BMJ Open

Integrating expert knowledge for dementia risk prediction in individuals with mild cognitive impairment (MCI): a study protocol

Meng Wang [1,2] Eric E Smith [1] Nils Daniel Forkert,[1,3] Thierry Chekouo,[4] Zahinoor Ismail,[1,2,5] Aravind Ganesh [1] Tolulope Sajobi [1,2]

To cite: Wang M, Smith EE, Forkert ND, *et al*. Integrating expert knowledge for dementia risk prediction in individuals with mild cognitive impairment (MCI): a study protocol. *BMJ Open* 2021;**11**:e051185. doi:10.1136/bmjopen-2021-051185

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-051185).

Received 13 March 2021
Accepted 13 October 2021

Check for updates

© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Tolulope Sajobi;
ttsajobi@ucalgary.ca

ABSTRACT

Introduction To date, there is no broadly accepted dementia risk score for use in individuals with mild cognitive impairment (MCI), partly because there are few large datasets available for model development. When evidence is limited, the knowledge and experience of experts becomes more crucial for risk stratification and providing MCI patients with prognosis. Structured expert elicitation (SEE) includes formal methods to quantify experts' beliefs and help experts to express their beliefs in a quantitative form, reducing biases in the process. This study proposes to (1) assess experts' beliefs about important predictors for 3-year dementia risk in persons with MCI through SEE methodology and (2) to integrate expert knowledge and patient data to derive dementia risk scores in persons with MCI using a Bayesian approach.

Methods and analysis This study will use a combination of SEE methodology, prospectively collected clinical data, and statistical modelling to derive a dementia risk score in persons with MCI . Clinical expert knowledge will be quantified using SEE methodology that involves the selection and training of the experts, administration of questionnaire for eliciting expert knowledge, discussion meetings and results aggregation. Patient data from the Prospective Registry for Persons with Memory Symptoms of the Cognitive Neurosciences Clinic at the University of Calgary; the Alzheimer's Disease Neuroimaging Initiative; and the National Alzheimer's Coordinating Center's Uniform Data Set will be used for model training and validation. Bayesian Cox models will be used to incorporate patient data and elicited data to predict 3-year dementia risk.

Discussion This study will develop a robust dementia risk score that incorporates clinician expert knowledge with patient data for accurate risk stratification, prognosis and management of dementia.

Strengths and limitations of this study

- ► Experts' clinical knowledge about the relative importance of potential predictors will inform predictor selection for the proposed dementia risk score.
- ► The use of multiple data sources for internal and external validation of the risk prediction models is another strength of this study.
- ► The experts may not be representative of all knowledge experts in this field.
- ► Training data are obtained from a single center memory clinic and may be subject to referral biases.
- ► A lack of patient data on all potential predictors identified by the clinician experts in the training or validation data cohorts might influence the accuracy of the developed risk scores

INTRODUCTION

Dementia is a global challenge, affecting over 46 million people in 2015 with about 10 million new cases annually worldwide.[1] In Canada, the prevalence of dementia for individuals aged 65 years and older is about 7.1% and an annual incidence rate of 14.3 new cases per 1000 people.[2] By 2031, it is expected that the total annual healthcare costs for Canadians with dementia will have doubled those from two decades earlier, from US$8.3 billion to US$16.6 billion.[3] Dementia is typically preceded by mild cognitive impairment (MCI), defined as cognitive concerns with poor cognitive test scores but preserved activities of daily living. On average, 10%–20% of MCI population progress to dementia per year,[4] and in patients who develop dementia due to Alzheimer's disease (AD), progression usually occurs within 2–3 years.[5] Prediction of individual risk could be used to inform patients and clinicians, and to motivate preventive lifestyle modification as well as advance care planning.[6] Therefore, it is important to predict future dementia risk for individuals with MCI. However, according to recent systematic reviews, there is no recommended or widely accepted dementia risk score to use for individuals with MCI, due to limited data sources, lack of validation and unavailability of predictors.[7 8]

When empirical evidence is limited, clinicians need to rely on their knowledge

obtained from previous experience or heuristics to reach a dementia prognosis. But biases can happen due to cognitive capacity and time pressure, thus resulting in highly variable practice. Expert judgement (belief or intuition) is often a combination of fact-based knowledge with subjective impressions established from experience. Structured expert elicitation (SEE) includes formal methods to quantity experts' beliefs, aims to help experts to express their beliefs in a quantitative form, and reduces biases in the process. It is recommended to consider the elicitation in the same way as empirical data, by using repeatable, transparent methods and addressing questions in the form of probabilities.[9 10] SEE encourages experts to think critically considering all relevant information[11] and promotes analytical thinking.[11] Evidence has shown that training and repeated feedback can improve accuracy and reduce overconfidence.[11 12]

Although SEEs have been applied in different types of decision making,[13 14] to date, its application in health sciences has largely been restricted to informing national level health technology assessment (HTA) decisions.[12] A few studies[15 16] have reviewed the SEE used in HTA, and found that most of the elicited information was related to cost effectiveness of interventions,[16] and usually in the form of proportions, event probabilities and diagnostic accuracy.[15 16] Other examples include the use of SEE in determining sample size calculations for clinical studies.[17] Most importantly, although expert knowledge can feed directly into a decision itself, with some data available, combining the two sources of information (formally) is usually preferred. For example, in a Bayesian framework, one can integrate expert knowledge as prior information, to estimate the probability of an outcome along with current data.[18] However, there has not been a formal investigation of this methodology to improve the accuracy of risk prediction models.

We hypothesise that integrating expert knowledge with patient data to develop dementia risk prediction model will lead to a more accurate risk score than risk scores developed using (patient) data alone. This study aims to (1) assess experts' beliefs about important predictors for 3-year dementia risk in persons with MCI and (2) integrate expert knowledge and patient data to improve dementia risk prediction in persons with MCI.

METHODS

Structured expert elicitation

A SEE will be used to obtain experts' belief on important predictors of dementia risk in individuals with MCI. To obtain each expert's best-considered answers, SEE will focus on helping experts to think analytically while minimising biases.[9–11] Bojke *et al*[12] reviewed 16 different SEE guidelines, and found that the underlying elicitation process is similar: pre-elicitation (eg, what quantities to elicit), elicitation conducting (eg, level of elicitation) and postelicitation (eg, aggregation).[12] Two SEE protocols will be followed including the elicitation protocol under the healthcare decision making (HCDM) setting and the Investigate, Discuss, Estimate and Aggregate (IDEA) protocol. The IDEA protocol has been applied in many fields (eg, education, psychology, ecology, and conservation) and shown to yield relatively reliable judgements.[10]

Table 1 describes the elicitation stages and other documents including the Invitation letter, Consent Form, Project Statement, Instructions, Introductory Meeting Script, Discussion Guide and Modelling details can be found in online supplemental support document.

Pre-elicitation

Expert recruitment

We plan to recruit 14 clinician experts for the elicitation process. Previous research have recommended 3–20 experts,[11 12] while one empirical study found that six experts was optimal (by assessing the benefits of adding additional experts). The study assumed that all experts perform to the best of their abilities and experts are independent of each other.[12] Experts in our study will be preselected by peer nomination, including representation from neurologists, psychiatrists and geriatricians. Inclusion criteria include recognition by peers and the expert's willingness and availability to participate. Specifically,

Table 1 The timeline of the study

Stage	Tasks	Timeline
Pre-elicitation	Expert Enrolment and Consent	Sep 2020–Jun 2021
	Questionnaire Pilot	Dec 2020–Jun 2021
	Introduction meetings	Jun 2021–Jul 2021
Elicitation	Round 1 estimate	Jul 2021– Aug 2021
	Analysis	Aug 2021–Sep 2021
	Group discussion	Sep 2021–Nov 2021 (current)
	Round 2 estimate	Nov 2021–Jan 2022
Postelicitation	Aggregation	Jan 2022–Feb 2022
	Model training and validation	Feb 2022–Apr 2022
	Sensitivity analysis	Apr 2022–May 2022

two subject matter experts (EES and ZI) will reach out to experts in dementia (informally) and possibly gain informal consent, then a formal invitation letter will be sent out along with attachments (eg, consent form). In order to migrate the possible bias due to our enrolment process, we plan to enrol experts with different specialisations (AD, vascular dementia, dementia with Lewy bodies and frontotemporal dementia), locations across Canada (Alberta, British Columbia, Ontario and Nova Scotia), age groups (25–75 years), sex (males and females), as well as career stages(early, middle and senior). Furthermore, contact details will be provided and the experts have the liberty to initiate any conversations about the project before the first meeting.

Four-step questions and piloting

The experts will be asked about the relevant predictors (for MCI conversion) and rate the importance of each predictor. The participants of the elicitation will be provided with a potential list of predictors and will be allowed to add additional predictors that they consider important. It is advised that elicited quantities are preferred to be observable, for example, regression coefficients may be difficult to elicit directly from experts.[11] We will elicit 3-year dementia risk, which is consistent with the average rate of progression to dementia in individuals with MCI. From the resulting probabilities, survival and hazard functions can be constructed.[19] As suggested by the IDEA protocol, the four-step questions will be used, asking experts to provide a minimum, maximum and best guess for each quantity as well as a 'degree of belief'.[10] Questions ordered as such in order to encourage experts to consider a wider range of possibilities, mitigate anchoring and overconfidence.[12] The questions are framed as uncertainty about frequencies in a large population to approximate probabilities.[12 20] Two subject matter experts will review the drafted questions, to ascertain the questions that are free from linguistic ambiguity, appropriate in the domain, and can be completed in 30 min. The following a–d explains the proposed questionnaire (developed in Qualtrics).

a. Variable selection

First, clinicians are requested to rank the importance of each predictor (seven Likert scale: not at all, not, less, neutral, somewhat, moderately and very important), based on the given candidate predictors. Only the preselected predictors and ranked as at least somewhat important will be asked for further information.

Question: What do you think the importance of the listed variables are, in terms of predicting dementia progression from MCI?

Question: Are there any other predictors you think we need to consider? Please be specific and rank the added predictor (somewhat, moderately and very important).

b. Reference group selection

We assume that each predictor has a reference value or group, corresponding to the lowest (or lower) risk of dementia from MCI. The reference group (or value) is prespecified for each predictor based on literature; the clinicians will be asked if they agree with the default choices.

c. Median and interval assessment for the reference group

Cues: imagine that there are 100 MCI patients at baseline and with every predictor at the lowest as shown in the previous page. How many of them do you think will develop dementia in 3 years? Then four-step elicitation will be followed:

Q1. Realistically, what do you consider the lowest plausible number out of 100 MCI patients to develop dementia in 3 years will be?

Q2. Realistically, what do you consider the highest plausible number out of 100 MCI patients to develop dementia in 3 years will be?

Q3. Realistically, what is your best guess? (how many out of 100 MCI patients to develop dementia in 3 years)

Q4. How confident are you that your interval, from the lowest to highest, could capture the true value? Please enter a number between 50% and 100%

d. Median and interval assessment for each predictor

Cues: the following questions will ask about each predictor, the ones which you have selected as at least somewhat important before. We would like you to consider modifying only one predictor at a time, while the other predictors remain at the lowest risk. We will select either the Montreal Cognitive Assessment (MoCA) or Mini-Mental State Examination (MMSE) in the model. The two tests will not be included in one model. Taking age variable as an example ('age 55 years and younger' was selected in (b). Question: Recall that 'age 55 years and younger' was considered as the reference group. Now, we would like you to consider that only age variable has changed to '75 years old', and all other predictors are still in the lowest risk group. How many people do you think will develop dementia in 3 years? This will be followed by the four-step elicitation, the same as (c).

Introductory meeting

An introductory meeting will be organised to discuss the motivation for the project, objectives of the elicitation study, and the roles of the participants. The questionnaires will be reviewed to ensure clarity of the wording and the training content. The training content includes introduction on probability and uncertainty, elicitation process, heuristics and biases, information on how elicitation will be used and details of any assumptions or definitions that are used in the elicitation.

Elicitation process

The elicitation will involve a combination of individual and group level elicitation: clinicians will first complete the questionnaire independently and then engage in a facilitated group discussion. Across all steps, the clinicians could communicate with the elicitation team any time during the elicitation for clarification or assistance.

Round 1

The first round will start with sending a questionnaire link to the consented clinicians, containing the questions as previously explained. Individual level elicitations is recommended for HCDM setting,[12] where experts provide estimates independently without interacting with each other. During the process, clinicians are encouraged to reflect on their answers. The questionnaire may take about 30–60 min, so 4 weeks will be allotted for completion with another 4 weeks for late responses.

Analysis

Prior to the discussion meeting, feedback on the results of Round 1 for the clinicians will be prepared. The analyst will provide the variable importance rank for each predictor from each expert and the summed score for each predictor from all experts. This way, experts can compare their results with pooled results. In addition, we will standardise the reported credible intervals (eg, 90%) using linear extrapolation, so that clinicians view the uncertainties of all clinicians across questions on a consistent scale. The main purpose of the adjusted intervals at this stage is to allow for comparisons during the discussion phase.[10] The clinicians will be encouraged to change their estimates in round 2 if the extrapolation does not represent what they believe. Analysis will likely take about 2–4 weeks.

Discussion phase

A online discussion session in which clinicians (ie, participants) will discuss their perspectives, ranking of the predictors and other estimates they provided, and any concerns they may have about the elicitation process.[21] The discussion session is an important step in SEE. It helps to assess face validity of the elicitation questionnaire (ie, asking experts whether the elicited questionnaire is unambiguous and measuring what it is intended to measure). The discussion meeting will give experts the opportunity to see how their estimates differ from the average of the group and allow experts to update their estimates based on new information from other experts. It may allow us to manage biases, as well as collect information on how experts responded the questions and challenges during the process. It helps to evaluate whether the questions are comprehensible and enable experts to share their knowledge.

Round 2

Following the discussion, experts will be asked to independently consider revising their responses and make a second, anonymous and independent estimate for each question. The clinicians will be given 4 weeks for the second round, and another 4 weeks for late responses.

Postelicitation

Following the completion of the elicitation, the elicited clinicians' variable importance ranking and final answers for each predictor will be aggregated and shared with the group for final review. All steps taken and results collected during the elicitation will be documented. When eliciting judgements from multiple experts, it is important to have a single distribution that characterise experts' knowledge that can be used in sequence modelling through aggregation.[11 12] In this study, a linear pooling method with equal weights will be used as suggested in the HCDM protocol.

Handling cognitive bias and elicitation evaluation

During each step of the elicitation, we will work diligently to minimise influence from cognitive biases. Besides the introductory meeting, feedback will be encouraged during every step of the elicitation. We will evaluate validity of elicitation through face validity, managing biases, evaluating whether the questions are comprehensible, checking for inconsistencies and internal and external peer reviews.[12] Additionally, we will examine uncertainty in the elicited distribution by using a number of alternative distributions that fit the elicited summaries. Lastly, we will report and document the feasibility of the exercise, including challenges in the task and logistics associated with the elicitation.

Data management

Patient data

Patient data are from three different sources: the Prospective Registry for Persons with Memory Symptoms (PROMPT) of the Cognitive Neurosciences Clinic at the University of Calgary (UCalgary); the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the National Alzheimer's Coordinating Center's Uniform Data Set (NACC-UDS; naccdata.org). The PROMPT Registry was established in July 2010 and enrols patients referred to the Cognitive Neuroscience Clinic, operating in two urban tertiary care centres in Calgary, for assessment of suspected impairment in cognitive or behavioural function. Consecutive patients are approached for consent, and all patients attending the clinic are eligible to participate. To ensure complete follow-up, we linked PROMPT participants to Alberta healthcare administrative data for surveillance of new dementia diagnoses. There were 452 patients with MCI (any types) in PROMPT (up until April 2020), with age 55 years and older who had at least one follow-up after the baseline visit.

The primary objective of ADNI study is to test whether biomarkers can be combined with demographic and clinical data to measure the progression of MCI and early AD.[22] Participants included in ADNI were between 55 and 90 years of age, English or Spanish speakers, and were accompanied by study partners. We identified a sample of 598 individuals with MCI that was defined as having an MMSE score between 24 and 30, reported subjective complaints, objective memory deficits defined as Wechsler Memory Scale Logical Memory II scores below education-adjusted thresholds, and a Clinical Dementia Rating score of 0.5.

The NACC was established by National Institute on Aging (NIA)-funded Alzheimer's Disease Research Centers that recruit and collect data on individuals

with diverse cognitive functions (ranging from normal to dementia). The National Alzheimer's Coordinating Center's Uniform Data Set (NACC-UDS) is a longitudinal dataset that includes demographic and standardised clinical data collected approximately annually. All test centres administered standardised forms and informed consent was collected from all subjects and their informants. There were 1233 patients with MCI (V.3 of NACC), aged 55 years and older who had at least one follow-up after the baseline visit within 3 years. Detailed information on the cohort and neuropsychological battery of tests included in the UDS is described elsewhere.[23–25] The identification of the study cohort and data linkage can be found in online supplemental support document.

The main outcome was dementia of any cause. Dementia incidence during 3 years follow-up after the diagnosis of MCI will be treated as the event. MCI and dementia definition are based on standard outcome definitions, including Diagnostic and Statistical Manual of Mental Disorders and the NIA - Alzheimer's Association.[26 27] The preselected candidate predictors include demographics (age, sex, education, marital status, race, employment history, number of children, living status, location and socioeconomic status), medical history (diabetes, hypertension, traumatic brain injury, cardiovascular disease, cerebrovascular disease, dyslipidaemia, hypothyroidism, obstructive sleep apnoea and mood disorder), lifestyle factors (smoke, alcohol abuse, physically active, sleep quality and healthy diet), genetic factors (Apolipoprotein E (APOE)), AD biomarkers (cerebrospinal fluid (CSF) profile pattern and fluorodeoxyglucose (FDG) positron emission tomography (PET) findings), and cognitive screening scores (either the MoCA or the MMSE). Patient data were prospectively collected on prespecified case report forms.

All data (patient data and elicited data) will be deidentified before using the data for research purposes. Data will be stored password protected on a secure server at the UCalgary, only accessible through a password-protected UCalgary computer. Any hard copy questionnaires or notes will be stored at a secure file cabinet. Only the study team will be granted access to the data.

The models to be developed in this study will be trained using the PROMPT registry data and validated in the ADNI and NACC datasets.

Patient and public involvement

No patient involved.

ANALYSIS PLAN

For aim (1), we will rank the potential predictors according to the sum of each expert's rated scale for the importance of each predictor, where the highest score refers to the most important predictor from the experts.

For aim (2), we will evaluate the relative contribution of the elicited expert knowledge by comparing model performance for (1) model based on patient data alone and (2) model based on a combination of elicited prior knowledge and patient data. We will train the models in PROMPT and validate in ADNI and NACC. For model (1), Cox regression will be used to develop and validate the prediction model to examine 3-year dementia risk for MCI persons in the PROMPT registry data. The linear relationship between continuous candidate variables and outcome will be assessed using linear splines and restricted cubic splines. The assumption of proportional hazards will be assessed based on Schoenfeld residuals.[28] Candidate predictors will be selected based on evidence from literature review and the availability of the predictors in the training and validation datasets. The least absolute shrinkage and selection operator (LASSO) will be used to obtain the most parsimonious models, while retaining age and sex. Nested cross-validation will be used to tune the LASSO hyperparameter and estimate the predictive performance of the models. The predictive accuracy of model will be assessed using c-index as a measure of discrimination and calibration by graphically comparing the predicted and observed values based on flexible hazard regression approach.[29 30]

For model (2), we will integrate expert knowledge with patient data based on Bayesian Cox regression with normal prior distribution for the regression coefficients. Means and variances of normal priors for regression coefficients can be testimated directly from elicitation, but it is not suggested since it is difficult for clinicians to think about coefficients and give estimates.[20] Instead, we can make use of the relationship between the survival probabilities and regression coefficients via Cox regression to generate the prior distributions. For continuous risk factor, a piecewise linear function will be used. The variance (or SD) for each regression coefficient will be calculated twice using the upper and lower fractiles of the credible interval (from the four-step elicitation) and we will average the two SDs.[31] Gibbs sampling will be used to approximate the posterior distributions (details about the mathematical modelling are provided in online supplemental support document). The posterior means of the regression coefficients will be used to calculate the dementia risk scores. The predictive accuracy (discrimination and calibration) of model will be reported. We will compare the risk prediction model with and without clinician knowledge, in terms of discrimination and calibration.

Sensitivity analysis

Given that death (before the end of the 3 years period) can be a competing event, a Fine and Gray model, that treats death as a competing event will be trained and validated to examine the robustness of the conclusions of these analyses to the presence of competing risk. On the other hand, sensitivity analyses will be conducted to examine the robustness of the accuracy of models that integrate expert knowledge with patient data to different choices of prior distributions and methods of aggregating

elicited data. All analysis will be performed using SAS V.9.4[32] and R.[33]

The reporting of the study will follow the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis,[34 35] SEE protocols and Bayesian analysis reporting guidelines.[36 37] The anticipated timeline for the study is outlined in table 1.

DISCUSSION

Expert elicitation is particularly valuable when questions are highly uncertain and difficult to address via other methods.[38] Evidence from previous literature shows that inherent uncertainty is deeply embedded in medicine[39 40] and that judgement of clinicians is ultimately required.[41] There is no broadly accepted dementia risk score for use in individuals with MCI. There are few large, well-characterised cohorts of MCI derived from the general population. Prior information gathered through years of clinical experience becomes more important, since the information contributed by current data is limited.[42] Unlike most dementia risk scores derived from patient data, our study will incorporate clinician expert knowledge with patient data to predict dementia risk in persons with MCI. In addition, it is suggested to involve clinicians in model development, because an understanding of the clinical context is the key for implementation.[42] Lastly, we believe that clinicians may be trained well at reasoning with uncertainty since they tend to make difficult decisions daily.[11]

One major concern in the development of clinical risk score is the methodology that is used to determine the important risk factors to be included in the risk score. Although many variable selection procedures (eg, stepwise, penalised likelihood, boosting, resampling based, machine learning based feature selection) are available, a recent review has concluded that there is no state-of-the-art guideline for variable selection.[43] Many risk prediction guidelines[42 44–46] have recommended using clinical expert knowledge to guide the selection of predictors but such an approach might reinforce expert cognitive biases without fully capturing experts' uncertainties, especially when done in an informal manner. This proposed study will use a structured elicitation process to elicit important risk factors from a group of experts.

To the best of our knowledge, this is the first attempt to integrate expert knowledge with patient data n multivariable dementia risk prediction models. Prior information plays a fundamental role in Bayesian statistics, which can incorporate expert knowledge into the modelling naturally. The strength of this study includes its integration of combined information from both data and experienced clinicians: estimating regression coefficients on the combination of findings in the sample under study with external information from experts. We expect that involving clinicians in model development will help to improve knowledge translation and model implementation.[42] Moreover, the use Bayesian methodology based on Gibbs sampling and other Markov chain Monte Carlo techniques allow us to make inference, especially in small-sampled studies. This is particularly meaningful in dementia research, where data are still limited and the effect size from each predictor (for MCI conversion) is likely to be small. We believe that SEE is a useful tool in clinical risk prediction, where variable importance can be elicited from experts in a systematic and transparent way. Bayesian model updating is a natural process: our model (individual regression coefficients) could serve as prior information and combine with new patient data using the likelihood function and Bayes rule. Additionally, for future expert elicitation on dementia risk, our elicited data (expert belief) can serve as prior data, which can be incorporated with new elicited data to update expert's belief on dementia risk in individuals with MCI.

However, this study is not without limitations. The lack of patient data on all potential predictors identified by the clinician experts in the training or validation data cohorts might influence the accuracy of the proposed models and its generalisability to other populations. In addition to time and budget constraints, the difficulties of recruiting a representative sample of experts may limit access to SEE. Finally, the experts to be recruited in this study may not be representative of all knowledge in this field. Data for model development are from the PROMPT registry which consists of cohort of patients with MCI seen at a tertiary care centre. The PROMPT data are subject to referral biases which might limit the generalisability of the conclusions of our study to other populations. To address this limitation, the models developed will be validated in two external cohorts (NACC and ADNI) to enhance the generalisability of our study findings.

In summary, this proposed study will develop a robust dementia risk score that incorporates clinician expert knowledge with patient data for accurate risk stratification, prognosis and management of dementia.

ETHICS AND DISSEMINATION

This study has been approved by the University of Calgary Conjoint Health Research Ethics Board (REB19-0469). The NACC database itself is exempt from IRB review. All ADNI subjects or their proxies provided written informed consent. Deliverables of the proposed research involve journal publications, conference presentations, the development of R codes and a web-based risk score as knowledge translation.

Our study will be modified in our postelicitation stage because of this peer review. Specifically, we will use two additional (patient) datasets for validation, was informed by the need to address the generalisability of the proposed model. In addition, we will have an additional sensitivity analysis given that death (before the end of the 3-year period) can be considered as a competing event, a competing risk model, a Fine and Gray model, that treats death as a competing event will be trained and validated to examine the robustness of the conclusions of these analyses to the presence of competing risk.

Author affiliations
[1]Department of Clinical Neurosciences & Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada

[2]Department of Community Health Sciences & O'Brien Institute of Public Health, University of Calgary, Calgary, Alberta, Canada
[3]Department of Radiology, University of Calgary Cumming School of Medicine, Calgary, Alberta, Canada
[4]Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada
[5]Department of Psychiatry, University of Calgary Cumming School of Medicine, Calgary, Alberta, Canada

Twitter Tolulope Sajobi @dr_sajobi

Contributors TS, EES and MW are responsible for the conceptualisation of this study. TS and EES oversee MW's doctoral work including clinician expert recruitment, ethics application, the planned expert elicitation process and data analysis. MW designed the questionnaire, drafted the protocol and modified it according to coauthors' suggestions. Wang will analyse all data, interpret results, and summarise findings. EES, ZI and AG have provided content expertise in the design of the expert elicitation questionnaires including a potential list of risk factors for MCI conversions. EES and AG participated in the piloting of the elicitation question for face validity. TC provided methodological expertise on Bayesian risk prediction modeling. NDF, TC and TS had been providing methodological inputs into the proposed derivation and validation of risk prediction models. EES supervised data collection, creation of the study cohorts (both experts and patient data), linkage to provincial administrative databases, expert elicitation meeting and clinician engagement activities. All authors critically reviewed and provided suggestions for the revision of this protocol manuscript.

Funding Meng was supported by Alberta Innovates Graduate Student Scholarship, University of Calgary Cumming School of Medicine Graduate Scholarship, and Harley N. Hotchkiss Doctoral Scholarship in Neuroscience. Dr Sajobi is supported by the Natural Sciences and Engineering Research Council Disvoery Grant. The PROMPT registry is funded by the Kathy Taylor Chair in Vascular Dementia the University of Calgary with additional support from the Brain and Mental Health Research Clinics initiative of the Hotchkiss Brain Institute, University of Calgary. The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADCs (see support document for complete list). Data collection and sharing for this project was funded by the ADNI (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the NIA, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Meng Wang http://orcid.org/0000-0001-7955-6981
Eric E Smith http://orcid.org/0000-0003-3956-1668
Aravind Ganesh http://orcid.org/0000-0001-5520-2070
Tolulope Sajobi http://orcid.org/0000-0002-5696-5552

REFERENCES

1 Prince M, Wimo A, Guerchet M. World Alzheimer Report. In: *The global impact of dementia: an analysis of prevalence, incidence, cost and trends*. 87, 2015.
2 Public Health Agency of Canada. Dementia in Canada, Including Alzheimer's Disease. Available: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/dementia-highlights-canadian-chronic-disease-surveillance.html
3 Public Health Agency of Canada. Mapping connections: an understanding of neurological conditions in Canada, 2018. Available: https://www.canada.ca/en/public-health/services/reports-publications/mapping-connections-understanding-neurological-conditions.html
4 Tschanz JT, Welsh-Bohmer KA, Lyketsos CG, *et al*. Conversion to dementia from mild cognitive disorder: the Cache County study. *Neurology* 2006;67:229–34.
5 Lopez OL, Becker JT, Chang Y-F, *et al*. Incidence of mild cognitive impairment in the Pittsburgh cardiovascular health Study-Cognition study. *Neurology* 2012;79:1599–606.
6 Arvanitakis Z, Shah RC, Bennett DA. Diagnosis and management of dementia: review. *JAMA* 2019;322:1589–99.
7 Hou X-H, Feng L, Zhang C, *et al*. Models for predicting risk of dementia: a systematic review. *J Neurol Neurosurg Psychiatry* 2019;90:373–9.
8 Goerdten J, Čukić I, Danso SO. Statistical methods for dementia risk prediction and recommendations for future work: a systematic review. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 2019;5:563–9.
9 Use MMG. And abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National academy of Sciences* 2014;111:7176–84.
10 Hemming V, Burgman MA, Hanea AM, *et al*. A practical guide to structured expert elicitation using the idea protocol. *Methods Ecol Evol* 2018;9:169–80.
11 O'Hagan A, ed. *Uncertain Judgements: Eliciting Experts' Probabilities*. John Wiley & Sons, 2006.
12 Bojke L, Soares MFO, Fox A. Developing a reference protocol for expert elicitation in healthcare decision making. *Health Technology Assessment Reports*. *Published online* 2019.
13 Authority EFS. Guidance on expert knowledge elicitation in food and feed safety risk assessment. *EFSA Journal* 2014;12:3734.
14 Bruno Soares M, Dessai S. Exploring the use of seasonal climate forecasts in Europe through expert elicitation. *Clim Risk Manag* 2015;10:8–16.
15 Soares MO, Sharples L, Morton A, *et al*. Experiences of structured elicitation for model-based cost-effectiveness analyses. *Value Health* 2018;21:715–23.
16 Grigore B, Peters J, Hyde C, *et al*. Methods to elicit probability distributions from experts: a systematic review of reported practice in health technology assessment. *Pharmacoeconomics* 2013;31:991–1003.
17 Dallow N, Best N, Montague TH. Better decision making in drug development through adoption of formal prior elicitation. *Pharm Stat* 2018;17:301–16.
18 Hoff PD. *A First Course in Bayesian Statistical Methods*. Vol 580. *Springer* 2009.
19 Chaloner K, Church T, Louis TA, *et al*. Graphical elicitation of a prior distribution for a clinical trial. *The Statistician* 1993;42:341.
20 Cooke R. *Experts in uncertainty: opinion and subjective probability in science*. Oxford University Press on Demand, 1991.
21 Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia - meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr Scand* 2009;119:252–65.
22 Petersen RC, Aisen PS, Beckett LA. Alzheimer's Disease Neuroimaging Initiative (ADNI). *Neurology* 2010;74:201–9.

23 Weintraub S, Salmon D, Mercaldo N, *et al*. The Alzheimer's disease centers' uniform data set (UDS): the neuropsychologic test battery. *Alzheimer Dis Assoc Disord* 2009;23:91–101.
24 Beekly DL, Ramos EM, van Belle G. The National Alzheimer's Coordinating Center (NACC) Database: an Alzheimer disease database. *Alzheimer Dis Assoc Disord* 2004;18:270–7.
25 Morris JC, Weintraub S, Chui HC, *et al*. The uniform data set (UDS): clinical and cognitive variables and descriptive data from Alzheimer disease centers. *Alzheimer Dis Assoc Disord* 2006;20:210–6.
26 McKhann GM, Knopman DS, Chertkow H, *et al*. 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. *Alzheimer's & Dementia* 2011;7:263–9.
27 Albert MS, DeKosky ST, Dickson D, *et al*. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia* 2011;7:270–9.
28 GRAMBSCH PM, THERNEAU TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;81:515–26.
29 Stone CJ, Hansen MH, Kooperberg C. Polynomial splines and their tensor products in extended linear modeling. *The Annals of Statistics* 1997;25:1371–425.
30 Kooperberg C, Stone CJ, Truong YK. Hazard regression. *J Am Stat Assoc* 1995;90:78–94.
31 Kynn M. *Eliciting expert knowledge for Bayesian logistic regression in species habitat modelling*, 2005.
32 SAS Institute Inc. *SAS/ACCESS® 9.4*. Interface to ADABAS: Reference, 2013.
33 RStudio Team. RStudio: Integrated Development Environment for R. 2021.
34 Collins GS, Reitsma JB, Altman DG, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2014;350:g7594.
35 Moons KGM, Altman DG, Reitsma JB, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.
36 Sung L, Hayden J, Greenberg ML, *et al*. Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. *J Clin Epidemiol* 2005;58:261–8.
37 Depaoli S, van de Schoot R. Improving transparency and replication in Bayesian statistics: the WAMBS-Checklist. *Psychol Methods* 2017;22:240–61.
38 Butler AJ, Thomas MK, Pintar KDM. Systematic review of expert elicitation methods as a tool for source attribution of enteric illness. *Foodborne Pathog Dis* 2015;12:367–82.
39 Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. *JAMA* 2017;318:517–8.
40 Ashby D. Bayesian statistics in medicine: a 25 year review. *Stat Med* 2006;25:3589–631.
41 Knopman DS, Haeberlein SB, Carrillo MC, *et al*. The National Institute on Aging and the Alzheimer's Association Research Framework for Alzheimer's disease: Perspectives from the Research Roundtable. *Alzheimer's & Dementia* 2018;14:563–75.
42 Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. *Springer International Publishing* 2019.
43 Sauerbrei W, Perperoglou A, Schmid M, *et al*. State of the art in selection of variables and functional forms in multivariable analysis—outstanding issues. *Diagnostic and Prognostic Research* 2020;4:1–18.
44 Moons KGM, Kengne AP, Woodward M, *et al*. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 2012;98:683–90.
45 Harrell Jr FE. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer, 2015.
46 Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014;35:1925–31.