

Evaluation of Lab Trials

Deliverable D7.3A

Project identification

Project Number	AAL-2016-069
Duration	30 months
Coordinator	Jon Arambarri
Coordinator Institution	Virtualware
Website	www.memento-project.eu

Document Identification

Deliverable ID	D7.3A
Release number/date	V4 / 25.09.2019
Checked and released by	Martina Pigliautile (UNIPG)
Work Status	Running
Review Status	Released

Key Information from "Description of Work"

Deliverable Description	Evaluation of lab trials
Dissemination Level	Restricted
Deliverable Type	Report
Original due date	30/03/2019

Authorship & Reviewer Information

Editor	Theresa König (MUW)
Partners contributing	MUW, Bidaideak, UNIPG, AIT
Reviewed by	Martina Pigliautile (UNIPG)

Release History

Release Nr	Date	Author(s)	Release Description/Changes made
V1	11.08.2019	Theresa König (MUW)	First version
V2	12.08.2019	Martina Pigliautile (UNIPG)	Revision
V3	27.08.2019	Theresa König (MUW)	Adapted Version
V4	24.09.2019	Martina Pigliautile (UNIPG)	Final revision
V5	21.10.2019	Theresa König (MUW)	Final version

Memento Consortium

Partner 1	VirtualWare
Contact Person	Jon Arambarri / jarambarri@virtualwaregroup.com
Partner 2	AIT Austrian Institute of Technology - GMBH
Contact Person	Christopher.Mayer / Christopher.Mayer@ait.ac.at
Partner 3	bkm design working group
Contact Person	Stefan Moritsch / moritsch@bkm-format.com
Partner 4	Medical University of Vienna
Contact Person	Elisabeth Stoegmann / elisabeth.stoegmann@meduniwien.ac.at
Partner 5	Wetouch
Contact Person	Christian Schueler / christian.schueler@wetouch.at
Partner 6	Integris
Contact Person	Teresa Pizzuti / teresa.pizzuti@integris.it
Partner 7	Università di Perugia
Contact Person	Patrizia Mecocci / patrizia.mecocci@unipg.it
Partner 8	Bidaideak - Sociedad Vasca de Personas con Diversidad Funcional
Contact Person	Oscar Aguila / oscar.bidaideak@gmail.com
Partner 9	Citard Services LTD
Contact Person	Eleni Christodoulou / cseleni@citard-serv.com

Table of Content

Executive Summary	9
1 About this Document	10
1.1 Role of the Deliverable	10
1.2 Relationship to other Memento Deliverables	10
1.3 Structure of this Document	11
2 Aim	11
3 Lab Trial Execution	12
3.1 User Characteristics	12
3.1.1 Recruitment	12
3.1.2 End Users	13
3.1.3 Caregivers	14
3.1.4 Investigators	15
3.2 Material and Methods	16
3.2.1 MEMENTO Prototype	16
3.2.2 Use Cases and Tasks	16
3.2.3 Evaluation Tools	18
4 Results	19
4.1 Observation forms	19
4.1.1 General observation and comments	19
4.1.2 General task	20
4.1.3 Medication	21
4.1.4 Appointments	23
4.1.5 Shopping	25
4.1.6 Getting ready	27
4.1.7 Lost outside	29
4.2 Task completion time	31
4.3 SUS Questionnaire	32

5	Main Results and Suggestions	35
5.1	SUS Report	35
5.2	Task Completion	37
5.3	User Experience	38
5.4	Target group related challenges	38
5.5	Technical problems	39
5.6	Design aspects	40
6	Conclusions	41
7	References	42

List of Figures

Figure 1 – Use cases that build the basis of tasks for lab trials	17
Figure 2 – SUS end-users and caregivers.....	36

List of Tables

Table 1: Relationship to other Memento Deliverables	10
Table 2: Patient Selection Criteria for Lab Trials	12
Table 3: End-User Characteristics.....	13
Table 4: Caregiver Characteristics	15
Table 5: Evaluation Tools	18
Table 6: Task completion time.....	31
Table 7: System Usability Scale	33
Table 8: SUS Interpretation	37

Abbreviations

Abbrev.	Description
AAL	Ambient Assisted Living
MUV	Medical University of Vienna
UNIPG	University of Perugia
MCI	Mild Cognitive Impairment
AD	Alzheimer's Disease
CRI	Cognitive Reserve Index
ADL	Activities of daily living
IADL	Instrumental activities of daily living
MMSE	Mini-Mental State Examination
NIA AA	National Institute on Aging and the Alzheimer's Association
SUS	System Usability Scale
P	Patient
C	Caregiver
Users	Patient and caregiver

Executive Summary

Memento aims to provide a persuasive system supporting memory and moments of people with early stage of dementia. In order to develop a functional and user-friendly solution with high user acceptance, we tested the system together with the target group. The Evaluation of Lab Trials deliverable describes the procedure of evaluating the first MEMENTO prototype, the characteristics of the trial participants and outcomes including observation forms, task completion times and system usability report.

Furthermore, difficulties from the end-user side, as well as technical and design problems are discussed.

1 About this Document

1.1 Role of the Deliverable

This deliverable provides information about user experience, engagement and acceptability of the first MEMENTO prototype according to the lab trials. Evaluating difficulties in performing tasks due to technical and design reasons, but especially due to challenges with regard to our end user group, will help to further improve our system.

1.2 Relationship to other Memento Deliverables

Deliverable	Relation
D2.2 – End users requirements	Description of end-users recruited for lab trials
D2.3 – Definition of Use Cases and Scenarios	Use cases defined in this deliverable are used as basis for lab trial tasks.
D3.1 – Specification of Hardware Design and User Interface	Describes the user interface design for the software components.
D4.1 A – Hardware Specification	Specifies the hardware design of the MEMENTO system.
D5.1 A – Software Specification	Specifies the software of the MEMENTO system.
D6.3 - Test plan for continuous expert tests	Ensures, that the implemented features are properly tested
D7.1 - Definition of lab trial protocol	Overall research design of the lab trials.
D7.2 A – Protocol for lab trials	Describes the strategy, preparation and execution of the lab trials in detail

1.3 Structure of this Document

This document describes the results of the lab trials performed together with patients and their caregivers. The first part of the document proposes the aim of the lab trials, describing the aspects of the MEMENTO device important to evaluate during this phase of development. It is followed by the section lab trial execution, which contains information about the participants of the trials and shortly describes material and methods used for testing and evaluating the MEMENTO system. Eventually, the fourth chapter provides the results of the lab trials, including detailed observation based on taking notes of users and caregivers thinking aloud during testing, amount of time required by user to complete the task and evaluation of the system usability questionnaire (SUS, Brooke 1996). Main results and suggestions are discussed at the end of the deliverable.

2 Aim

The lab trials were executed using a functional prototype of the MEMENTO system by the clinical partners. Evaluation techniques such as observation, thinking aloud, interview and questionnaires took place. The lab trial phase focuses on the user experience. User engagement and user acceptability are fundamental requirements of technology and if user perceptions do not transfer well and the design misleads them, it is likely to result in poor performance and product abandonment.

Therefore, the evaluation of those aspects after testing the MEMENTO device with users and caregivers during lab trials is of high value for the further development of the system.

Additionally, data about difficulties in performing tasks from end-user side, as well as about technical problems and design aspects were collected and will provide valuable information for further improvement of MEMENTO.

3 Lab Trial Execution

This section describes the execution of the lab trials, including the characteristics of patients and caregivers, materials used and a description of the protocol and evaluation methods.

3.1 User Characteristics

3.1.1 Recruitment

In line with the inclusion criteria defined within deliverable D2.2, clinical partners recruited patients treated at the dementia outpatient clinic MUV, Bidaideak and UNIPG. Before starting with the lab trials, written informed consent was provided by the patient or their legal guardian. In each clinical center 5 Patients with a diagnosis of MCI due to AD or mild AD according to the NIA AA criteria (McKhann, Knopman et al. 2011) with an MMSE 28 – 24 (inclusive) and their caregivers participated in the trials.

As stated in deliverable 7.1A, additional information was collected about each patient and respective caregiver. Both mandatory criteria for patient recruitment and optional patient traits are summarized in Table 2.

Table 2: Patient Selection Criteria for Lab Trials	
Mandatory Criteria	
Diagnosis of MCI due to AD and mild AD (amnestic type) (McKhan criteria)	
Activities Of Daily Living	Lawton - Brody Instrumental Activities Of Daily Living Scale (IADL) equal or below 5 a) subjects must be able to dial a few well-known numbers on the cellular phone b) subjects that are able to get around (or travel) outside of the home (alone or accompanied)
Correct total score Mini-Mental State Examination (MMSE)	RANGE 24-28

Optional Traits
Different levels of cognitive reserve (CRIq scores)
Different levels of technical proficiency

3.1.2 End Users

5 End-users were recruited in each outpatient clinic in Italy, Austria and Spain. An overview is provided in Table 3: End-User Characteristics.

- Cognitive reserve established with Cognitive Reserve Index (CRI): the concept of "reserve" has been used to explain the difference between individuals in their capacity to cope with or compensate for pathology. Considering the importance of the cognitive reserve, the CRI (Nucci, Mapelli et al. 2012) was taken into account. The CRI is established by a semi-structured interview that gathers and quantifies all the experiences that a person has acquired throughout their life.
- Technical proficiency patient: defined as the skills required to operate an information system (i.e., a hardware/software solution). Our ambition was to test the MEMENTO device with end users having different levels of technical skills.
- Age and Sex: both aspects should be considered in terms of the general attitude towards technology, design requirements and needs regarding the individual life phase.

Table 3: End-User Characteristics					
Outpatient Clinic: UNIPG					
PATIENT #	1	2	3	4	5
AGE	81	81	84	66	67
SEX	F	F	M	F	M
MMSE	28	28	24	26	28
CRI	medium-high	medium	medium	medium	high
TECHNICAL	low	high	low	low	high

PROFICIENCY					
Outpatient Clinic: MUV					
PATIENT #	6	7	8	9	10
AGE	54	52	76	60	74
SEX	M	M	M	F	F
MMSE	25	25	24	27	24
CRI	medium	medium	high	medium-high	high
TECHNICAL PROFICIENCY	high	high	medium	high	low
Outpatient Clinic: Bidaideak					
PATIENT #	11	12	13	14	15
AGE	72	79	86	84	76
SEX	F	F	F	F	M
MMSE	27	24	24	24	26
CRI	medium	medium-high	medium	medium	medium
TECHNICAL PROFICIENCY	low	low	medium	low	low

3.1.3 Caregivers

The patients were accompanied by their caregivers. An overview is provided in Table 4: Caregiver Characteristics

- Caregiver status: the caregiver status is relevant regarding their availability in daily live. Subjects living with their spouse or in a family context, as well as subjects living alone with an informal supervisor (son/daughter/niece/...) will be included in the trial.
- Technical proficiency caregiver: the technical skills of the caregiver are important for supporting the patient and using various software solutions of the MEMENTO system (i.e., accessing the calendar from another technical device).

Table 4: Caregiver Characteristics					
Outpatient Clinic: UNIPG					
	1	2	3	4	5
STATUS	daughter (no co-living)	aunt's niece (living at down floor)	wife	husband	wife
TECHNICAL PROFICIENCY	high	high	low	low	high
Outpatient Clinic: MUV					
	6	7	8	9	10
STATUS	wife	wife	wife	husband	husband
TECHNICAL PROFICIENCY	high	low	low	high	medium
Outpatient Clinic: Bidaideak					
	11	12	13	14	15
STATUS	daughter (no co-living)	certified caregiver	certified caregiver	daughter (no co-living)	sister
TECHNICAL PROFICIENCY	high	high	high	high	medium

3.1.4 Investigators

The lab trials took place in the outpatient clinics in the presence of one or more persons from the clinical team. We observed the trials, took notes and monitored the user acceptability, engagement and perception. We also focused on detection of open UX issues and challenges for the system by means of a thorough test in a collaborative setup.

3.2 Material and Methods

Each patient and their caregiver were be given enough time (1-2 hours each) to test the device, in particular to try out all developed use cases. The functionalities of the system to assess were presented to the users in different scenarios and each user had time to discover them before the trial tasks. Different evaluation techniques were used, such as observation, notes of thinking aloud of patients and caregivers, and a questionnaire based on System Usability Scale (SUS) (Brooke 1996). Detailed information about the protocol used for lab trials are contained in deliverable 7.2A.

3.2.1 MEMENTO Prototype

A first prototype of the MEMENTO system was used for the lab trials. Hardware design and user interface specifications can be found in deliverable D3.1. Deliverable D6.2A describes the features subjected to testing for the first prototype. The tasks for lab trials were created based on those features and four selected use cases.

3.2.2 Use Cases and Tasks

High priority features were implemented, which refer to the most important features needed by MEMENTO users, as described in D6.3. Those features, based on four use cases will be tested during the lab trials (see Figure 1):

- Medication
- Appointments
- Shopping
- Getting Ready

The mentioned use cases are described in deliverable D2.3 and were developed by analysis of several workshops with patients and their care givers. Tasks to be performed by the users and caregivers were defined for each use case in addition to general tasks, such as creating caregiver and patient accounts. A detailed description of the tasks can be found in deliverable D7.2A.

mə'mento keeps my mind

Additionally, the participants were asked about their preferences regarding orientation and navigation when lost outside, which is another use case that will be implemented for field trials. This information is referred to as “Lost Outside” in the following evaluation.



Figure 1 – Use cases that build the basis of tasks for lab trials

3.2.3 Evaluation Tools

The main objective of this phase was testing the device with real users during the prototype development phase using a Living Usability Labs methodology (Dias et al., 2015). It is important to collect feedback on user engagement by observation of the participants and collect feedback on usability using an established questionnaire.

As described in deliverable D7.1 and D7.2A, data was collected through direct observation during task performance and questionnaires.

From this, an observation form was specifically developed to collect measures such as task execution time, task completion rate (and how easily the participant completed the task), assistances during task completion, and the participant's visible emotional state. The results of these observations are described for each patient and caregiver group (referred to as "users") in detail in chapter 4.1.

The SUS was administered after completion of the tasks.

Evaluation tools are summarized in Table 5.

Table 5: Evaluation Tools	
Observation	Observation of users and caregivers and taking notes of their thinking aloud.
Task completion	Percentage of users who complete the task
	Amount of time required by users to complete the task.
Questionnaire	SUS (Brooke 1996)

4 Results

4.1 Observation forms

During lab trials, notes were taken of problems during task performance, engagement of the participants, suggestions and comments of the patients (P) and caregivers (C) thinking out loud and whether support was needed. Information for each user pair and each task is listed below. Additionally, the preferences regarding orientation and navigation when lost outside are included in 4.1.7.

4.1.1 General observation and comments

Users 1	Both P and C appreciated design and system in general. They were engaged and positively affected.
Users 2	P appreciated the “welcome” message on the main device. C suggested the possibility to use only vocal command to manage the main device. Both of them were engaged and positively affected.
Users 3	They put in evidence their unfamiliarity and inexperience with technology. P was positively engaged with the main device while appearing to be extremely frustrated by the all-day device interaction.
Users 4	They declared their inexperience with technology. The C is interested and expressed curiosity. The P was positively engaged.
Users 5	They seemed to be engaged with the system and expressed signals of pleasure. Both of them defined the system as simple and intuitive. In particular, they appreciated that the system includes reduced functionalities. The users judged the C interface on the computer as obsolete and suggested a mobile app.
Users 6	Users were positively affected by the all-day device in general. They liked the font sizes and appreciated the good writing recognition on the main device.
Users 7	They stated that the main device is very nice because of its clear layout. Additionally, handwriting on the main device was appreciated. The Users were also engaged with the design of the all-day device.
Users 8	They liked the overall design of both main and all-day device.

mə'mento keeps my mind

	<p>They thought that the symbols and font sizes on the main device are good, whereas the icons on the all-day device are not intuitive.</p> <p>The C interface was not received well.</p> <p>They stated that a “home” button on the main device would be nice.</p>
Users 9	<p>The users were very critical and seemed to be not overly engaged with the system, but were interested and expressed curiosity.</p>
Users 10	<p>They appreciated the all-day device, but expressed concern about the price of the main device (two tablets). They asked about advantages compared to existing devices.</p>
Users 11	<p>C appreciated design and system in general. They were engaged and positively affected. P found the main device a bit too heavy. Both P and C got frustrated because of the impossibility to use the all-day device.</p>
Users 12	<p>They put in evidence their unfamiliarity and inexperience with technology. Both P and C were positively engaged with the main device while appearing extremely frustrated by the all-day device interaction. P found the main device too heavy.</p>
Users 13	<p>They put in evidence their unfamiliarity and inexperience with technology. C points out that there should be more time before the screen of the all-day device goes off after touching it. C suggested the possibility to use vocal commands to manage the main device. Both of them were engaged and positively affected.</p>
Users 14	<p>They declared their inexperience with technology. The C is interested and expressed curiosity. The P was positively engaged but seemed to express some anxiety. She was afraid to confuse the icons.</p>
Users 15	<p>They seemed engaged but expressed signals of rejection with the technology. Both P and C defined the system as complex and hard to understand. Operating the C interface on a computer was judged to be too complex. After performing medication task, the P was so frustrated and annoyed that he suddenly stood up and left the room. The remaining tasks were not performed.</p>

4.1.2 General task

Users 1	<p>Task was completed.</p> <p>Support was given for the insertion of a special character (@ in the mail) and the time zone.</p>
Users 2	<p>Task was completed.</p> <p>Support was necessary to explain some field content and for time zone insertion.</p>

mə'mento keeps my minc

Users 3	Task was not proposed. They had never used a computer.
Users 4	Task was not proposed. They had never used a computer.
Users 5	Task was completed. Support was necessary to explain some field content.
Users 6	Task was completed. Support was given to log in on all-day device due to small font size.
Users 7	Task was completed. Support was given to log in on all-day device due to small font size.
Users 8	Task was completed. Support was given to log in on all-day device due to small font size and to explain some field content.
Users 9	Task was completed. Support was given due to technical problems Users asked if it is possible to log into the all-day device via computer.
Users 10	Task was completed. Support was given due to technical problems
Users 11	Task was not completed. Support was given for the insertion of the time zone. For P and C, it was impossible to log into the all-day device. Support was given but it was impossible to log in for the researcher too and the P and C had to leave the activity without performing any other task.
Users 12	Task was not proposed. The researcher created the accounts to assure the proper synchronization between devices.
Users 13	Task was completed. Support was necessary to explain some field content and for time zone insertion, also to log into the all-day device.
Users 14	Task was not proposed. The researcher created the accounts to assure the proper synchronization between devices.
Users 15	Task was not proposed. They had never used a computer but the C operates a smartphone daily.

4.1.3 Medication

Users 1	Tasks were completed. Support was given for return to "home page". There were some problems with handwriting and difficulties to insert the time in the format hh:mm.
Users 2	Tasks were completed. P was insecure and it was necessary to

mə'mento keeps my mind

	<p>guide her step by step. Anyway, she explained that this is her usual manner to approach novelty. She gave some examples of changes between traditional and new methods to approach memory problems (e.g. she recently changed completely the system to remind her of pharmacological therapy). She recently learned to use google assistance on the smartphone.</p> <p>Her C needed support to comprehend some fields of the C interface.</p>
Users 3	<p>C task was not proposed. P task was completed without problems.</p>
Users 4	<p>C task was not proposed. P task was completed. No problems with writing recognition despite use of italic. Sometimes P failed to return to "home page". There was confusion between the icons "list" and "new insertion". The medicine reminder was correctly displayed on the all- day device but not on the main device.</p>
Users 5	<p>Tasks were completed. P was disappointed that he cannot add a medication and C commented that she doesn't like to "control" him. Anyway, they commented that it is an important function in case of more severe cognitive problems. Intuitively, C guessed that the insertion of data in the C interface causes an alert on the devices. At the same time, P guessed that a change will appear in the medicine list and checked this.</p>
Users 6	<p>Task was completed, minor support was needed. There were problems with inserting times on the main device (format hh:mm). They suggested to set some markers, such as "insert hours" or insert minutes". Checkmarks on the C interface after choosing dates/times are not intuitive. Additionally, they stated that the font size is very small and the design is confusing.</p>
Users 7	<p>Task was completed, minor support was needed. Users stated that the font size of the C interface is very small. They asked if it is necessary to confirm reminders on both devices and suggested that the confirmed medications should appear on the device with a checkmark.</p>
Users 8	<p>Task was completed, major support was needed. P was insecure and it was necessary to guide him step by step. He had general problems with selecting icons on the main device (pressing too long/too hard,...) and concentrate on the task. The C was insecure on the computer as well and needed guidance. They asked if there are also notifications on the C interface.</p>
Users 9	<p>Task was completed, no support was needed. Writing recognition was working well, but simply choosing a day on a calendar would be preferred.</p>

mə'mento keeps my minc

	Additionally, they stated that showing the weekday would be good. They pointed out that there is too much information on the C interface. A reminder for new prescription/meds would be nice
Users 10	Task was completed, minor support was needed. There were problems with inserting times on the main device (format hh:mm). They think that instead of exact times, it would be better to choose between “morning”, “noon” and “evening”. They suggested, that information about the medicine would be nice (e.g. “Aspirin helps against pain”) The users voiced their concerns about more than one medication at the same time point. They also stated that the checkmarks after choosing dates/times on the C interface are not intuitive.
Users 11	Tasks were not performed.
Users 12	C task was not proposed. P task was completed with assistance due to difficulties to read on main device. Support was given for return to “home page”. Some problems with writing recognition and difficulties to insert the time in the format hh:mm. Both P and C were frustrated with the all-day device. C remarks, that the volume of the all-day devices sound alarm is a bit too low.
Users 13	Tasks were completed. P was insecure and it was necessary to guide her step by step. She doubts that the systems are able to add a reminder. Support was given to understand the all-day device. The P was fine with the weight of main device. She strongly pushed the tablet screen (like pushing a button) instead of touching it so the tablet did not recognize the action. The P wrote and set the hour by herself without problems. Support was given to the C to comprehend some fields of the C interface.
Users 14	C task was not proposed. P task was completed with assistance. No problems with writing recognition. Support was given due to P failing to return to “home page”. Confusion between icons: list and new insertion. The all-day device reminder did not work. Both P and C were frustrated with the all-day device.
Users 15	Tasks were not completed. C was disappointed that she couldn't write the hour on the screen because the systems response is too fast. P commented that he doesn't like the system at all. C commented that she already uses the smartphone alarm to remind her to take her medication.

4.1.4 Appointments

Users 1	Tasks were completed. Support was given to return to “home page” and to click “done”, at the end of any step. There were
---------	--

mə'mento keeps my minc

	problems with month data insertion (P inserted month name instead of number). Additionally, there were problems with comprehension of some fields of the C interface. Technical problems appeared due to internal server errors.
Users 2	Tasks were completed. Support was given to return to “home page” and to click “done”, at the end of any step. There were problems with month data insertion (P inserted month name instead of number). Technical problems due to internal server errors and with alerts and writing recognition. P had difficulties with reading (size of characters). Interpretation of icons was not immediately. Problems with comprehension of some fields of the C interface.
Users 3	C task was not proposed. P task was completed with difficulties due to writing recognition and interruptions (alert activation that makes it necessary to restart the procedure).
Users 4	C task was not proposed. P task was completed without difficulties. Support was given to insert time in hh:mm format. The system recognized the italic handwriting also in this case. The appointment reminder presented an incorrect date.
Users 5	Tasks were completed. The writing recognition required too long and the P seemed to be annoyed to rewrite without the possibility to correct. Problems with comprehension of some fields of the C interface, especially “start date” and “repetition”. In particular it is not clear that start date and repetition concerns the alert set up. It could be replaced with “Set-up Alert date and time” and “Alert Repetition”. Alerts functioned correctly. An alert was funny for grammatically incorrectness (“hai un Titta in 30 minuti” instead of “hai un appuntamento con Titta tra 30 minuti”) so the words “un appuntamento con” were removed.
Users 6	Task was completed, minor support was needed. Problems with comprehension of some fields of the C interface.
Users 7	Task was completed, minor support was needed. Problems with comprehension of some fields of the C interface. The term “start date” is confusing; it could be replaced with “reminder date”. If there is an error on the C interface, all data gets deleted and makes it necessary to fill in the information again. Users asked, if it is possible to see appointments for the next weeks on the main device.
Users 8	Task was completed, major support was needed. Handwriting was a key issue. The P had shaky hands when writing more slowly; therefore he tried to stabilize his hand by

mə'mento keeps my minc

	<p>putting it down on the device. This lead to unwanted actions on the tablet and was very frustrating for the user.</p> <p>Support was given to return to “home page” and to click “done”, at the end of any step.</p>
Users 9	<p>Task was completed, no support was needed.</p> <p>The term “start date” on the C interface is confusing; it could be replaced with “reminder date”. The C was very annoyed using the C interface. He also stated that the checkmarks after choosing dates/times are not intuitive.</p>
Users 10	<p>Task was completed, minor support was needed.</p> <p>Information about weekdays would be great (e.g. appointment next Monday).</p> <p>The term “start date” on the C interface is confusing; it could be replaced with “reminder date”.</p> <p>Support was given to return to “home page” and to click “done”, at the end of any step.</p>
Users 11	<p>Tasks were not performed.</p>
Users 12	<p>C task was not proposed. P task was completed with difficulties and assistance due to problems with writing recognition and interruptions. The tablet digitalizes handwriting too fast for the Ps writing speed and sometimes gets stuck. C writes the month instead of the date number.</p>
Users 13	<p>Tasks were completed. Support was given to return to “home page” and to click “done”, at the end of any step. P had difficulties with icons interpretation and misunderstood “Citas” with “having a date” with some boyfriend. The P suggested to change the word “Citas” for “Agenda”. Problems with month data insertion (P inserted month name instead of number). Technical problems with writing recognition for both P and C. Problems with comprehension of some fields of the C interface. The C suggests to change the word “Editar” to “Añadir nueva actividad/entrada” (Add a new activity/entrance).</p>
Users 14	<p>C task was not proposed. P tasks were completed with writing assistance. Support was given to return to “home page” and to click “done”, at the end of any step. P had difficulties with icons interpretation and misunderstood “Citas” with “having a date”.</p>
Users 15	<p>Tasks were not performed.</p>

4.1.5 Shopping

Users 1	<p>Tasks were completed. Support was given to return to “home page” and to click “done”, at the end of any step. Technical problems with alerts. All devices were appreciated for touch function and reading list function.</p>
---------	---

mə'mento keeps my minc

Users 2	Tasks were completed. From P side, there were problems with icons recognition and the size of characters (all-day device). No problems to listen the vocal message, which was appreciated. Initial difficulties to scroll down the list.
Users 3	Task with all-day device was completed with assistance. No problems with main device. However, the P was frustrated and the task was repeated two times in order to encourage him. Main problems concerned size of characters, scroll down list and touch function. No problems with icon recognition.
Users 4	Tasks were completed. Difficulties with all-day device due to size of characters and scrolling down the list.
Users 5	Tasks were completed. Difficulties due to confusion about including or excluding items when ticking them off on the main device. Users suggested a list of 20 typical items plus the possibility to add some items. The sign of a "tick" would be assigned to "include" items to the list shown on the all-day device. It is not clear if the list is completed after ticking off all items because no message (written or vocal) appears.
Users 6	Task was completed, minor support was needed. They stated that it would be better, if the all-day device reads only the next item on the list. They also mentioned that the reading is too fast. Difficulties due to confusion about including or excluding items when ticking them off on the main device. Also, P was confused about where to press on the main device (on the item itself or the field in front). The icons on all-day device are thought to be confusing. Going back on the all-day device to choose another list was complicated.
Users 7	Task was completed, minor support was needed. They thought that the lists are great in general. Users asked, what happens if I have more than one item of a sort (e.g. 3 apples). Users suggested a list of pre-selected, frequently used items.
Users 8	Task was completed, major support was needed. Difficulties due to confusion about including or excluding items when ticking them off on the main device. Also, P was confused about where to press on the main device (on the item itself or the field in front).
Users 9	Task was completed, no support was needed Users asked, what happens if I have more than one item of a sort (e.g. 3 apples). The C suggested an extra field for adding numbers. They stated that the technical voice is not very nice and that

mə'mento keeps my mind

	the reading is too fast.
Users 10	Task was completed, minor support was needed Difficulties due to confusion about including or excluding items when ticking them off on the main device. Users suggested a list of pre-selected, frequently used items. They stated that the font size on the all-day device is too small. They also think that blending out items that are already ticked off would be better. Going back on the all-day device to choose another list was complicated.
Users 11	Tasks were not performed.
Users 12	Tasks were completed. P had problems with icon recognition and the size of characters of the all-day device. The reading of the list was appreciated. Difficulties to scroll down the list. The function of the voice reading the list was perceived wrong - confusion because of the voice reading the unticked objects on the list. Reading was too fast at the first time to understand for the P.
Users 13	Task with all-day device was completed with assistance. No problems with main device. Main problems concerned size of characters, scroll down list and touch. No problems in icons recognition. The function of the voice reading the list was perceived wrong. Confusion due to the voice reading the unticked objects of the list. Reading was too fast at the first time to understand for the P. The P was frustrated. Both P and C suggested that it would be better if the voice reads the ticked objects on the list and more slowly.
Users 14	Tasks were completed. Difficulties with ticking main device list and with all-day device due to size of characters and scrolling down the list. The function of the voice reading the list was perceived wrong - confusion because of the voice reading the unticked objects on the list. Reading was too fast at the first time to understand for the P. The P was frustrated.
Users 15	Tasks were not performed.

4.1.6 Getting ready

Users 1	Tasks were completed. Support was given for return to "menu" and for scroll done the list. Icons were not clear for the P.
Users 2	Tasks were completed. Sometimes appeared signals of anxiety

mə'mento keeps my minc

	but she was engaged and curious. She was worried to forget the procedures of the system.
Users 3	Task with all day device was completed with assistance and time was not considered. No problems with main device. Main problems are size of characters, scroll down list and touch. No problems in icons recognition. P was frustrated. From technical side: an alert appears with an incorrect time.
Users 4	Tasks were completed. Difficulties due to confusion about including or excluding items when ticking them off on the main device and due to the size of characters. When the list is completed no message (written or vocal) appears.
Users 5	Tasks were completed without problems. It was not clear when the list was completed because no message (written or vocal) appears.
Users 6	Task was completed, minor support was needed. Difficulties due to confusion about including or excluding items when ticking them off on the main device. Going back on the all-day device to choose another list is complicated.
Users 7	Task was completed, minor support was needed. Going back on the all-day device to choose another list was complicated.
Users 8	Task was completed, major support was needed. Difficulties due to confusion about including or excluding items when ticking them off on the main device. Difficulties also with scrolling and font size.
Users 9	Task was completed, no support was needed.
Users 10	Task was completed, minor support was needed. Difficulties due to confusion about including or excluding items when ticking them off on the main device. Going back on the all-day device to choose another list was complicated.
Users 11	Tasks were not performed.
Users 12	Tasks were completed with assistance. Support was given to return to "menu" and to scroll down the list. Icons were not clear for the P and there were difficulties to understand the all-day device list. The function of the voice reading the list was perceived wrong - confusion because of the voice reading the unticked objects on the list. Reading was too fast at the first time to understand for the P. The P was frustrated. Both P and C suggested that it would be better if the voice reads the ticked objects on the list and more slowly.
Users 13	Task with all-day device was completed with assistance. No problems with main device. Main problems are size of characters, scroll down list and touch. No problems in icons recognition. P was frustrated. Too many steps to get into the

mə'mento keeps my mind

	list on all-day device (too small for both P and C). The function of the voice reading the list was perceived wrong - confusion because of the voice reading the unticked objects on the list. Reading was too fast at the first time to understand for the P. The P was frustrated. Both P and C suggested that it would be better if the voice reads the ticked objects on the list and more slowly.
Users 14	Tasks were completed. Difficulties due to confusion about including or excluding items when ticking them off on the main device and due to size of characters. The function of the voice reading the list was perceived wrong - confusion because of the voice reading the unticked objects on the list. Reading was too fast at the first time to understand for the P. The P was frustrated.
Users 15	Tasks were not performed.

4.1.7 Lost outside

Users 1	To call emergency numbers function was appreciated. The possibility to receive a message with the closest familiar place was preferred to a message with the address.
Users 2	To call emergency numbers function was appreciated. The possibility to receive a message with the closest familiar place together with the address was suggested. The C was concerned that the emergency call is activated inadvertently or too often.
Users 3	To call emergency numbers function was appreciated. The possibility to receive a message with the closest familiar place together with the address was suggested.
Users 4	To call emergency numbers function was appreciated. The possibility to receive a message with the closest familiar place was preferred to a message with the address.
Users 5	To call emergency numbers function was appreciated along with a message sent to the C including the P position (coordinates). The possibility to receive a message with the closest familiar place together with the address was suggested. Also a route to come back to the starting position it could be nice.
Users 6	To call emergency numbers function was appreciated. The users preferred GPS coordinates/maps, as well as information on the name of the street and a point of interest (e.g. the "bakery around the corner").
Users 7	To call emergency numbers function was appreciated. The users preferred GPS coordinates/maps, as well as information

mə'mento
keeps my mind

	on a point of interest (e.g. the “bakery around the corner”).
Users 8	To call emergency numbers function was appreciated. The users preferred GPS coordinates/maps, as well as information on a point of interest (e.g. the “bakery around the corner”).
Users 9	To call emergency numbers function was appreciated. The users displayed concerns, that sights as well as points of interest could be confusing. Also, they pointed out that there is a lot of data to type in when setting up the points of interests.
Users 10	To call emergency numbers function was appreciated. The users preferred GPS coordinates/maps, as well as information on a point of interest (e.g. the “bakery around the corner”). In the users opinion, the optimal solution would be a GPS system leading the way home.

4.2 Task completion time

Time was taken during the performance of each task. Data is missing for users where tasks were not proposed. Results of task completion time is given in seconds in Table 6: Task completion time.

Table 6: Task completion time																
	Users 1	Users 2	Users 3	Users 4	Users 5	Users 6	Users 7	Users 8	Users 9	Users 10	Users 11	Users 12	Users 13	Users 14	Users 15	∅
General Task	321	222	-	-	100	183	124	310	-	192	-	-	294	-	-	218
Medication																
P task	40	133	49	89	48	71	53	99	40	25	-	341	165	261	132	110
C task	20	80	-	-	60	140	186	200	30	96	-	-	255	-	236	130
Appointments																
P task	115	380	366	150	235	188	31	-	64	125	-	215	217	141	-	186
C task	170	220	-	-	179	151	-	-	110	85	-	-	212	-	-	161
Shopping																
Task 1	10	110	15	79	4	104	14	120	10	50	-	95	78	86	-	60
Task 2	50	135	265	180	80	55	15	240	16	164	-	249	244	222	-	147
Getting ready																
Task 1	40	56	60	95	30	50	16	65	12	20	-	74	51	59	-	48
Task 2	79	230	-	136	40	72	-	153	33	50	-	254	248	186	-	135

4.3 SUS Questionnaire

Results of the SUS questionnaire are listed in detail in Table 7: System Usability Scale. The participants scored 10 questions (Q1-10) with one of five responses that range from strongly agree (5) to strongly disagree (1).

A final score was calculated from these answers.

Patient 15 did not fill in the questionnaire, patient 12 only partially.

- **Q1 - I think that I would like to use this system frequently.**

Patients rated the first question with 4 out of 5 points. 5 out of 14 patients scored the question with 5 points and therefore strongly agree to use the system frequently, 1 patient strongly disagrees. The caregivers rated the question with an average of 3 points.

- **Q2 - I found the system unnecessarily complex.**

Most of the patients didn't think of MEMENTO as unnecessarily complex and disagreed with question 2. Again, caregivers scored the question with an average of 3 points.

- **Q3 - I thought the system was easy to use.**

This question was rated with 4 points from both patients and caregivers. None of the patients and one caregiver strongly disagreed.

- **Q4 - I think that I would need the support of a technical person to be able to use this system.**

Both patients and caregivers rated the question with an average of 3 points.

- **Q5 - I found the various functions in this system were well integrated.**

Both patients and caregivers rated the integration of the functions with 4 points.

- **Q6 - I thought there was too much inconsistency in this system.**

Both patients and caregivers rated the question with 3 points.

- **Q7 - I would imagine that most people would learn to use this system very quickly.**

mə'mento keeps my mind

Patients agreed with 4 points with this statement. Caregivers rated the question with three points.

- **Q8 - I found the system very cumbersome to use.**

Both patients and caregivers rated the questions with 3 points.

- **Q9 - I felt very confident using the system.**

Most patients and caregivers felt confident using the system and rated the question with 4 points. 5 out of 14 patients strongly agreed, as well as 4 out of 15 caregivers, whereas none of the users strongly disagreed.

- **Q10 - I needed to learn a lot of things before I could get going with this system.**

Question 10 was rated with an average of 3 points from the users.

Table 7: System Usability Scale											
SUS Patients											
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Score
P1	3	2	3	3	4	3	3	4	3	4	50
P2	4	3	4	5	4	3	3	1	4	5	55
P3	3	1	3	4	5	4	3	2	5	3	63
P4	5	1	5	3	5	4	4	2	4	3	75
P5	4	2	5	1	5	1	5	1	4	1	93
P6	2	1	4	2	4	4	4	2	4	2	68
P7	5	1	4	2	5	1	3	2	5	1	88
P8	3	4	4	3	4	4	4	3	4	4	53
P9	5	3	4	4	3	4	3	3	4	4	53
P10	5	1	3	1	5	1	5	1	5	1	95
P11	3	3	4	3	5	3	3	5	3	4	50
P12	1	5	4	5	2	-	-	5	5	-	-
P13	4	2	3	4	3	3	4	2	5	1	68
P14	5	1	5	3	5	4	4	2	4	3	75
P15	-	-	-	-	-	-	-	-	-	-	-
Mean	4	2	4	3	4	3	4	3	4	3	68

keeps my mind mə'mento

mə'mento
keeps my mind

SUS Caregivers											
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Score
C1	5	5	5	2	4	4	5	5	5	1	68
C2	4	1	5	4	4	1	4	2	4	1	80
C3	3	3	4	3	2	4	5	3	4	4	53
C4	5	2	4	4	5	1	4	1	4	2	80
C5	2	1	5	1	5	1	3	1	5	1	88
C6	1	1	3	3	3	2	3	2	3	2	58
C7	4	3	4	5	5	4	3	3	3	3	53
C8	4	4	1	4	4	4	2	4	4	4	38
C9	5	2	3	2	4	5	1	4	3	1	55
C10	2	3	2	1	3	4	2	4	5	1	53
C11	3	5	5	3	3	4	4	5	3	5	40
C12	3	4	4	5	4	3	4	2	5	5	53
C13	3	3	4	3	4	4	4	2	5	1	68
C14	3	4	4	4	3	3	2	1	4	2	55
C15	1	5	5	5	2	5	3	5	2	5	20
Mean	3	3	4	3	4	3	3	3	4	3	57

5 Main Results and Suggestions

5.1 SUS Report

The mean score of the individual questions provided information about different aspects of usability.

Summarized, most of the patients stated that they would use the system frequently and wouldn't describe the MEMENTO system as unnecessary complex, whereas the caregivers rated the question neutrally. Both think the system is easy to use, but for some of the users, technical support might be needed.

In systems where functionality and interactions are distributed across more than one device, surveys about a coherent user experience across the devices with which the user interacts are important. The system integration, hence the connection of the different components of MEMENTO into a single larger system that functions as one, was thought to be good. The consistency of the system, meaning if the cross-device system feels coherent, was rated as average.

Patients think that most people would learn to use this system very quickly and the users felt confident using the system. Due to different technical proficiencies, some patients might need some time to learn and practice before using the system, whereas others could get going fast.

Calculating and interpreting the SUS score is complex and a detailed explanation is given by Bangor and colleagues (Bangor, Kortum et al. 2009).

The SUS underlines, that there is great variability between users, mostly due to their different age and technical proficiency and interest.

The score ranges between 50 and 95 from patients' side of view and between 20 and 88 according to the caregivers (out of 100), shown in Figure 2 – SUS end-users and caregivers.

The raw SUS scores can be converted into percentile ranks, indicating how well the raw score compares to others in the database. Table 8: SUS Interpretation shows the percentile ranks for SUS thresholds.

The average score (at the 50th percentile) of the SUS is 68. The exact mean score of our patient group was 67.9 (SD=16.2), whereas the caregivers rated the usability with a mean of 57.2 (SD=17.6).

The usability of the first MEMENTO prototype is therefore ranked between “OK” and “Good” compared to other systems. Overall, end-users rated the questionnaire better than the caregivers.

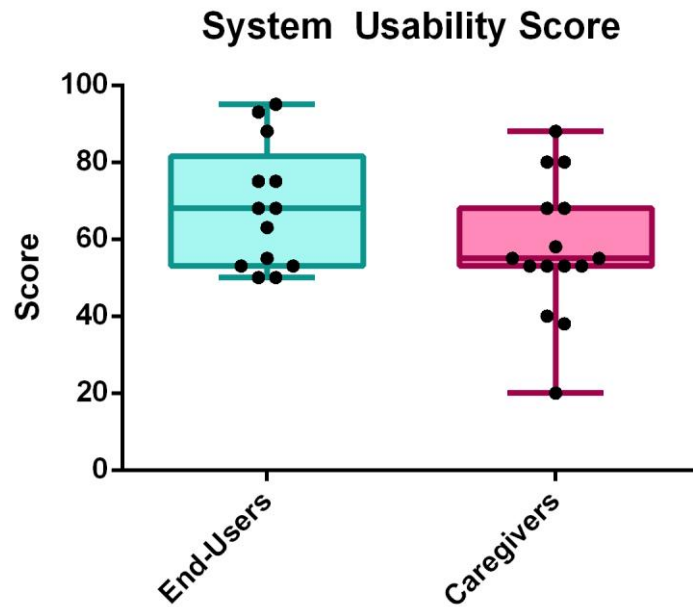


Figure 2 – SUS end-users and caregivers
The boxplot shows the mean value, and extends from the 25th to 75th percentile. Whiskers indicate min. to max. values.

Grade	SUS	Percentile	Adjective
A+	84.1-100	96-100	Best Imaginable
A	80.8-84.0	90-95	Excellent
A-	78.9-80.7	85-89	
B+	77.2-78.8	80-84	
B	74.1 – 77.1	70 – 79	
B-	72.6 – 74.0	65 – 69	
C+	71.1 – 72.5	60 – 64	Good
C	65.0 – 71.0	41 – 59	
C-	62.7 – 64.9	35 – 40	
D	51.7 – 62.6	15 – 34	OK
F	25.1 – 51.6	2– 14	Poor
F	0-25	0-1.9	Worst Imaginable

5.2 Task Completion

All but one of the tasks proposed to the users were completed, with different levels of support. One task was not completed due to difficulties with handwriting, whereas other tasks were not proposed to the users. Apart from one patient cancelling the trial, this was due to investigators sensing that participants started to get tired or annoyed during testing or to prevent an excessive frustration (e.g. the general task was not proposed to users who never been used a computer).

Points where support was needed regarding problems from end-user side are further discussed in 4.5.

The mean task completion time ranged from 48s for simple tasks (open and checking lists) to 186s for more complex tasks, such as scheduling appointments. Most time was spent on creating an account and logging into the devices (218s). Taking time for this general task was cancelled in some cases or was taken over by the investigators, due to problems from end-user side and low technical proficiency of the users.

This task needs to be done only once when receiving the MEMENTO system.

Generally, the task completion times were fine considering the users using the devices for the first time and are expected to improve when getting used to the system. The data can be used as reference after further refining the system for upcoming field trials.

5.3 User Experience

Most of the users were very engaged with the system and its design in general. They liked the idea of a personalized system, including a photography on the left tablet and the individual greeting on the right tablet. The patients and their caregivers expressed curiosity for the general concept and the possibilities of the system. Most patients appreciated the handwriting recognition when scheduling new appointments and medication and enjoyed the all-day device function to let the list be read aloud.

Frustration was expressed especially when technical problems occurred, which emphasizes the importance of stability of the system.

5.4 Target group related challenges

Due to our specific end-user group, a special focus during lab trials was on the challenges related to their needs and problems.

Regarding the main device, problems occurred when inserting the date during scheduling appointments. Patients were confused where to write the hours and minutes and whether to insert the month's name (e.g. "June") or number of the month (e.g. "6"). They mentioned that labelling of the respective fields would be very helpful.

Also, it was not intuitive for some patients to confirm their insertion by clicking done at the end of each step during the tasks. Especially the end-users with low technical proficiency had problems to go back to the starting page to start a new action. Patients and caregivers stated, that they would appreciate that the device returns back to the starting page after the action is completed.

For some of the patients, handwriting was challenging due to shaking hands. For the device to recognize handwriting properly, a relatively even speed of writing is required. Digitalizing the handwriting during slow writing before finishing the word caused frustration in those patients, particularly

because correction is not possible and the data had to be deleted and inserted again. Those patients would prefer inserting the dates by selecting them on a list or calendar.

Additionally, the caregivers needed a lot of assistance during inserting data on the caregiver interface. Most of the elderly caregivers had problems comprehending the meanings of the input fields and navigating through the interface. They would also appreciate an extra button to go back to the patient overview page.

Almost all end-users and caregivers misunderstood the “ticking off” of items of the shopping and getting ready lists. Intuitively, they thought a checkmark on the main device is inclusive (as in “I have to buy/pack this item”), while on the all-day device it is exclusive (“I already bought/packed this item”). This confusion was probably also strengthened by the already generated demo lists during lab trials. However, the users were in favour of prepared general lists with common items (such as butter and milk for shopping), from which to choose items and with a possibility to add additional items manually.

The function to let the lists be read aloud by all-day device was perceived very well from both patients and caregivers. For some of the users, the speed of reading was too fast and going back to choose another list or start another action was too complicated. Scrolling through the list on the all-day device was also a problem for users with low technical proficiency. Additionally, one of the caregivers stated, that it would be nice if there is a message when all items are ticked off and the list is completed, such as a pop-up or a vocal message.

For almost all users, the log in had to be performed from investigator side due to the small screen of the watch.

In general, a very important statement concerned the insertion of more than one medicine at the same time point, which was not possible during lab trials but is crucial for our end-user group during the field trials. Some of the patients would also prefer a categorization of medicine into “morning”, “noon” and “evening”. Also, a patient stated that she would appreciate short information about the effects of the medicine.

5.5 Technical problems

During the lab trials, technical problems occurred. Many of the points stated below were probably due to bad Wi-Fi connection, which underlines the importance of a stable internet access.

After setting reminders, they were sometimes not shown on the main device or at a wrong time point. During two trials, there was a problem with a reminder showing up repeatedly every minute. Additionally, users were annoyed to start a task from the beginning after a reminder popped up in-between.

In some cases, the screens froze, making technical support necessary. This is especially unfavourable in case the left side containing the clock freezes, causing further disorientation for the patients in everyday life.

There were also problems with handwriting in the Austrian trials due to failing recognition of umlauts.

Also, there was sometimes slow synchronization between main and all-day device. Otherwise, there were no technical problems regarding the all-day device.

As for the caregiver interface, there was in one trial a major problem with registration. After correctly entering the data, the user profile was created in the background but not logged in, showing an error message.

5.6 Design aspects

Regarding the design, the caregiver interface was not perceived well. The interface was too complicated for the mostly elderly caregivers. They stated, that there is too much information, which makes the site confusing and intimidating.

The overall design of the main device was appreciated. There was confusion about which button to use to add new appointments and medication.

The participants emphasized, that personalization of the device is crucial to them, such as adding personalized photographs on the left tablet. There is also need to attract more attention to the different headlines.

A patient also stated that it would be nice to have a weather forecast on the left tablet. When inserting dates and looking them up, showing the weekday would be also appreciated for better orientation. A replacement pen in case one is lost would be good.

The design of the all-day device was also considered to be very good. One patient commented that he doesn't like the plastic straps of the prototype. Recognizing some of the icons (panic button, lists) on the all-day device was difficult for some of the patients and the font size was too small for patients having problems with vision.

6 Conclusions

In this deliverable, the usability and the acceptability of the first prototype and engagement of the end-users and their caregivers was evaluated.

Overall, the MEMENTO system was well received. The general concept and design were appreciated by the participants of the lab trials and they showed great interest.

Regarding the usability, several valuable suggestions for improvement were pointed out by the users and technical, as well as design problems were detected. The findings of the lab trials will be useful for further development of the MEMENTO system to provide a solution that reliably supports our target group, also in the aspects of engagement, effectiveness, efficiency, and overall ease of use.

7 References

Bangor, A., P. Kortum and J. Miller (2009). "Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale." Journal of Usability Studies **4**(3): 114-123.

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. Usability evaluation in industry. London, Jordan PW, Thomas B, Weerdmeester BA, McClelland AL.

Dias, M. S., Vilar, E., Sousa, F., Vasconcelos, A., Pinto, F. M., Saldanha, N., & Eloy, S. (2015, August). A living labs approach for usability testing of Ambient Assisted Living technologies. In International Conference of Design, User Experience, and Usability(pp. 167-178). Springer, Cham.

McKhann, G. M., D. S. Knopman, H. Chertkow, B. T. Hyman, C. R. Jack, Jr., C. H. Kawas, W. E. Klunk, W. J. Koroshetz, J. J. Manly, R. Mayeux, R. C. Mohs, J. C. Morris, M. N. Rossor, P. Scheltens, M. C. Carrillo, B. Thies, S. Weintraub and C. H. Phelps (2011). "The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." Alzheimers Dement **7**(3): 263-269.

Nucci, M., D. Mapelli and S. Mondini (2012). "Cognitive Reserve Index questionnaire (CRIq): a new instrument for measuring cognitive reserve." Aging Clin Exp Res **24**(3): 218-226.

Partners



blkm



CiTARD Services Ltd



VIRTUALVARE

WETOUCH

Supporting Organizations

