RESEARCH PROTOCOL

Primary care diagnostics or diagnostics in a memory clinic in older persons with memory complaints –
A long-term cost-effectiveness trial with non-inferiority design

Protocol ID	CMO dossiernummer NL83486.091.22 Panama dossiernummer 114143
Short title	PRIMED
Version	2.4
Date	18-11-2024
Coordinating investigator/project leader	**Drs. D. Ronner** *Coördinerend onderzoeker, AIOTO (arts in opleiding tot onderzoeker en specialist ouderengeneeskunde)* Radboudumc Eerstelijnsgeneeskunde Geert Grooteplein Zuid 10, 6525GA Nijmegen E-mail: demi.ronner@radboudumc.nl **Dr. M. Perry** *Senior Onderzoeker, Huisarts* Radboudumc Eerstelijnsgeneeskunde, Geriatrie Geert Grooteplein Zuid 10, 6525GA Nijmegen E-mail: Marieke.Perry@radboudumc.nl Huisartsenpraktijk Velp Rozendaalselaan 34, 6881LD Velp **Prof. dr. E. Richard** *Neuroloog* Radboudumc Neurologie Geert Grooteplein Zuid 10, 6525GA Nijmegen E-mail: Edo.Richard@radboudumc.nl **Prof. dr. H. Schers** *Hoogleraar huisartsgeneeskunde, Huisarts* Radboudumc Eerstelijnsgeneeskunde Geert Grooteplein Zuid 10, 6525GA Nijmegen E-mail: Henk.Schers@Radboudumc.nl
Principal investigator(s) (in Dutch: hoofdonderzoeker/ uitvoerder)	**Dr. M. Perry** *Senior Onderzoeker, Huisarts* Radboudumc Eerstelijnsgeneeskunde, Geriatrie Geert Grooteplein Zuid 10, 6525GA Nijmegen E-mail: Marieke.Perry@radboudumc.nl
Sponsor (in Dutch: verrichter/opdrachtgever)	Radboudumc
Subsidising party	ZonMW, programma Doelmatigheid
Independent expert	**Dr. Jurgen Claassen** *Klinisch geriater, senior onderzoeker* Radboudumc Geriatrie Geert Grooteplein Zuid 10, 6525GA Nijmegen E-mail: Jurgen.Claassen@Radboudumc.nl
Laboratory	Not applicable
Pharmacy	Not applicable

PROTOCOL SIGNATURE SHEET

Name	Signature	Date
Head of Department Primary and Community Care: *A. Timen*		15-12-2022 03-03-2023 21-04-2023 14-06-2023 18-11-2024
Principal Investigator: *M. Perry*		15-12-2022 03-03-2023 21-04-2023 14-06-2023 18-11-2024

DOCUMENT HISTORY

Document	Date of version	Summary of Changes
Protocol version no. 2.4	18-11-2024	- Addition of longitudinal prospective cohort - Extension of recruitment period
Protocol version no. 2.3	14-06-2023	- Removed Lawton - Added/changed minor details in Chapter 5 (Treatment), 6 (Methods), and 8 (Ethical considerations)
Protocol version no. 2.2	21-04-2023	- Removed section about additional training offered to participating GPs
Protocol version no. 2.1	03-03-2023	- Added document history - Correction of dates and version numbers - Added project coordinator
Protocol version no. 2.0	28-2-2023	- Added NFU risk classification - Adjusted text in 6.3.1
Original protocol 1.0	20-12-2022	Not applicable

Tabel of contents

Appendices 22

Summary

Rationale

The number of older persons with memory complaints rises. Dutch Dementia guidelines support diagnostics in primary care. General practitioners (GPs) are competent to observe and interpret changes in cognition and behavior on their patients' functioning. However, > 60% of 20.000 diagnoses per year in the Netherlands are made in a memory clinic (MC). An early MC diagnosis may be faster and more accurate on the short term. But how do these advantages outweigh the possible drawbacks of expensive and burdensome tests such as neuropsychological tests, MRI, lumbar puncture or PET scan, which may lead to over-diagnosis and incidental findings? GPs argue for a timely diagnosis, as with no effective treatment available, the urgency for an early diagnosis is limited and determined by patient and caregiver preferences.

Aim

Research question: What is the comparative efficacy and safety between dementia diagnostics in primary care and dementia diagnostics in a memory clinic for older persons with memory complaints?

Hypothesis: In absence of disease-modifying treatment for dementia, a diagnostic trajectory in primary care is not inferior to a memory clinic with respect to long-term outcomes relevant to patients and caregivers and generates less healthcare costs.

Study design

Mono-center randomized controlled non-inferiority trial. To be able to generalize study findings to persons with strong preferences to either primary care diagnostics or memory clinic diagnostic, older persons who do not want to be randomized are included in a longitudinal prospective cohort. To acknowledge the complexity of the issue under trial, the RCT is complemented with a mixed-methods process evaluation.

Study population: Patients > 70 years presenting to the GP with memory problems without signs of uncommon types of dementia or intracranial pathology. Patients are recruited at primary care practices and asked for written informed consent.

Study procedure: Intervention: dementia diagnostics in primary care. Comparison: dementia diagnostics in a memory clinic. Diagnostic procedures in both arms are left to the discretion of the diagnosing physician. During follow-up, all professionals involved are asked to perform care as usual.

Outcome measures (for RCT and cohort): primary outcome measure is daily functioning using the Amsterdam instrumental Activities of Daily Living questionnaire (A-iADL-Q-SV). Secondary outcomes are measured using validated questionnaires for quality of life, costs, cognition, behavior and mood, caregiver burden and perseverance time, anxiety or dissatisfaction after diagnostic trajectory. Secondary outcomes derived from primary care electronic data are acute admissions, time to institutionalization, time to mortality, accuracy of the initial diagnosis and data on diagnostic and follow-up processes. Outcomes are measured every 6 months up to 30 months.

Sample size/data analysis: with a minimal clinically important difference on the A-iADL-Q-SV of 6.3 points and a non-inferiority margin of 60% (α=0.025, β=0.8), we need 2 x 79 participants in the RCT, which will be increased to 2 x 91 to account for attrition. Primary analysis is performed according to the intention-to-treat principle, using linear regression analysis to compare A-iADL-Q-SV scores between both groups at (24-)30 months. A per protocol analysis is performed as participants from the primary care arm may be referred during follow-up. Cohort data are analyzed

separately using similar protocols as RCT data and outcomes are used to compensate the minimal loss of power that occurred by setting a non-inferiority margin of 60% instead of the more common 50%.

Cost-effectiveness analysis: we will perform a cost-effectiveness analysis according to the Dutch manual for costing research. Outcome measures are quality-adjusted life years (QALYs), based on combining the Euro-QOL 5D (EQ5D-5L) utility scores with survival. To determine robust confidence intervals surrounding the Incremental Cost-Effectiveness Ratio, we will use (non)parametric bootstrapping. Volumes of care are determined using the Resource Utilization in Dementia (RUD) questionnaire at baseline, and every 6 months during follow-up.

Process evaluation: To acknowledge the complexity of the issue under trial, the RCT is complemented with a mixed-methods process evaluation. We will draw on the fundamental ideas of the MRC guidance for process evaluations. We will consider factors related to the GP, the patient, the informal caregiver, the diagnostic and follow-up trajectories, collecting a combination of quantitative and qualitative process measures and patient/professional reported experience measures (PREMs).

Burden and risks to participants

Both diagnostic trajectories comprise forms of regular care. Risks in both study arms are low. NFU guidelines recommend to identify and register these risks in low-risk studies.
In the MC arm, the main risks are 1) complication of a diagnostic procedure (lumbar puncture is rarely performed and has a very low risk of serious complications such as subdural hematoma and meningitis) 2) incidental findings on brain CT or MRI, leading to anxiety or additional (diagnostic) procedures and treatments and 3) burden of referral.
In the primary care arm, the main consequence may be a delayed diagnosis. This can result from the fact that patient and/or GP decide together that diagnostic certainty is not yet desirable. On the other hand, it can be caused by lack of explicit decision making on the wish for diagnostic certainty, possibly resulting in a period of uncertainty for patient and caregiver. Appropriate counseling may mitigate the burden of this uncertainty. Most importantly, a potentially delayed diagnosis does not affect the prognosis, because no disease-modifying treatment exists. Chances of missing other intracranial pathology than a neurodegenerative disease as cause are very low.
Burden for participating patients is limited as questionnaires take less than one hour to complete per time point.
Burden for GPs includes recruiting patients, completing a short baseline questionnaire, answering three to five additional questions three times for every included patient, and uploading the primary care electronic medical records of all included patients at the end of follow-up. Participation in an additional interview for both patients and GPs is voluntarily.

1. Introduction and rationale

1.1 Problem definition

Ageing of our society causes the number of older persons with memory complaints to rise.[1, 2] Dementia is a clinical diagnosis and Dutch Dementia guidelines promote dementia diagnostics in primary care without ancillary investigations in all older patients without signs of uncommon forms of dementia (such as focal signs, concomitant Parkinsonism).[3, 4] General practitioners (GPs) are well-positioned to observe and interpret changes in cognition and behavior or their patients' daily functioning. In spite of this excellent position and the guideline recommendations, over 60% of the 20,000 new diagnoses per year in the Netherlands are made in memory clinics (MC).[5]

Medical specialists tend to promote an early dementia diagnosis to be made in an MC, as it is faster and may be more accurate than in primary care.[6] But how do these advantages outweigh the possible drawbacks of extensive investigations such as neuropsychological tests, MRI, lumbar puncture or PET scan? Hospital visits and such tests are burdensome for older patients and caregivers and may lead to over-diagnoses and incidental findings. The value of a more exact dementia-subtype diagnosis -made in a MC- may be limited for the vast majority of older persons with dementia, because of (1) the large overlap between Alzheimer pathology and cerebrovascular pathology underlying the clinical dementia syndrome, (2) the limited prognostic value of a more exact diagnosis, and (3) the lack of disease- modifying treatments that can slow the decline.

Without treatments available, the urgency for a diagnosis is determined by patient and caregiver preferences and sometimes a requirement for initiating care.[7] There is wide consensus that such a *timely* diagnosis is more important than an *early* diagnosis. GPs emphasize the importance of such a timely diagnosis, and are well-positioned to make a shared decision with individual patients on the optimal timing of diagnosis. This may however lead to diagnostic delay.[8]

It is unknown how pros and cons of both diagnostic approaches outweigh each other: which makes patients and caregivers fare better in the long-term and which is most cost-effective.
Randomized studies performed over 10 years ago showed that structured MC diagnostics improved health related QoL (HRQL) compared to diagnostics as usual at 6 months. This effect was not sustainable at 12 months. Usual diagnostics could include both primary care diagnostics and referral to 'normal' MCs, but data on diagnostic setting in the usual diagnostics group are not reported.[9] Follow-up care in MCs after specialist diagnosis did not lead to better QoL or decreased informal caregiver burden after one year compared to follow-up in primary care.[10,11] Both studies thus essentially differed from the design proposed in the current proposal. One non-controlled observational study on long-term (cost-) effectiveness of MCs in the UK reported minimally increased HRQL and QALYs two years after referral.[12]
Current diagnostics in Dutch MCs is very heterogeneous and less structured than in the aforementioned studies.[13] Meanwhile, structured elderly care programs and specialized nurses were introduced in primary care.[14] This means the diagnostics and care landscape has changed considerably since the above-mentioned studies were conducted. With this increasing relevant expertise in primary care within and in collaboration with general practices, the question about the value of a specialist diagnosis has become relevant again. Moreover, the guidelines and policy makers calling for primary care diagnoses need an evidence-base for this claim.[3,4]

1.2 Aim

Dementia is a complex syndrome for which no effective disease-modifying treatment is available. Therefore, it is crucial for the evaluation of a diagnostic trajectory to focus on clinical consequences, rather than diagnostic accuracy

alone.[15,16] As relatively short diagnostic trajectories are likely to have long-term disease-management consequences, long-term follow-up is essential.

Therefore, we aim to investigate if diagnostics in primary care is not inferior to diagnostics in a memory clinic in terms of long-term outcomes that matter to patients and against equal or lower costs.

2. Research questions

Primary research question

What is the comparative efficacy and safety between dementia diagnostics in primary care and dementia diagnostics in a memory clinic for older persons with memory complaints?

Additional research questions

a) What is the comparative effectiveness between diagnostics in primary care and memory clinic referral in older persons with memory complaints with regard to outcomes that matter to patients?

b) How are diagnostic and subsequent care trajectories performed and experienced, and which patient characteristics and quality of care indicators are associated with daily functioning and healthcare costs in both study arms?

Hypotheses

- Primary care diagnostics is not inferior to MC diagnostics with regard to clinically relevant outcomes including daily functioning (primary outcome), informal caregiver burden, number of acute admissions in hospitals or nursing homes, and time to institutionalization (clinical effectiveness)
- Outcomes of primary care diagnostics are not inferior to outcomes after memory clinic referral and generate less healthcare costs (cost-effectiveness)
- A primary care diagnostic trajectory does not cause more insecurity, anxiety or dissatisfaction in older persons and their informal caregivers (safety)

If GP diagnostics is not inferior to specialist diagnostics, GPs are supported to take up a more prominent role in diagnostics for older people with memory complaints, which is concordant with the National Research Agenda General Practice, and consistent with national guidelines and regional initiatives.[3,4] Moreover, a shift towards more primary care diagnoses leads to decrease of unnecessary and potentially burdensome testing and healthcare costs, and to more efficient use of available specialist expertise for those who really need it.

3. Study design

3.1 Context

We will perform a long-term cost-effectiveness trial with non-inferiority design complemented with a mixed-methods process evaluation. The choice for an RCT to investigate the proposed issue is pertinent following Gluud & Gluud emphasizing the importance of thorough evaluation of diagnostic tests and procedures already in BMJ in 2005 and recently emphasized again by Kennedy in JAMA [15,16] These researchers underline that once the specificity and sensitivity of a test have been established (labelled as a phase II study), the ultimate question relating to diagnostic value is whether tested patients fare better than similar untested patients, preferably in an RCT.

Particularly in the absence of treatment, diagnostic accuracy alone is insufficient as key outcome for patients, family and society.
To be able to generalize study findings to persons with strong preferences to either primary care diagnostics or memory clinic diagnostics, older persons who do not want to be randomized are included in a longitudinal prospective cohort with the same outcomes and follow-up protocol.

Since the start in 1980s, the value of memory clinics seemed self-evident: dementia was under-recognized and under-diagnosed and GPs experienced many barriers to diagnosing dementia.[8, 17] Pro-active care for older persons was non-existent. With increasing relevant expertise in primary care within and in collaboration with general practices, the question about the incremental value of a specialist diagnosis has regained relevance. Moreover, the growing number of guidelines and policy makers calling for primary care diagnoses need a robust evidence-base for this claim.

NON-INFERIORITY of primary care diagnosis compared to memory clinic diagnosis is of interest on the premise that primary care diagnosis has some other advantages for patients and society:
- Patients profit from greater availability, less invasiveness and fewer possible adverse effects (harms), integrating care on co-existing chronic diseases, GPs' deep routed knowledge on their individual history, personal and societal circumstances, and from greater patient and family continuity of care.
- Society benefits from reduced costs and better availability of specialist services for those who need it.

The interpretation of trial results in this complex field requires special attention for the effects of variations within both study arms (diagnostic trajectory, initial diagnosis) that may lead to heterogeneity in diagnostic disclosure, advice, follow-up and long-term outcomes. Therefore, we will include a mixed-methods process evaluation based on the principles of MRC guidance for complex interventions [18], to investigate how the 'diagnostic intervention' in each arm leads to certain outcomes and to far greater depth of understanding about how outcomes are influenced by both the medical and socio-cultural context. We will consider factors related to GPs, patients (including their preferred allocation), caregivers, diagnostic and follow-up trajectories. This information will additionally help GPs (and specialists) to better tailor diagnostic trajectories and follow-up in the future.

3.2 Study setting and duration
Recruitment of older persons with memory complaints takes place in primary care practices. GP cooperatives in the Nijmegen, Amsterdam, Boxmeer and Arnhem areas have already committed to including practices and patients. Follow-up after inclusion will be 24-30 months. For participants enrolled during the final six months of the inclusion period, the follow-up duration will be reduced from 30 months to a minimum of 24 months.

4. Population
4.1 Population (base)
The study population consists of persons over 70 years who visit their GPs with memory complaints.
The incidence of dementia (20,000 annually) is high. This underlines the relevance of the issue under study and is an important facilitator for recruitment. An average GP practice has 14 patients with memory problems who have not been diagnosed, and an additional 2-3 patients first presenting with memory problems annually [2], who are all potentially eligible.
If half the eligible patients are willing to participate, 60 practices should be sufficient to recruit 300 participants within 12 months. We have previously successfully recruited GPs (who included four patients on average during 6-9 months) for RCTs with persons with memory complaints (N> 100) or dementia (N=38).[19, 20] Generally, the project team has extensive experience with (large-scale) clinical trials in the field of dementia, some with long follow-up up to eight years with minimal drop-out on the primary outcome.[21,22]

Specific challenges are the willingness to randomize or to be randomized between diagnostic options that may be preference sensitive. Interviews with 15 GPs and a survey among a heterogeneous sample of 26 GPs show that GPs feel capable diagnosing dementia (96%) and are willing to randomize patients (89%). The Network 100 elderly panel recognized the issue of patient preference with regard to diagnostic options, but therefore emphasized the relevance of the proposed trial (See 4.20). The patient panel of the NFU commented on the current research proposal: 'this is useful care' (Dit is zinnige zorg).

4.2 Inclusion criteria

All patients 70 years and older consulting the GP with memory problems are eligible if patient and GP consider starting diagnostic evaluation.

4.3 Exclusion criteria

Patients are excluded if referral is clearly indicated due to suspicion of an uncommon form of dementia or another brain disease with accompanying memory problems, based on history or focal signs on neurological examination, such as extrapyramidal signs, hemiparesis or Babinski's sign. Patients are also excluded if the GP considers referral undesirable, e.g. in case of concomitant terminal illness.

4.4 Sample size

We will recruit 182 older persons with memory complaints in the RCT. Our sample size calculation is based on the primary outcome of the A-iADL-Q-SV. In this mixed population with subjective memory complaints, MCI and early dementia, we expect an average yearly decline of 4.2 points (SD=7.25), and an estimated SD at each measurement of 9.0. The Minimally clinically important difference (MCID) is estimated to be 6.3 points (for all estimates, see Appendix 1). We defined the non-inferiority margin as 60% of the MCID during follow-up. With an average decline of 10.5 points in 30 months, an SD adjusted for baseline A-IADL-Q-SV and diagnosis of 9.0, a non-inferiority margin of 3.8 points, 80% power, and a one-side alpha of 0.025, we need 2 x 79 participants to show that primary care diagnostics is not inferior to referral to a memory clinic. We increased this to 2 x 91 participants to compensate for 15% drop-out. We do not take cross-over from the primary care arm to the referral arm into account in the sample size calculation, since in this pragmatic trial the primary analysis will be based on the intention-to-treat principle. Cohort data are analyzed separately using similar protocols as RCT data and outcomes will be used to compensate the minimal loss of power that occurred by setting a non-inferiority margin of 60% instead of the more common 50%.

The shorter follow-up period for participants enrolled in the last six months has only a minor impact on the power calculation, as this calculation is based on a comparison made at the end of the study rather than on repeated measurements. Additionally, only a small subset of participants near the end of the inclusion period will be affected.

5. Treatment of subjects

Intervention

For the primary care arm, the diagnostic procedures are described in the Dementia Guideline of the Dutch College of General Practitioners. The minimum diagnostic work-up includes: History taking with patient (and caregiver if available), cognitive screening (MMSE and clock drawing test) and assessment of daily functioning. Additional diagnostics may be performed at the discretion of the GP, and may include blood tests, practice nurse consultation, consultation of an Elderly Care Physician (ECP), as available in the practice or regional settings. The diagnostic criteria according to McKahn are used.[23]

Participating GPs will receive a summary of the GP guideline in a flow-chart format and an overview of local diagnostic services and collaboration agreements. All in line with current guidelines, and without additional diagnostic procedures.

Comparator

In the referral arm, patients will be referred to a MC of the referring GP's and referred patient's choice. The diagnostic procedures at the MC are completely at the discretion of the physician/multidisciplinary team, and may include neuroimaging with CT/MRI, neuropsychological examination, CSF examination, EEG and nuclear imaging.

The use of diagnostic instruments, initial diagnosis and the time to diagnosis are collected in both arms. All professionals provide care as usual over the long-term follow-up.

6. Methods

6.1 Study parameters/endpoints

6.1.1 Primary outcome

Our primary outcome is daily functioning, since most aspects of dementia (cognition, behavior, mood, motor function), including factors related to caregivers, eventually impact on daily functioning. In addition, functioning determines the need for home care, institutionalization and thus costs. Moreover, living at home independently is a priority outcome for older persons.[24]

Outcome assessment is performed by a researcher blinded to allocation. Because of the heterogeneity of the population, we use both disease specific and more generic instruments, developed in the Netherlands or validated in the Dutch healthcare setting:

Daily functioning:

- Amsterdam Instrumental Activities of Daily Living Questionnaire Short Version (A-iADL-Q-SV). (Time to complete 20 min; dementia-specific)[25]

This 30-item instrument was developed using Item Response Theory (IRT) to ensure robust psychometric properties. It has a linear scale with mean of 50 and standard deviation of 10. It includes items on seven categories: household activities (e.g. doing groceries or cooking), household appliances (e.g. using microwave or dishwasher), finances (e.g. paying bills), work, computer (e.g. using internet), appliances (e.g. remote control) and leisure activities (e.g. driving or playing games).

6.1.2 Secondary outcomes

- Diagnostic accuracy. For this we will use the first dementia diagnosis as index diagnosis, and the diagnosis at 30 months as reference standard. Dementia diagnosis will be assessed using an algorithmic approach, with all clinical information available in electronic medical records.
- Time to dementia diagnosis. We will not analyze time to MCI diagnosis, since this is not a diagnostic entity as used in primary care.
- Acute admissions and time to institutionalization (months)
- (Time to) mortality

- Quality of life: EQ-5D-5L (Time to complete 5 min) [26]

- Costs: modified version of the Resource Utilization in Dementia (Time to complete 15 min) [27]. Section A1.4 Caregiver Health Care Resource Utilisation will be omitted, following a similar approach as the RUD Lite, a shortened version of the RUD [28]. This section will be excluded because caregiver resource use is generally low and has little influence on total societal costs.

- Cognition: Short cognitive screening test Mini Mental State Examination (MMSE) (Time to complete 5-10 min)

- Behavior: Neuro Psychiatric Inventory Questionnaire (NPI-Q) (Time to complete 5-15 min) [29]
  The NPI was developed to assess psychopathology in patients with dementia. It evaluates 12 neuropsychiatric disturbances common in dementia: delusions, hallucinations, agitation, dysphoria, anxiety, apathy, irritability, euphoria, disinhibition, aberrant motor behavior, night-time behavior disturbances, and appetite and eating abnormalities. The severity and frequency of each neuropsychiatric symptom are rated on the basis of scripted questions administered to the patient's caregiver. The NPI also assesses the amount of caregiver distress engendered by each of the neuropsychiatric disorders.

- Mood: Geriatric Depression Scale (GDS-15) (Time to complete 5 min) [30]
  The GDS is a self-rating instrument for depression in older persons. The 15-item version is validated for the primary care setting and for patients with memory complaints and dementia.

- Caregiver burden: Perseverance time (Time to complete 1 min) [31]
  Perseverance time measures the time for which a caregiver expects to be able to continue providing care if the caregiving situation remains as it currently is, and it includes six ordered answering categories: <1 week, 1 week–1 month, 1–6 months, 6 months–1 year, 1–2 years, and >2 years. It integrates the aspect of perceived burden with the caregiver's capacity to cope with the burden, in contrast to most available instruments, which measure solely the burden of caregiving.

Follow-up:
Outcomes are measured at 6 months, 18 months and (24-)30 months (4 time points including baseline) or attrition due to death, because the relatively short diagnostic trajectories are likely to have long-term consequences due to disease management decisions following diagnosis. Measurements may take place either one month earlier or two months later. In case of further deviation, we document the reason. In case of institutionalization, primary outcome assessment continues, other outcomes if feasible and appropriate.
We will administer a short questionnaire if caregivers no longer wish to complete the A-IADL-Q-SV at follow-up measuments. Its purpose is to gather a minimal dataset, ensuring these participants can still be included in the main analysis. The questionnaire contains four questions addressing mortality, changes in place of residence, and hospital admissions/ER visits. These questions are taken from the RUD and are also used at T6 for other participants.

Questionnaire	**Type of COA**	**Administration method**	**M0**	**M6**	**M18**	**M30**
A-IADL-Q-SV	ObsRO	Self-administered	X	X	X	X
EQ-5D-5L	PRO	Interviewer-administered	X	X	X	X
RUD	ObsRO	Interviewer-administered	X	X	X	X
MMSE	PerfO	Interviewer-administered	X			X
NPI-Q	ObsRO	Self-administered	X	X	X	X
GDS-15	PRO	Interviewer-administered	X	X	X	X
Perservance time (Pt)	ObsRO	Self-administered	X	X	X	X

COA: Clinical Outcome Assessment; ObsRO: Observer-reported outcome measure. In this study the caregiver is the observer; PRO: Patient reported outcome measure; PerfO: Performance outcome. "X" represents administration of a specific questionnaire at a given timepoint (M0: baseline; M6: after 6 months; M18: after 18 months; M30: after 30 months).

6.1.3 Process evaluation

In the mixed-methods process evaluation, we aim to explore which mechanisms lead to which effects in both arms in the complete trajectories of diagnosis and follow-up and how the different contexts influence these mechanisms and the (RCT) outcomes. We will use the MRC guidance for process evaluations.[18] We will consider factors related to the GP, the patient, the caregiver, the diagnostic and follow-up trajectories, collecting a combination of quantitative and qualitative process measures and patient/professional reported experience measures (PREMs). These include:

- Professionals: diagnostic certainty GP at the moment of diagnosis and at 18 and 30 months (VAS);
- Patient and caregiver: satisfaction with diagnostic trajectory (VAS/interview), acceptance of diagnosis (VAS/interview), coping with diagnostic uncertainty (VAS/interview)
- Diagnostic trajectory: number and type of diagnostic tests used including consultation of other healthcare providers in primary care or in memory clinic and adverse events, number of additional specialist consultations (electronic medical records)
- Follow-up trajectory: care use in general (electronic medical records: Minimum DataSet integrated dementia care (including indicators on involvement case managers, yearly multidisciplinary meeting, medication review, advance care plan)[32], Resourse Utility in Dementia (RUD)).[27]

6.1.4 Participant characteristics

We will collect the following participant characteristics:

- Patients: gender, age, education, comorbidity, medication, living situation, relationship to informal caregiver
- GPs: gender, age, type of practice, attitude towards dementia diagnosis

6.2 Randomization, blinding and treatment allocation

Randomization will be in a 1:1 ratio and performed at the individual level, to avoid selection bias that may result from cluster randomization at GP level, as GPs may tend to recruit patients for the trial that fit their allocation. Additionally, GPs may be less inclined to participate if they are randomized to referring all patients or doing all diagnostics in primary care. Randomization procedures will be performed digitally in Castor-EDC. Block randomization with different block sizes will be used.

In this study, participants cannot be blinded. The researcher will be blinded to the allocation of the patient when performing baseline and follow-up assessments and when extracting data from electronic medical records. GPs will be blinded for all their patients' study measurements. Patients' experiences and satisfaction may be influenced based on the match of their preferred and actual allocation. The primary outcome iADL is measured with an instrument developed for and validated in the population of study, which prevents this lack of blinding to significantly impact our results.

6.3 Study procedures

6.3.1 Recruitment of participants

Recruitment takes place in participating primary care practices. Practices will be recruited via GP cooperations in and around Gelderland, part of Limburg and North Brabant, Amsterdam and Leiden (several cooperations signed a letter of intent during the proposal phase) during the first year of the study. Patient inclusion will start after six months and continue during 27 months. Patients will be recruited by the GPs when visiting the practice for memory complaints. The GP will shortly explain the study and if patients are interested to participate, the GP shares the Patient Information Leaflet (PIL), including an informed consent form (IC). Permission will be asked to send the patient's contact data to the project's research assistant via secured e-mail (zorgmail), who subsequently calls the patient within a week to make an appointment to sign the IC form and perform a baseline assessment. Depending on the preferences of patient and informal caregiver, the baseline assessment will take place at the patient's home, or in the practice. The research assistant will complete a written IC form together with patient and informal caregiver. At this moment the baseline assessment consisting of a series of questionnaires is performed.
To assess coverage/reach of the study, GPs will be asked to register age and sex of all patients that were eligible and informed about the study, but did not wish to participate. GPs are also asked to register the same data of patients with memory complaints that were not eligible, either because referral was clearly indicated or because referral was undesirable.

A purposeful sample of the participants will be recruited to participate in an interview for the process evaluation. Relevant predefined criteria for sampling of patients are: allocation (preference), process measures and PREMS, and also gender, age and relation to the informal caregiver will guide purposive sampling of patients. GP sampling is based on their gender and age (as these characteristics are associated with attitudes to cognitive decline and their regional support services). For this qualitative sub-study, additional IC will be obtained.

6.3.2 Data collection

We will use Castor-EDC to build an e-CRF to collect and store quantitative data. We will assign a unique identification code to each research participant to prevent data obtained and analyzed for scientific research from being directly traceable to the participant. In a secured identification log, we will separately record and store the directly traceable data of the enrolled study participants (name, address, telephone number, medical file number (MDN), etc.) of data collected and used for research (eCRF, data file for analysis, etc.). In the Data Management Plan (DMP) we will specify where the identification log and research data are stored and who has access to these locations. Coding is composed and assigned chronologically, based on IC signing date. The GP practice will be part of the identification code.

6.3.2.1 Demographic data and questionnaires

Validated questionnaires on primary and secondary outcomes, information on participant characteristics and randomization preferences of both patients and GPs will be completed at baseline. For self-administered questionnaires, this can take place digitally by sending a survey link directly from Castor to the participants or on paper where after the research assistant transfers the data to Castor. Questionnaires that were developed as an interview will be administered by the researcher and entered directly into the Castor e-CRF. All questionnaires will be repeated at 6, 18 and 24-30 months. Questionnaires at six months will be supplemented with items on processes and experiences of the diagnostic trajectory. Questionnaires at 24-30 months will be supplemented with items on processes and experiences of the follow-up trajectory. (see 6.1.3)

6.3.2.2 Data from electronic medical records

At the end of follow-up (M30), participants' GPs will be asked to upload the patient electronic medical record from the complete study period via a secured link (Surf File Sender). Outcomes and process measures derived from primary care electronic medical records will be collected by a research assistant using a data extraction form. The first 20 records and every 10th record will be checked by a second researcher. In case of disagreement, discussion in the project group will lead to consensus.

6.3.2.3 Qualitative data

Data will be collected with individual and dyadic semi-structured interviews with patients and caregivers after finishing a diagnostic trajectory (around 6 months after inclusion) with a on the expectations, experiences and preferences during the diagnostic trajectory.) and 24 (focus on follow-up trajectory) months after inclusion. GPs will be interviewed after finishing a diagnostic trajectory on diagnostic certainty.
A topic list will be used by an independent interviewer, which will be adapted based on the data collected. All interviews will be audio-taped and transcribed verbatim. Data collection will end upon saturation (expected after 10-15 interviews per target group) and be performed in parallel with thematic analysis of the transcripts.

7. Statistical Analysis

7.1 Primary analysis

For the primary analysis, we analyze the difference in A-IADL-Q-SV between the intervention and control group at 24-30 months using linear regression analyses, adjusting for diagnosis and A-IADL-Q-SV score at baseline, using a one-sided alpha of 0.025%. We account for clustering within GP practices using mixed effects models with a random intercept per GP if this improves model fit according to the Akaike Information Criterion by ≥2 points. If residuals for linear regression models are not normally distributed, we use confidence intervals from bootstrapped analyses.

Primary analyses are based on the intention-to-treat principle. A per protocol analysis is performed and becomes more relevant if more participants from the primary care arm are referred during the course of the follow-up. The main analyses include all available data (i.e. pairwise deletion), and was powered on a drop-out of 15% from baseline. We conduct a sensitivity analysis using multiple imputation with chained equations (MICE) to account for missing data, based on the measurements acquired at each time-point.

7.2 Secondary analyses

Secondary analyses assess differences in yearly change using a mixed effects multivariable linear regression model, with the score on the A-IADL-Q-SV at 6, 18 and 24-30 months as outcome, adjusted for the A-IADL-Q-SV score at baseline, with a random intercept per participant to account for repeated measurements. Secondary outcomes are assessed using linear regression for continuous outcomes and Poisson regression for dichotomous outcomes. Differences between the intervention/control group in diagnostic accuracy, when the initial diagnosis is a dementia diagnosis, compared to the diagnosis at the end of follow-up (as gold standard) will be assessed using Fischer-exact tests, c-statistic, sensitivity/specificity and positive/negative predictive values. Potential effects of differences in mortality are assessed using Cox proportional hazard models with age at diagnosis at time of entry, and age at death/censoring as time to event.

We conduct several exploratory subgroup analyses, examining the following sub-groups: baseline MMSE (split at the median), age (split at the median), gender, living status (single vs partnered), highest attained education. P-values for the differences between subgroups are calculated using interaction terms.

Cohort data are analyzed separately using similar protocols as RCT data.

7.3 Economic evaluation

The economic evaluation investigates, along-side the clinical trial, the value for money of full implementation of a

diagnostic trajectory in primary care in patients with memory problems compared to the MC trajectory (usual care). This will be done from a health care/societal perspective (applying the Resource Utilization in Dementia (RUD)[27] on patient and informal caregiver level). Cost and QALYs will be measured on a per patient/informal caregiver basis over a relevant time path (24-30 months) in which the (most important) differences in both QoL and costs between both arms are supposed to manifest themselves. The design of the economic evaluation follows the principles of a cost-utility analysis and adheres to the Dutch guideline for performing economic evaluations in health care.[33] Cost-effectiveness over the 24-30 months evaluation period will be expressed in terms of cost per QALY gained. To determine robust confidence intervals surrounding the Incremental Cost-Effectiveness Ratio (ICER), we use (non)parametric bootstrapping. Further uncertainty will be dealt with by one-way sensitivity analysis (deterministic) and by parametric statistics ultimately presenting cost-effectiveness planes and acceptability curves.

7.3.1 Cost analysis

The cost analysis, based on the RUD, exists of two main parts. First, on patient/informal caregiver level, volumes of care will be measured prospectively over the time path of study, if necessary, completed by data from the GP and hospital electronic medical records. Productivity losses of informal caregivers, which we expect to be minor as many informal caregivers will be (retired) spouses, are measured using the friction cost approach (ZIN, 2016). Second, per item of health care consumption standard cost-prices will be determined using the appendix 1 of the guideline for performing economic evaluations (ZIN, 2016). If standardized prizes are not available, by full cost prices via activity-based costing.
Between group differences in costs will be evaluated using regression-based techniques that are able to deal with skewness and heteroscedasticity in the data (for example a GLM with gamma distribution and log/identity link function).

7.3.2 Patient outcome analysis

The effect analysis adheres to the design of the randomized controlled trial and measures at baseline, and at 6-18-(24-)30 months during the clinical trial. To measure the quality of the health status of the patients, a validated health-related quality of life (HRQoL) instrument will be used, the EuroQol-5D (EQ-5D) [34], which was recently also validated in persons living with dementia.[35] This HRQoL instrument will be completed by the patients or caregivers and is available in a validated Dutch translation (EQ-5D 5L). The EQ-5D is a generic HRQoL instrument comprising five domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The EQ-5D index is obtained by applying predetermined weights to the five domains. This index gives a societal-based global quantification of the patient's health status on a scale ranging from 0 (death) to 1 (perfect health). Patients will also be asked to rate their overall HRQoL on a visual analogue scale (EQ-5D VAS) consisting of a vertical line ranging from 0 (worst imaginable health status) to 100 (best imaginable).

7.4 Process evaluation

Quantitative data analyses will be exploratory. All data will be descriptively analyzed. Linear and logistic mixed-models comparing processes and PREMS in both arms help us to answer questions as e.g. 'how is satisfaction with the diagnostic trajectory/acceptance of the initial diagnoses associated and do these associations differ between the study arms'? Mixed-models are also used to study associations between mediators as diagnostic and/or process measures and relevant contextual factors as a living situation of the informal caregiver, and trial outcomes.

Qualitative data will be thematically analyzed. Open coding will start directly after the first interview, will be performed by one researcher and supervised by a second researcher. Two researchers will perform axial coding, compile categories and eventually themes. In case of disagreement discussion will lead to consensus. Atlas-Ti will be used to support the analysis process.

Integration of data will take place at a methodological level.[36] Quantitative baseline data will guide purposive sampling for the interviews. Interviews may reveal hypotheses on relevant contextual factors and mediators/mechanisms to inform quantitative analyses. Additionally, combined description and interpretation of quantitative and qualitative data will further add to data integration.

8. Ethical considerations

8.1 Regulation statement

This study will be conducted according to the principles of the Declaration of Helsinki (version 2013). This research protocol was submitted to the Medical Research Ethics Committee (MREC Oost-Nederland), which concluded that the study is subject to the Medical Research Involving Human Subjects Act (WMO). The entire research file was then submitted for review by the MREC for ethical approval.

8.2 Recruitment and consent

See 5.2.1.

8.3 Benefits and risks assessment, group relatedness

8.3.1 Safety and efficacy

Both diagnostic trajectories comprise forms of regular care. Risks in both study arms are low. We will identify and register these risks according to NFU guidelines 'Kwaliteitsborging Mensgebonden Onderzoek' on low-risk studies (see 6.1.3).

In the MC arm, the main risks are 1) complication of a diagnostic procedure as lumbar puncture, which has a very low risk of serious complications such as subdural hematoma and meningitis) and 2) incidental findings on brain CT or MRI[37], potentially leading to anxiety or additional (diagnostic) procedures or, very rarely, even to inappropriate neurosurgical procedures.

In the primary care arm, the main consequence may be a delayed diagnosis, resulting in a period of uncertainty for a patient and caregiver. Appropriate counseling may mitigate the burden of this uncertainty. Most importantly, a potentially delayed diagnosis does not affect the prognosis, because no disease-modifying treatment exists. People diagnosed with dementia by a GP are very likely to be true positives: a 2012 literature review and a recent UK study report specificity of GP clinical judgement for the diagnosis of dementia between 85-100% and a positive predictive value around 85%.[38,39] Sensitivity of dementia diagnosis in primary care is low [38], as general practitioners perceive various persistent barriers to dementia diagnosis, including lack of knowledge, lack of confidence with diagnosis, and a lack of time.[17] Whether that is a problem is to be answered by this study. It is highly unlikely that GPs miss uncommon causes of cognitive impairment with therapeutic consequences. The risk of missing intracranial pathology is very low: the prevalence of subdural hematoma or brain tumor in older patients in primary care with cognitive impairment without other symptoms or signs on neurological examination is negligible.

As for efficacy, comparative long-term (cost-)effectiveness of primary care and MC diagnostics including post-diagnostic care was never studied in an RCT.

8.3.2 Burden to participants

Burden to patients and caregivers

The burden is limited. Time to complete the combination of questionnaires for additional outcome assessment per time point (M1, M6, M18 and M30) is less than one hour. Patients and caregivers do not experience the possible burden of a new intervention, as diagnostics either in primary care or memory clinic are part of usual care. Interviews are held with a purposive sample only, which is voluntary.

Burden to GPs

Because we investigate two usual care procedures, the burden for participating GPs is limited to asking older persons (and their caregivers) to participate and have them sign an informed consent form. Additionally, they are asked to complete a short baseline questionnaire including individual and practice characteristics. Per patient, we ask them to share their experiences and diagnostic confidence directly after the diagnostic trajectory, and after 18 and (24-)30 months. At the end of follow-up, we will ask them provide the patient's medical record. Because the number of participants per practice is low, we expect compliance to be high. Interviews are held with a purposive sample only, which is voluntary.

9. Administrative aspects, monitoring and publication

9.1 Handling and storage of data and documents

We have drafted a digital Data Management Plan (DMP) according to the digital Radboudumc format. Data will be collected and stored in Castor-EDC. Monitoring will be conducted in compliance with the SOP Monitoring of Radboudumc. Handling and storage of collected data and documents will comply with the General Data Protection Regulation (AVG). Data will be stored in a secured folder at the department's H-schijf or the Radboudumc Digital Research Environment. Data can only be accessed by authorized members of the research team. Sensitive, personal data will be saved in an identification file in a separate folder. This file will be destroyed after the database is locked. At this point, only coded data is available to use, so this data cannot be traced back to a single person. This coded data is used when publishing the results of this study. After interview transcripts are pseudonymized, the audio records will also be destroyed. Other documents will be stored for fifteen years at the research location.

This is a prospective study, collecting new data. Where possible, we will use data that are routinely registered in general practice and memory clinics that will be carefully pseudonymized. New quantitative data will be collected for (cost-)effectiveness outcomes, which need to be aligned in both study arms. New qualitative data from interviews with patients, caregivers and doctors will inform on mechanisms leading to these outcomes. Data will be safely stored in the Radboudumc Digital Research Environment and archived according to the FAIR principles and made available via the national archive DANS. Anonymized data will be made available for future research questions by other researchers, given the data are adequate to answer their question and it is of sufficient scientific relevance and quality.

9.2 Public disclosure and publication policy

Data will not be traceable to persons in reports (PhD thesis, medical scientific publication). The team commits to making all scientific publications resulting from this project publicly available free of charge via open access avenues. Costs are largely covered due to contracts of the overarching Dutch Federation of Universities with most major publishing companies. If open access publishing is not directly possible via the journal, then a copy will be made available via the Radboud repository (after the embargo period). Non-scientific publications, such as reports and recommendations will be made publicly available, for example via websites (ZonMw, DementieNet, Network 100, GP cooperatives, NHG/thuisarts.nl) and our social media channels LinkedIn and Twitter.

10. References

1. *Dutch Alzheimer Association: Factsheet cijfers en feiten over dementie*. p. www.alzheimer-nederland.nl.
2. Nielen, M.M.J., K. Hek, and T.R.J. Schermer, *Jaarcijfers aandoeningen: incidenties en prevalenties. Uit: Nivel Zorgregistraties Eerste Lijn.* 2020.
3. *NHG-Standaard Dementie*. 2020, Nederlands Huisartsengenootschap Utrecht.
4. *Richtlijn Diagnostiek en Behandeling van Dementie*. 2020, Nederlandse Vereniging voor Klinische Geriatrie: Utrecht.
5. Gruters, A.A.A. et al. Int J Geriatr Psychiatry, 2019. **34**(8): p. 1267-1274.
6. Jack, C.R. et al. Alzheimers Dement, 2018. **14**(4): p. 535-562.
7. Dhedhi, S.A. et al. BMJ Open, 2014. **4**(3): p. e004439.
8. Lang, L., et al. BMJ Open, 2017. **7**(2): p. e011146.
9. Wolfs, C.A., et al. Br J Psychiatry, 2008. **192**(4): p. 300-5.
10. Meeuwsen, E., et al. PLoS One, 2013. **8**(11): p. e79797.
11. Meeuwsen, E.J., et al. BMJ, 2012. **344**: p. e3086.
12. Gomes, M., et al. Int J Geriatr Psychiatry, 2019. **34**(3): p. 439-446.
13. Hafdi, M., et al. Ned Tijdschr Geneeskd, 2021. **165**.
14. Dutch Organization for Research and Development. *Beter oud worden - 10 jaar investeren in betere zorg en ondersteuning voor ouderen*. https://www.beteroud.nl/docs/beteroud/over-ons/krachtboekje-npo.pdf.
15. Gluud, C. and L.L. Gluud. BMJ, 2005. **330**(7493): p. 724-6.
16. Kennedy AG. JAMA. 2022 Apr 12;327(14):1335-1336.
17. Mansfield, E., et al. Gerontologist, 2019. **59**(6): p. e697-e708.
18. Moore, G.F., et al. BMJ, 2015. **350**: p. h1258.
19. Perry, M. *Development and evaluation of a Dementia Training Programme for Primary care*. 2011, Radboud UMC: Nijmegen.
20. Tilburgs, B., et al. J Am Med Dir Assoc, 2020. **21**(6): p. 837-842 e4.
21. Richard, E., et al. Lancet Digit Health, 2019. **1**(8): p. e424-e434.
22. Moll van Charante, E.P. Lancet, 2016. **388**(10046): p. 797-805
23. McKhann, G.M., et al. Alzheimers & Dementia, 2011. **7**(3): p. 263-269.
24. Robben, S.H., et al. J Am Geriatr Soc, 2011. **59**(8): p. 1552-4.
25. Sikkes, S.A., et al. Alzheimers Dement, 2012. **8**(6): p. 536-43.
26. EQ-5D
27. Wimo, A., et al. Alzheimers Dement, 2013. **9**(4): p. 429-435 e17.
28. Wimo, A., et al. Brain Aging, 2003. **3**(1): p. 48– 59
29. Cummings, J.L., et al. Neurology, 1994. **44**(12): p. 2308-14.
30. Mitchell, A.J., et al. J Affect Disord, 2010. **125**(1-3): p. 10-7.
31. Richters, A., et al. Alzheimers Res Ther, 2017. **9**(1): p. 26.
32. Oostra, D.L., et al. BMJ Open Qual, 2020. **9**(2).
33. Zorginstituut Nederland. *Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg*. 2016: https://www.zorginstituutnederland.nl/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg.
34. Versteegh, M.M., et al. Value in Health, 2016. **19**(4): p. 343-352.
35. Michalowsky, B., et al. Value in Health, 2020. **23**(6): p. 760-767.
36. Fetters, M.D. et al. Health Serv Res, 2013. **48**(6 Pt 2): p. 2134-56
37. Vernooij, M.W., et al. N Engl J Med, 2007. **357**(18): p. 1821-8.
38. van den Dungen, P., et al. Int J Geriatr Psychiatry, 2012. **27**(4): p. 342-54.
39. Creavin, S.T., et al. BJGP Open, 2021. **5**(5).

Appendices

Appendix 1: literature review on expected decline in daily functioning

Our study is a randomized controlled diagnostic clinical trial. As primary outcome of our comparison we use functioning, assessed with a linear ADL scale, because this captures all aspects related to the disease, but also to possible downstream consequences of a diagnostic pathway, including changes in disease management, complications of diagnostic procedures and the consequences of timing of diagnosis. We focus this systematic review on the expected decline in daily functioning in a mixed population of dementia, mild cognitive impairment and subjective cognitive decline.

Search terms

We have systematically searched the literature the expected decline in functioning – Amsterdam Instrumental activities of daily living (A-iADL-Q)

We searched Medline using the keywords in the table below, with terms for the Amsterdam IADL questionnaire, cross-referenced with terms for the target population using "AND". All terms were exploded to their subject headings if possible:

Terms for IADL scale	Terms for population
(Amsterdam Instrumental activities of daily living) OR (Amsterdam IADL)	(Dementia OR (cognitive impairment) OR memory clinic)

Methodological filters

We included studies up to level 4 on our search (expected decline in daily functioning in a mixed population of dementia, mild cognitive impairment and subjective cognitive decline)

Levels of evidence:

1a. Systematic review of RCT's
1b. RCT
2. Cohort study with controls
3. Case-control study
4. Patient series without controls
5. Expert opinion

Databases

We searched the following databases:

- National library of medicine (Medline)
- Cochrane Library
    - Cochrane Database of Systematic Reviews (CDSR)
    - Database of Abstracts on Reviews and Effectiveness (DARE)
    - Cochrane Controlled Trial Register (CCTR)
- Current Controlled Trials (CCT): http://controlled-trials.com/
- ClinicalTrials.gov: http://clinicaltrials.gov/
- NHS Centre for Reviews and Dissemination (CRD): http://www.york.ac.uk/inst/crd/
    - NHS Economic Evaluation Database (NHS-EED)

- Health Technology Assessment Database (HTA)

Selection procedure, validity assessment

The search was performed by MP. MP, ER and JWvD, all post-doc level researchers with experience in systematic reviews, were involved in abstract screening and selection. All titles and abstracts were screened by two researchers. In case of discordance, consensus discussion took place.

We used the following quality criteria:

- Prospective clinical studies, RCT or cohort
- Longitudinal assessment of daily functioning using the A-iADL-Q
- Minimum of 25 patients
- Minimum of 12 months follow-up with a minimum of two measurements

Results

Determination of expected IADL change

As second input for our sample size, we required data on the expected course of IADL according to the Amsterdam IADL Questionnaire (A-IADL-Q) [2], in our target population. Because in absence of sufficient longitudinal data, the expected standard deviations (SD) of the mean scores is also informative for the sample size calculation, we also collected studies that published mean and SD's of the different populations likely to be present in our dataset.

Our search yielded 59 papers, of which, 34 full-texts were examined, and 9 were included. Two papers assessed the longitudinal change in the A-IADL-Q in our target population. As described in the table below. One study reported yearly change in whole points, estimated for individuals with AD as -7.4, SD=9.01 per year, translating to a standardized beta of 0.47, St.SD=0.58 based on the baseline SD (15.6).[3] Yearly change estimates for the subjective cognitive complaints (SMC) and mild cognitive impairment (MCI) populations were not provided, but these were stable. The other study only provided standardized estimates of the change per year, which resulted in an estimated standardized beta of -0.15 SD=0.54 per year for individuals with MCI, and -0.35 St.SD=0.46 per year.[4] The average of the estimates by Jutten and Koster for change and SD in individuals with dementia would be -0.41 SD=0.52.

Study	Pop	Duration	Diagnosis	Results (95%CI)	Change (SD)*	St.change (SD)*
Jutten 2018 [3]	Memory clinic	12 months	MCI n=62 m age: 74	Stb: -0.15 (-0.29 ; -0.02)		-0.15 (0.54)
			AD n=65 m age: 71	Stb: -0.35 (-0.46 ; -0.23)		-0.35 (0.46)
			SMC n=12	stable		
Koster 2015 [2]	Memory clinic	6-36 months	AD n=57 m age: 71	Beta: -7.4 (5.01-9.69)	-7.4 (9.0)	-0.47 (0.58)
			MCI n=23 SMC n12	stable		

Abbreviations: m age: mean age, MCI: mild cognitive impairment, SMC: subjective cognitive complaints, stb: standardized beta, st.change: standardized change with SD of change expressed as proportion of SD of baseline SD

**SD for change was calculated as the SE derived from the confidence intervals, times the square root of the number of individuals*

To estimate the approximate SD according to diagnosis for our population composition, we collated data from studies that reported mean and SD scores on the A-IADL-Q for individuals according to baseline cognitive status. Based on the literature, we expect an approximate composition of 25% individuals with SMC, 25% with MCI, and 50% with dementia.[5] The table below shows the means and standard deviations for the A-IADL-Q score, and it's shorter version, which we also collated because it has similar properties to the full version regarding score range and distribution.[6] We identified seven potentially relevant studies.[2,7-12] Overall, SD's ranged from 3.7-8.3 in individuals with SMC, 6.9 to 10.1 for individuals with MCI, and 8.6 to 10.1 for individuals with dementia. The average SDs accounting for baseline diagnoses with the expected population composition ranged from 8.4 to 9.0. Based on these data, we estimated the population SD at 9.0.

Study	Country	Population	Version	Diagnosis	N	Mean	SD	Exp.SD*
Milosevic 2022	Serbia	Memory clinic	Normal	SMC MCI Dementia	30 36 53	63.9 47.8 32	5.7 10.1 10.1	9.0
Jutten 2019	Netherl.	Memory clinic	Normal	SMC MCI Dementia	49 25 86	58.6 54.7 48.1	8.3 6.9 8.6	8.1
Dubbel-man 2020	Netherl.	Memory clinic	Normal	SMC MCI AD non-AD	29 19 75 46	59.2 53.6 45.7 47.4	8.3 8.3 7.9 10.1	8.7
Verrijp 2021	Netherl.	Volunteers	Normal	SMC+NC (50/50)	3288	65.9	4.8	
Villneuve 2019	France	Memory clinic	Short	SMC	289	67	3.7	
Sikkes 2013	Netherl.	Memory Clinic	Normal	No dementia Dementia	138 140	54.4 45.7	9.5 8.3	8.9
Brudener Hofstetter 2020	Switz.	Memory clinic	Short	MCI (48%) Dem (46%) SMC (6%)	56	54.7	8.4	8.4

** Exp.SD:* expected SD based on the study data and a population composition of 25% SMC, 25% MCI, and 50% AD.

Estimation of change (SD) for MCI and Dementia patients

By multiplying the baseline SDs of the different populations above, with the standardized decline estimates according to Jutten and Koster, it is possible to estimate an expected decline based on these baseline standard deviations. Based on the 8.6 to 10.1 range in SD for individuals with dementia and the average of the standardized scores by Jutten and Koster above (-0.41, St.SD=0.52), we would expect yearly decline ranging from -3.5 SD=4.5 to -4.3 SD=5.2. For individuals with MCI, the baseline range of 6.9 to 10.1 and the standardized findings by Jutten in MCI (-0.15, St.SD=0.54), we would expect a yearly decline ranging from -1.0 SD=3.7 to -2.0 SD=5.5. Since both Jutten and

Koster reported the scores of the SMC group to be stable, without giving an estimate in change, we expect the change and SD in the SMC group to be smaller than the MCI group.

Estimations used for the power calculation

For the power calculation, the SD determines the required sample size, with higher SD leading to larger samples. Conservatively assuming that the SD of the SMC population would be equal to that of the MCI population, the SD for the yearly change in the "no dementia" group (SMC + MCI) would be 5.5. In the dementia group, the SD for yearly change would be approximately 5.2. Given that Koster found an SD of 9.0 for the decline in the dementia population, we conservatively chose to use that higher estimate. Assuming a population composition of 50% dementia, and 50% MCI or SMC, this would result in an average overall expected SD in yearly change of (5.5+9.0)/2=7.25.

References

1. Sikkes SA, Pijnenburg YA, Knol DL, de Lange-de Klerk ES, Scheltens P, Uitdehaag BM. Assessment of instrumental activities of daily living in dementia: diagnostic value of the Amsterdam Instrumental Activities of Daily Living Questionnaire. J Geriatr Psychiatry Neurol. 2013 Dec;26(4):244-50.

2. Koster N, Knol DL, Uitdehaag BM, Scheltens P, Sikkes SA. The sensitivity to change over time of the Amsterdam IADL Questionnaire(©). Alzheimers Dement. 2015 Oct;11(10):1231-40.

3. Jutten RJ, Peeters CFW, Leijdesdorff SMJ, Visser PJ, Maier AB, Terwee CB, Scheltens P, Sikkes SAM. Detecting functional decline from normal aging to dementia: Development and validation of a short version of the Amsterdam IADL Questionnaire. Alzheimers Dement (Amst). 2017 Mar 31;8:26-35.

4. Gruters AAA, Ramakers IHGB, Kessels RPC, Bouwman FH, Olde Rikkert MGM, Blom MM, de Vugt ME, Verhey FRJ. Development of memory clinics in the Netherlands over the last 20 years. Int J Geriatr Psychiatry. 2019 Aug;34(8):1267-1274.

5. Jutten RJ, Peeters CFW, Leijdesdorff SMJ, Visser PJ, Maier AB, Terwee CB, Scheltens P, Sikkes SAM. Detecting functional decline from normal aging to dementia: Development and validation of a short version of the Amsterdam IADL Questionnaire. Alzheimers Dement (Amst). 2017 Mar 31;8:26-35.

6. Milošević V, Malobabić M, Stojanović I, Bašić J. Importance of a functional measure in the evaluation of patients in a memory clinic: Validation of the Serbian version of the Amsterdam instrumental activities of daily living questionnaire. Clin Neurol Neurosurg. 2022 Feb 3;214:107165.

7. Roos J. Jutten, Ellen Dicks, Lieke Vermaat, Frederik Barkhof, Philip Scheltens, Betty M. Tijms, Sietske A.M. Sikkes. Impairment in complex activities of daily living is related to neurodegeneration in Alzheimer's disease–specific regions. Neurobiology of Aging. 2019; 75: 109-116

8. Mark A. Dubbelman, Merike Verrijp, Leonie N.C. Visser, Caroline B. Terwee, Wiesje van Der Flier, Philip Scheltens, Sietske A.M. Sikkes. A mixed-methods approach to establish clinically meaningful categories of impairment in instrumental activities of daily living. Alzheimers Dement.2020;

9. Verrijp M, Dubbelman MA, Visser LNC, et al. Everyday Functioning in a Community-Based Volunteer Population: Differences Between Participant- and Study Partner-Report. *Front Aging Neurosci*. 2022;13:761932. Published 2022 Jan 5.

10. Villeneuve SC, Houot M, Cacciamani F, Verrijp M, Dubois B, Sikkes S, Epelbaum S; MEMENTO study group and the INSIGHT-preAD study group. Latent class analysis identifies functional decline with Amsterdam IADL in preclinical Alzheimer's disease. Alzheimers Dement (N Y). 2019 Oct 8;5:553-562.

11. Bruderer-Hofstetter M, Dubbelman MA, Meichtry A, et al. Cross-cultural adaptation and validation of the Amsterdam Instrumental Activities of Daily Living questionnaire short version German for Switzerland. *Health Qual Life Outcomes*. 2020;18(1):323