - Bathtub shower bar	BC_08 Bath or shower boards and benches - Savannah Slatted Bathboard (WITH handle) 76 cm long (30 in); Bath or shower boards and benches - Aquasense bathboard; Swedish bath rails - Bath safety bar; Swedish bath rails - Tub clamp rail; Swedish bath rails - Bathtub shower bar
No product/No advice	No AT/advice recommended at either T1 or T2	This case happens only when there are no AT/advice at T1 AND at T2.	e.g., Participant PH2 No AT	BC_07 No AT

Legend: AT: Assistive technology; T1 session 1; T2 Session 2

Table 2: Characteristics of the participants (n=43)

Variable	n (%)
Sex (women)	28 (65)
Age (years)	
50-60 y. o.	9 (21)
61-70	14 (33)
71-80	13 (30)
81-90	7 (16)
Canadian province of residency	
Quebec	12 (28)
Ontario	13 (30)
British Columbia	18 (42)
Education	
College or other non-university certificate	9 (21)
University certificate (above bachelor's level)	9 (21)
High school	7 (16)
Bachelor's level	6 (14)
University certificate (below bachelor's level)	6 (14)
Trade certificate or diploma	4 (9)
Less than high school	2(5)
Marital status	
Divorced	12 (28)
Married	9 (21)
Single, never married	9 (21)
Widowed	7 (16)
Separated	3 (7)
Living common-law	3 (7)
Primary occupational status	
Retired	35 (81)
Volunteering	6 (14)
Working at a paid job or business	4 (9)
Long term illness	4 (9)
Others (e.g., caring for children, household work, looking for a paid job)	4 (9)
ICT devices used	
Desktop	34 (79)
Laptop	26 (60)
Tablet	16 (37)
Smartphone	23 (53)
Others (flip phone, old fashioned phone, smart TV)	3 (7)

Internet use for medical or health-related information	
Yes	37 (86)
No	5 (12)
Missing data	1 (2)
Living environment	
Urban area (city, town, conurbation)	41 (95)
Rural area (small settlement, village, hamlet)	2 (5)
Remote (isolated sector, lack of transportation)	0 (0)
Has used internet	
Yes	42 (98)
No	1 (2)
Internet use frequency	
At least once a day	33 (77)
At least once a week (but not every day)	6 (14)
At least once a month (but not every week)	1 (2)
Less than once a month	0 (0)
Missing data	3 (7)
Ownership of an Information and Communication Technology device	
Yes	40 (93)
No	3 (7)

Table 3: Number (min and max) of AT (Assistive Technology) and advice recommended to participants, total number of participants and questions per AON (Area Of Need)

AON	Situations introduced by DSS	AT		Advice		n	Number of questions
		Min	Max	Min	Max		
1	<i>I have difficulty getting on and off the toilet</i>	0	18	4	8	7	30
2	<i>I have difficulty cleaning myself after using the toilet</i>	0	2	0	1	2	10
3	<i>I have difficulty cutting my toenails</i>	0	3	1	4	18	10
4	<i>As a man, when sitting on the toilet I have difficulty directing my urine into the bowl</i>	-	-	-	-	0	4
5	<i>Can you help me find the right Bath Lift</i>	0	0	5	7	2	31
6	<i>I am finding it quite hard to get in and out of the tub to soak in the bath</i>	0	14	6	17	4	49
7	<i>I am finding it quite hard to step into the bathtub and (or) stand to take a shower</i>	0	15	3	15	10	45

AT: assistive technology; AON: Area of need; ADL; Activity of daily living; DSS: Decision support system; Min: minimum; Max: maximum; n= number of participants who selected the area of need

Table 4: Examples of discrepancies between data entered at T1 and T2

Question	Answer T1	Answer T2	Field notes
Q5 (AON3): Can you squeeze someone's hand tightly? (Type of response : see Figure 1 b) Multiple choice)	Not with either hand	Yes, with both hands	Participant : PH2_ON_03 T1 : Condition is variable: "Not all the time, it depends on the pain in my hands" T2: Participant said she could not open jars but selected that "yes" she could squeeze with both hands.
Q9 (AON1): When facing your most used toilet, are there any pipes or obstacles where the toilet paper rolls are in the	Obstacle on right hand side of	Obstacle on left hand side of	Participant : PH2_ON_02 T1: The image shows toilet paper rolls on the floor, on the right and/or left side of the

<p>pictures? (In other words, there is not enough clear space to put 1 toilet paper roll as shown.) (Type of response : see Figure 1 c) Multiple choice with illustrations)</p>	toilet (facing view)	toilet (facing view)	<p>toilet. Participant says: “My paper is on the wall, this doesn’t make sense.” T2: Participant realized that the question describes the front view, not the view when sitting on the toilet. Participant did not read properly at T1, may explain the different response at T2.</p>
<p>Q8 (AON1): For your most used toilet, what is the height from the floor to the top of the toilet bowl with the seat up. Example: 15 inches (40 cm). (Type of response : see Figure 1 a) Circle slider)</p>	36cm	73cm	<p>Participant : PH2_BC_07 T1: Participant measured in inches. Interviewer noticed a mistake when entering the number due to the display being in ft/in and not only inches. T2: The participant measured in inches but converted from inches to cm before entering the response.</p>

Table 5: Usability issues: themes and examples

1. Navigation: Related to elements of navigation (widgets, buttons, ease of navigation), links, steps		
1.1 Clarity and constancy of navigation:	Simplicity of navigation and possibility to perform it in a similar fashion from page to page	e.g., Participant didn't know how to go back to previous question.
1.2 Hyperlinks and widgets:	Ease of use regarding hyperlinks and widgets (e.g. measurement tool, menus)	e.g., Participant cannot figure out how to use the round widget to enter measurements.
1.3 Icons and buttons:	Size, intuitiveness, labelling and positioning of icons and buttons	e.g., Participant not sure if the grey rectangle is a clickable button.
1.4 Loading:	Loading information and time vs consequences	e.g., Loading time long and no indicator after clicking on "Start Assessment" button.
1.5 Help and information:	Possibility to get help and information related to or to facilitate the progress of an assessment on DSS and the application of the recommendations	e.g., Participant would like a "request further help" button.
1.6 Number of steps:	The number of steps influencing the length and fluidity of the session and the participant's degree of cognitive and physical effort	e.g., Participant says it's not practical to go back and forth for measurements.

2. User behavior or system error: Related to the user or system but not clearly linked to the DSS interface		
2.1 User preference and capacity:	Physical or cognitive abilities or knowledge (e.g. knowledge of computers). Also, statements of preference from the participant.	e.g., Participant has difficulty using the mouse. Clicks on right button instead of left.
2.2 Accuracy of entered info:	The participant does not enter accurate information in response to a question, for a reason which is not linked to a specific problem with the interface	e.g., Participant seems to have chosen a device that she CAN use (instead of CANNOT as asked in the question).
2.3 System malfunction:	Technical problems unrelated to DSS's interface (e.g. internet connexion)	e.g., Response was marked as clicked (blue) but nothing happened. Had to click again.
3. Information content: Related to textual content of information presented in the DSS		
3.1 Question and response options:	Related to textual information presented in the DSS's questions and response options	
3.1.1 Clarity of question:	Vocabulary and style of the sentences and general comprehension of the question, clarity and length	e.g., Participant wonders if the question is about hand or toenails.
3.1.2 Clarity of response options:	General comprehension of the response options, of what each of them entails	e.g., Participant doesn't know the difference between "poor" and "very limited" vision.
3.1.3 Adequacy of question:	Adequacy in the specific context of the participant, participant's perception of the question's usefulness	e.g., Participant hesitated because he has good sight from far but bad from up close.
3.1.4 Adequacy of response options:	Adequacy in the specific context of the participant, participant's perception of the response option's usefulness	e.g., Participant would like to have the option "cannot use any of them" or "all of them".
3.2 Recommendations:	Related to textual information presented in the DSS's recommendations	
3.2.1 Clarity of recommendations:	General comprehension of the recommendations, clarity and length	e.g., Participant does not know what "rails need to be grounded" means.
3.2.2 Pleasantness of recommendations:	How pleasing it is to read the recommendation	e.g., Participant likes the one-handed nail clipper - thinks it is a good idea
3.2.3 Adequacy of recommendations:	Adequacy in the specific context of the participant, participant's perception of the recommendations' usefulness and helpfulness	e.g., Participant perceives that soaking toenails before clipping is helpful.
3.3 Area of need:	Related to textual information in the sentence related to the area of need at the beginning of an assessment	

3.2.5 Adequacy of area of need:	Adequacy in the specific context of the participant, participant's perception of the area of need's usefulness and helpfulness	e.g., Participant says it is not always necessary to “get on the toilet” as said in area of need.
4. Information presentation: Related to the presentation of information in the DSS		
4.1 Image:	Related to images presented in the DSS	
4.1.1 Relevance of image:	Image’s relevance and depiction of the context or relevance of adding an image/video for better understanding	e.g., The recommendations mentioned a bath lift would not fit. The participant would have liked a picture of a bath lift.
4.1.2 Understandability of image:	Clarity of what the image represents in relation to the question asked	e.g., It is not obvious for the participant that the image is a side view.
4.1.3 Size and proportions of image:	Size and proportions of images	e.g., Participant wants a bigger image.
4.1.4 Pleasantness of image:	How pleasing the image is	e.g., Participant says the “diagrams are terrific”.
4.2 Text:	Related to textual and numerical information presented in the DSS	
4.2.1 Positioning and grouping of info:	Grouping of corresponding information and positioning relative to one another or to the global context	e.g., Participant wants advisory messages to all be on one page.
4.2.2 Adequacy of measurement related info:	Adequacy of the measurement related information in the specific context of the participant (e.g. units)	e.g., Participant does not know how to convert to kg.
4.2.3 Text readability:	Ease of reading text	e.g., Images and text size are fine for the participant.
4.3 Ease of finding important info:	Highlighting and ease of finding important information	e.g., Participant had not seen the thumbnail leading to advice.
4.4 Overall presentation:	Design of the website, general presentation including text and images	e.g., Participant says there are too many text colors (white, blue, black and grey).

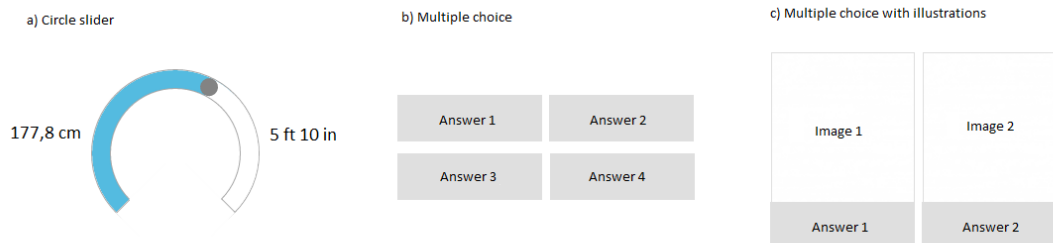


Figure 1: Templates of the three types of response options. 1a. Circle slider to enter height, body weight and other bathroom measurements, 1b. Multiple choice with text, 1c. Multiple choice with illustration and text.

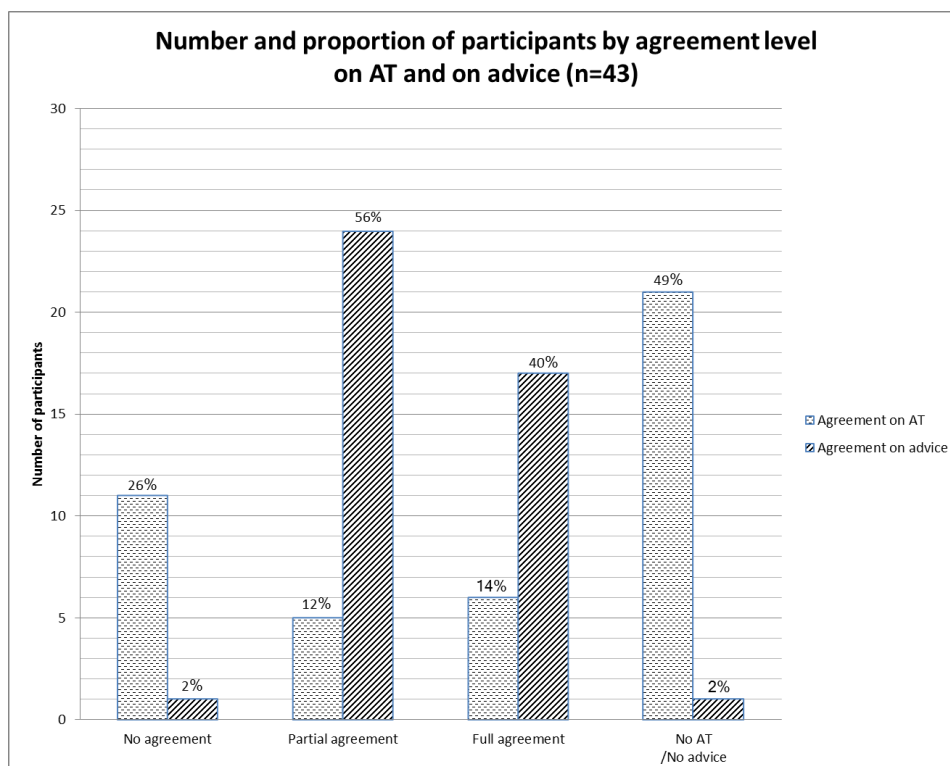


Figure 2: Stability of AT (Assistive Technology) and advice recommended from T1-T2

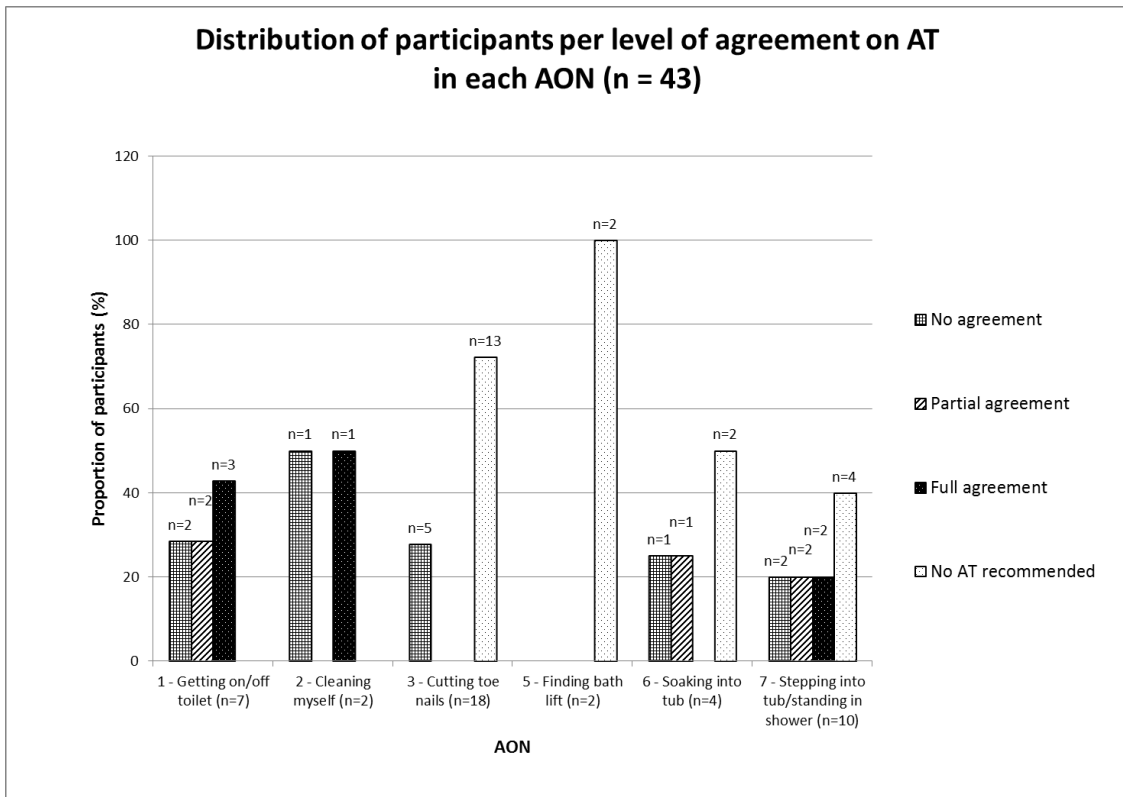


Figure 3: Agreement on AT (Assistive Technology) per AON (Area Of Need)

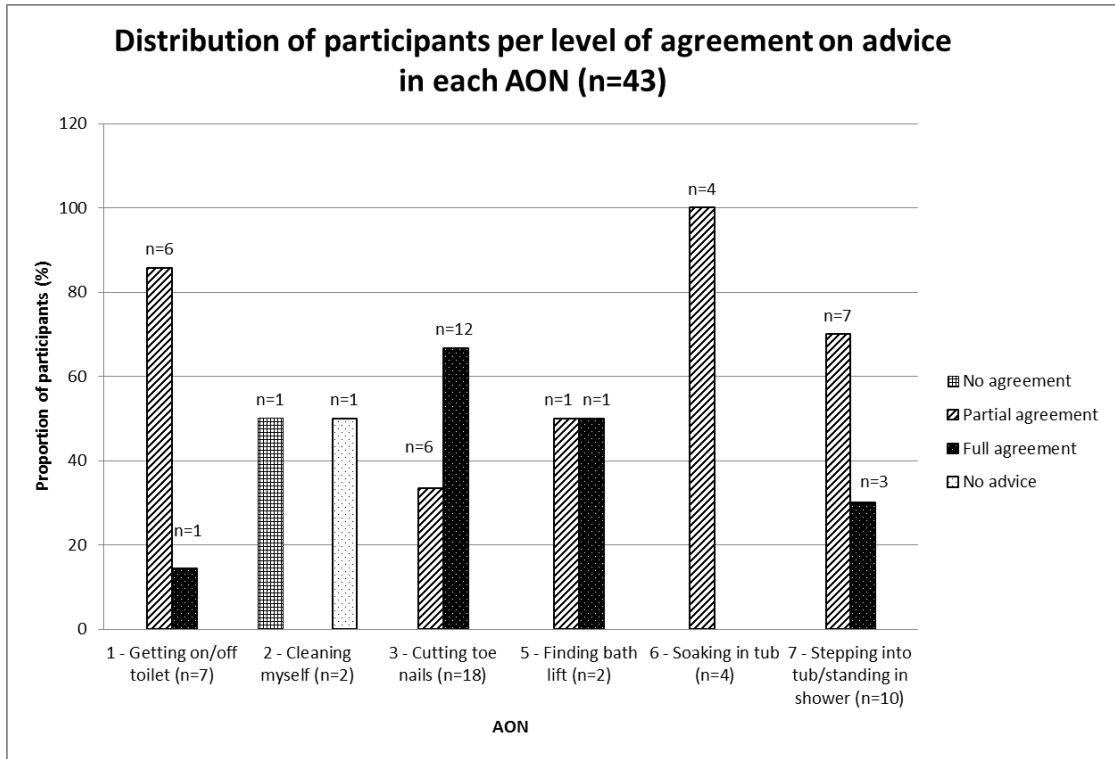


Figure 4: Agreement on advice per AON (Area Of Need)