

Original Paper

Early Identification of Cognitive Impairment in Community Environments Through Modeling Subtle Inconsistencies in Questionnaire Responses: Machine Learning Model Development and Validation

Hongxin Gao¹, MSc; Stefan Schneider^{2,3,4,5}, PhD; Raymond Hernandez^{2,5}, PhD; Jenny Harris¹, PhD; Danny Maupin¹, PhD; Doerte U Junghaenel^{2,3,4,5}, PhD; Arie Kapteyn⁵, PhD; Arthur Stone^{2,3,5}, PhD; Elizabeth Zelinski⁴, PhD; Erik Meijer⁵, PhD; Pey-Jiuan Lee^{2,5}, MS; Bart Orriens⁵, PhD; Haomiao Jin¹, PhD

¹School of Health Sciences, University of Surrey, Guildford, United Kingdom

²Center for Self-Report Science, University of Southern California, Los Angeles, CA, United States

³Department of Psychology, University of Southern California, Los Angeles, CA, United States

⁴Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, United States

⁵Center for Economic and Social Research, University of Southern California, Los Angeles, CA, United States

Corresponding Author:

Haomiao Jin, PhD
School of Health Sciences
University of Surrey
Kate Granger Building
30 Priestley Road
Guildford, GU2 7YH
United Kingdom
Phone: 44 7438534086
Email: h.jin@surrey.ac.uk

Abstract

Background: The underdiagnosis of cognitive impairment hinders timely intervention of dementia. Health professionals working in the community play a critical role in the early detection of cognitive impairment, yet still face several challenges such as a lack of suitable tools, necessary training, and potential stigmatization.

Objective: This study explored a novel application integrating psychometric methods with data science techniques to model subtle inconsistencies in questionnaire response data for early identification of cognitive impairment in community environments.

Methods: This study analyzed questionnaire response data from participants aged 50 years and older in the Health and Retirement Study (waves 8-9, n=12,942). Predictors included low-quality response indices generated using the graded response model from four brief questionnaires (optimism, hopelessness, purpose in life, and life satisfaction) assessing aspects of overall well-being, a focus of health professionals in communities. The primary and supplemental predicted outcomes were current cognitive impairment derived from a validated criterion and dementia or mortality in the next ten years. Seven predictive models were trained, and the performance of these models was evaluated and compared.

Results: The multilayer perceptron exhibited the best performance in predicting current cognitive impairment. In the selected four questionnaires, the area under curve values for identifying current cognitive impairment ranged from 0.63 to 0.66 and was improved to 0.71 to 0.74 when combining the low-quality response indices with age and gender for prediction. We set the threshold for assessing cognitive impairment risk in the tool based on the ratio of underdiagnosis costs to overdiagnosis costs, and a ratio of 4 was used as the default choice. Furthermore, the tool outperformed the efficiency of age or health-based screening strategies for identifying individuals at high risk for cognitive impairment, particularly in the 50- to 59-year and 60- to 69-year age groups. The tool is available on a portal website for the public to access freely.

Conclusions: We developed a novel prediction tool that integrates psychometric methods with data science to facilitate “passive or backend” cognitive impairment assessments in community settings, aiming to promote early cognitive impairment detection.

This tool simplifies the cognitive impairment assessment process, making it more adaptable and reducing burdens. Our approach also presents a new perspective for using questionnaire data: leveraging, rather than dismissing, low-quality data.

(*JMIR Form Res* 2024;8:e54335) doi: [10.2196/54335](https://doi.org/10.2196/54335)

KEYWORDS

machine learning; artificial intelligence; cognitive impairments; surveys and questionnaires; community health services; public health; early identification; elder care; dementia

Introduction

Background

Cognitive impairment encompasses both mild cognitive impairment (MCI) and dementia. MCI is considered to be an intermediate state between normal aging and dementia [1] and may initially not interfere with daily life but potentially progress to dementia and cause significant functional impairment [2]. Recent biomedical research continues to yield novel treatments, such as lecanemab (Eisai and Biogen) and Donanemab (Eli Lilly and Company) [3,4], which can effectively slow cognitive decline if applied at its early stages (MCI or mild dementia). This underscores the importance of detecting cognitive impairment early for the cognitive health of older adults.

Nevertheless, cognitive impairment underdiagnosis remains a concerning issue around the world. In the United States, cognitive impairment assessment is included in Medicare's Annual Wellness Visit [5], but less than 20% of older adults use this preventive service [6,7]. In countries such as the United Kingdom, routine cognitive impairment assessment is not recommended for all older adults [8]. This underuse or lack of provision of preventive cognitive impairment assessment, combined with challenges faced by primary care providers in identifying early cognitive impairment on top of managing other health issues, often leads to cognitive impairment underdiagnosis [9,10].

Considering these challenges, health professionals working in community environments such as community health and social workers (CHSWs) have become increasingly important in cognitive health management for older adults. CHSWs primarily focus on promoting overall well-being within community settings, such as nonprofit organizations, government agencies, and community centers, and often work with specific populations such as older adults. The World Health Organization's Global Action Plan has made dementia a priority for public health action, accelerating a paradigm shift in the prevention and care of cognitive impairment from clinical systems toward family and community-based services [9]. This shift emphasizes the essential roles of CHSWs in identifying and referring older adults with early cognitive impairment symptoms to clinical care providers for further assessment and treatment.

Although CHSWs have a growing role in cognitive health management, several challenges may hinder the implementation of timely cognitive impairment assessment in community environments. Administering a cognitive impairment assessment, even with a so-called "brief" tool such as the Telephone Interview for Cognitive Status (which consists of 11

items and takes about 10 minutes with a trained administrator), is easier said than done [11]. Given the broad range of responsibilities of CHSWs, integrating cognitive impairment assessments into their existing tasks could stretch their capacity and limit the time available for other essential services. Besides, CHSWs may encounter challenges related to a lack of specific training and expertise in cognitive health assessment, because their educational background often focuses on social work and public health. This knowledge gap could decrease their ability to accurately identify symptoms of cognitive impairment and confidently refer individuals for further clinical evaluation. Finally, stigma and cultural barriers surrounding cognitive impairment may also present challenges for CHSWs [12]. Older adults or their families might be hesitant to disclose cognitive impairment symptoms within community settings, leading to inaccurate results of cognitive impairment assessment [12].

Low-Quality Response in Questionnaires

In conducting surveys or questionnaires, researchers usually expect respondents to answer questions honestly and accurately. However, some respondents might lack the cognitive capability to think carefully and answer each question genuinely, leading to a lower quality of response [13-16]. The definition of low-quality response (LQR) in this paper is relatively neutral and is not motivated by intentional deception, but rather by invalid, unreliable, or erroneous responses [17,18]. Manifestations of the LQR are diverse [19-21], including but not limited to (1) skipping questions—respondents may selectively answer and bypass certain queries; (2) contradictory answers—for similar questions, respondents might provide conflicting responses; (3) oversimplified responses—for instance, always choosing "agree" or "do not know" for all questions; and (4) inaccurate or unreliable answers. This could be due to the respondents not fully understanding the question or lacking the time or motivation to ponder over it.

Completing a questionnaire is a task that requires multiple cognitive functions working in synergy [22], such as attention, working memory, and executive functions. Respondents with subtle cognitive deficits might resort to suboptimal answering strategies to cope with the psychological demands of responding to questions [17,23]. Some studies have found that older adults with cognitive deficits may exhibit stronger signals of LQR during questionnaires [19-21,24]. For instance, they might more frequently skip questions or opt for the "do not know" answer. Early cognitive decline might first manifest in complex tasks in the daily life of people, among which, completing a questionnaire is one.

Gaps in Research

Recent research on early identification of cognitive impairment has been using advanced computational statistics techniques such as machine learning to develop predictive models of cognitive impairment [25-28]. Nevertheless, challenges persist in effectively implementing such cognitive impairment predictive models within community environments. Existing studies on machine learning and cognitive impairment classification mostly used clinical biomarkers such as magnetic resonance imaging and cerebrospinal fluid analysis as predictors [29-31]. These approaches are not feasible in community settings due to their reliance on high-cost or invasive methods for obtaining predictor information. Furthermore, most current studies on machine learning and cognitive impairment classification have been based on small clinical samples (sample size typically ranged from 140-550), where resampling or case-controlling techniques were widely used to derive analytic samples with balanced cognitive impairment and noncognitive impairment proportions [32-34]. The proportion of cognitive impairment cases in such datasets can be significantly higher than the estimated prevalence (about 15%-20%) observed among older adults in communities [35]. While these resampling and case-controlling techniques may enhance the model's sensitivity to the minority class (ie, the cognitive impairment) [36,37], existing research has shown that reliance on these techniques can lead to discrepancies in the distribution of data and introduce model bias [38,39].

It is worth noting that recent studies have started to explore the use of large community-based samples with questionnaire data to develop machine learning models for cognitive impairment classification [40-43]. However, there remain several important gaps in this genre of research. First, issues related to LQRs, such as nonresponse, extreme answers, and acquiescence, are ubiquitous in the questionnaire data of older adults [19,44-46]. Direct use of questionnaire response data without proper handling of the potential LQR issue may reduce the validity of the prediction models. Second, the number of predictors used in these studies was generally substantial (approximately 22-44),

which can make rapid data collection and cognitive impairment assessment challenging. Lastly, widely used machine learning techniques such as regularization are not designed to produce unbiased predictions, because bias-variance trade-off is an essential component of these techniques [47]. Consequently, the predicted probability from these models does not necessarily reflect the true likelihood of cognitive impairment. This poses a challenge for end users, such as CHSWs, in selecting an appropriate threshold. Existing studies often fall short in providing guidance on determining this threshold.

This Study

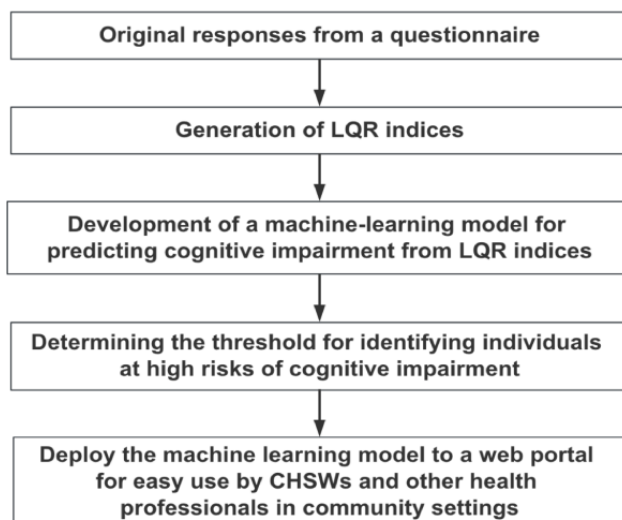
This study aimed to develop a novel machine learning tool for the early identification of cognitive impairment in community settings using psychometric indices generated from subtle inconsistencies in questionnaire responses. A distinctive feature of this study is that rather than filtering out the original questionnaire responses due to potential LQR issues (eg, eliminating low-quality answers), we relied on LQR indices, derived from psychometric methods, as predictors in our machine learning models to classify cognitive impairment. In community environments, CHSWs routinely administer brief questionnaires to identify well-being issues [48], generating a rich amount of questionnaire response data despite that the contents of these questionnaires are usually not directly related to assessing cognitive impairment. This study will investigate whether the relationships between LQR and cognitive impairment could be exploited to develop a machine learning tool for early cognitive impairment identification.

Methods

Study Design

An overview of the methodology adopted in this study is provided in [Figure 1](#). We followed the guidelines outlined in the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) statement for reporting the development and validation of the multivariate predictive models ([Multimedia Appendix 1](#)) [49].

Figure 1. Overview of the workflow and methods for the proposed tool development. CHSW: community health and social worker; LQR: low-quality response.



Study Sample

Development of the machine learning tool was based on individuals aged 50 years and older from the Health and Retirement Study (HRS), a community-based longitudinal aging study involving a representative sample of older adults in the United States [50]. As a part of this study, demographics including race and ethnicity are collected using surveys. In 2006 (wave 8) and 2008 (wave 9), the HRS administered the leave-behind survey, which explored various aspects of participants' psychological well-being and life experiences [51,52]. A random half (n=8681, 51.13%) of participants completed the survey in 2006, and the remaining participants (n=8296, 48.87%) completed it in 2008. Our analysis included participants from the leave-behind survey, excluding those whose questionnaires were completed by proxy, those with missing responses to all scale items, or those lacking key follow-up records (such as survival status). This resulted in a total of 12,942 individuals included in our analysis. While our analysis examined all 21 available questionnaires in the HRS leave-behind survey (Multimedia Appendix 2) [51], 4 questionnaires, that is, the optimism, hopelessness, life purpose, and life satisfaction scales, were used as the primary data source to develop the machine learning tool. These questionnaires were brief and assessed aspects that are closely related to the common focus of CHSWs on overall well-being as a positive psychological attitude, satisfaction with life, clear life goals, and optimism significantly contribute to enhancing the quality of life, extending the anticipated lifespan, and reducing mortality risk in those aged 50 years or older [53-56].

LQR Indices

To model subtle inconsistencies in questionnaire responses as predictors in our machine learning model, we created two types of LQR indices using psychometric methods. For each questionnaire, we first fitted a graded response model—a common psychometric model based on the item response theory—to the data to obtain an estimate of the respondents' "latent trait" score [57]. We did this to remove information related to content (eg, people's optimism levels) and to isolate statistically "misfitting" response patterns indicative of LQR regardless of question content. The first index, the squared residuals, was then derived by calculating the squared differences between the observed questionnaire item responses and the statistically expected responses given a respondent's latent trait [58]. Larger differences between observed and expected responses indicate severe problems with LQR. The second index was defined as the probability of a respondent choosing the observed response given their latent trait [59]. Selecting statistically fewer probable responses can indicate problems with LQR. The LQR indices were generated using R (version 4.2.2; R Foundation), mirt (version 1.38.1; York University), and tidyverse (version 2.0.0; RStudio) [59,60].

Predicted Outcomes

Cognitive impairment status was ascertained using the validated criteria developed by Langa et al [61] for the HRS. In brief, a cognitive status score was calculated from cognitive tests of immediate and delayed word recall, an attention and working memory task (serial sevens), and counting backward from 20.

The total score has a range from 0 to 27 with higher scores indicating better cognitive performance. Thresholds of dementia (0-6 points), cognitively impaired with no dementia (7-11 points), and cognitively normal (12-27 points) were applied [61]. The cognitively impaired with no dementia and dementia categories were then combined into a single cognitive impairment category to indicate whether the individual currently has cognitive impairment. To investigate the longitudinal prognostic utility of the LQR indices, we also created a secondary predicted outcome in the supplemental analysis, that is, dementia or mortality in the next 10 years, which was derived from the HRS follow-up records.

Model Development and Training

Using the LQR indices derived from each questionnaire, we trained and compared seven machine learning models for the binary classification task of forecasting cognitive impairment. Specifically, we used the multilayer perceptron (MLP) technique. MLP is a feed-forward neural network that is structured with input, hidden, and output layers [62]. Compared to conventional machine learning, it possesses robust adaptive learning capabilities, enabling the automatic extraction of abstract information from raw predictor variables without the need for labor-intensive manual transformations [62]. Moreover, the MLP is particularly adept at discerning nonlinear relationships between predictors and outcomes [63], facilitating the modeling of potential complex associations between the LQR indices and cognitive impairment. In contrast to certain advanced hybrid deep learning models, the MLP boasts a lighter architecture and parameter setup, which reduces computational burdens [64]. In addition to training the models using only LQR indices, another version of the models was developed using the LQR indices in combination with two easily obtained demographic variables, that is, age and gender, as predictors.

The dataset had 1.25% missing values (3898 of 310,608 data points), with 1499 of 12,942 (11.58%) participants having at least one missing item. These missing values were imputed using a regression-based iterative imputation method (details in Multimedia Appendix 3) [65]. The imputed data was then split into training (n=9059, 70%) and testing sets (n=3883, 30%) using stratified random sampling to ensure a consistent cognitive impairment prevalence (n=1641, 18.11% for the training set and n=704, 18.1% for the testing set) across both sets. We used batch normalization and class weighting strategies during the training process [66,67]. This allowed our model to support raw data input without scaling and resampling steps.

The MLP model comprised four hidden layers with a rectified linear unit activation function [68]. To reduce overfitting, we added dropout layers after each hidden layer [69]. The output layer used a sigmoid activation function with an initialization strategy to accelerate model convergence [68,70]. During the training phase, we used 10-fold cross-validation on the training sets [71]. Model weights and biases were adjusted based on the binary cross-entropy loss. The data were further split into 80% (n=7247) training subsets and 20% (n=1812) validation subsets in each fold. The area under the curve (AUC) was monitored on the validation subset to guide mechanisms such as early stopping [72]. We also used the Adam optimizer with a dynamic

learning rate to optimize the training process, ensuring that only the best-performing model was retained for the final evaluation of the test dataset [66,73]. Hyperparameter selection was initially guided by KerasTuner (Google) [74]. Based on its recommended values, extensive iterative experiments were conducted to ascertain the selected configurations, as listed in [Multimedia Appendix 4](#). The model training was carried out using Python (version 3.9.16; Python Software Foundation), TensorFlow (version 2.12.0; Google, Google Brain), and scikit-learn (version 1.2.1; French Institute for Research in Computer Science and Automation) [66,75,76]

We compared the performance of the MLP to 6 different predictive models. These included 2 classical machine learning models (logistic regression and decision trees) [77,78], 2 ensemble learning models (XGBoost and LightGBM) [79,80], and 2 hybrid deep learning models bidirectional-gated recurrent unit and convolutional neural network-long short-term memory [81-83]. We compared bidirectional-gated recurrent unit and convolutional neural network-long short-term memory models due to the potential dependencies among LQR indices. The LQR indicators are derived from multi-item questionnaires, where multiple items are designed to load onto a single latent factor. Consequently, there may be correlations between LQR indices within items from the same questionnaires. For the classical machine learning and ensemble learning models, we used grid search to determine the optimal hyperparameters. In the case of the hybrid deep learning models, we adopt the same training approach as the MLP and a similar architecture for the fully connected layers. All models were trained by 10-fold cross-validation and evaluated on the same unseen test dataset.

Threshold Determination

We determined the threshold for cognitive impairment risk scores predicted by the model based on the ratio of underdiagnosis cost to overdiagnosis cost, aiming to minimize the total cost of underdiagnosis and overdiagnosis [84]. Underdiagnosis refers to when the assessment tool identifies an individual with cognitive impairment as cognitively normal, while overdiagnosis refers to when the tool identifies a cognitively normal individual as having cognitive impairment. When applying the cognitive impairment assessment tool in community settings, individuals whose predicted risk scores are higher than the threshold would be referred to follow-up clinical cognitive impairment assessment. Thus, the cost of overdiagnosis would be equal to the cost of conducting one follow-up clinical cognitive impairment assessment. The cost of underdiagnosis would be more profound, including the cost of not receiving timely clinical assessment and treatment of cognitive impairment. Users of the cognitive impairment assessment tool may choose different cost ratios tailored to the needs of their practices. A larger cost ratio would be appropriate when the follow-up clinical assessment and treatment of cognitive impairment is readily available. With the emergence

of effective medications such as lecanemab and donanemab [3,4], this cost ratio would more likely to become larger in the future.

Evaluation of Model Performance and Efficiency

Predictive accuracy of the assessment tool was evaluated using the AUC value. AUC values range from 0 to 1, with a higher score indicating better predictive performance [72]. Moreover, the efficiency of using the tool in identifying individuals at high risk of cognitive impairment was evaluated using an efficiency curve approach. The first step was to calculate the proportion of individuals referred to further clinical assessment for cognitive impairment based on predicted results from the tool, which served as a proxy measure of the resources required to conduct follow-up clinical assessment and treatment. Then, the proportion of individuals with cognitive impairment identified by the follow-up clinical assessments was calculated, which acted as a proxy measure for the desired outcome of the tool's implementation and subsequent cognitive impairment assessment. These proportions were computed at different predicted risk score thresholds from the tool, and the paired proportion values were plotted on a curve. The tool's efficiency was then compared with plausible rule-based screening strategies, which included conducting cognitive impairment assessments for all individuals aged 65 years and older, or those with cardiovascular diseases such as high blood pressure, heart diseases, possible stroke or transient ischemic attack, or diabetes that is one of major risk factors for cardiovascular disease. This approach of comparing screening efficiency based on proxy measures of needed resources and desired outcomes has been previously implemented by research to compare machine learning-based assessment against rule-based screening strategies [85].

Ethical Considerations

This study is approved by the University of Southern California institutional review board (UP-22-00147) and the University of Surrey Research Integrity and Governance Office (FHMS 21-22 216 EGA).

Results

Sample characteristics are described in [Table 1](#). In addition, [Multimedia Appendix 5](#) presents sample characterization of the training and testing sets. There were 9059 individuals in the training dataset with an average age of 68.9 (range 50-104.6, SD 9.85) years, of which 18.11% (n=1641) had cognitive impairment according to the Langa-Weir criteria. The testing set included 3883 individuals. The average age of the testing sample was 68.8 (range 50-100.7, SD 9.72) years and 18.1% (n=704) in the testing set had current cognitive impairment. The comparison between training and testing sets showed no significant differences.

Table 1. Sample characteristics (N=12,942).

Variables	Participants, n (%)
Age (years)	
50-59	2613 (20.19)
60-69	4113 (31.87)
70-79	3935 (30.4)
80 and older	1938 (14.97)
Gender	
Female	7357 (56.85)
Male	5242 (40.5)
Marital status	
Married	8112 (62.68)
Not married	4486 (34.66)
Race	
African American	1511 (11.68)
White	10,588 (81.81)
Other	500 (3.86)
Ethnicity	
Hispanic	903 (6.98)
Not Hispanic	11,696 (90.37)
Education level	
High school and below	6833 (52.8)
Some college	2918 (22.55)
College graduate and above	2847 (22)
Self-reported diseases	
High blood pressure	7029 (54.31)
Diabetes	2402 (18.56)
Heart disease	2963 (22.89)
Possible stroke or transient ischemic attack	947 (7.32)
With 2 or more of the above diseases	3728 (28.81)

The MLP has the highest predictive performance on the testing dataset compared to other predictive methods. As shown in [Multimedia Appendix 6](#), when using the LQR indices as predictors, the MLP ranked first or tied for first in AUC values for 16 of the 21 questionnaires. After including age and gender as additional predictors, the MLP ranked first or tied for first in AUC for 18 of the 21 questionnaires. As a result, we chose the MLP as our predictive model. In addition, we evaluated the performance of using raw questionnaire responses as predictors in logistic regression models as baseline performance. Among the 21 questionnaires, the baseline model's AUC ranged from 0.47 to 0.64, with a mean of 0.56 (SD 0.04) and a median of 0.55 (IQR 0.04). The performance of using raw questionnaire responses for cognitive impairment prediction slightly outperformed random guessing (AUC=0.5), but was inferior to using LQR indices for prediction in our machine learning models.

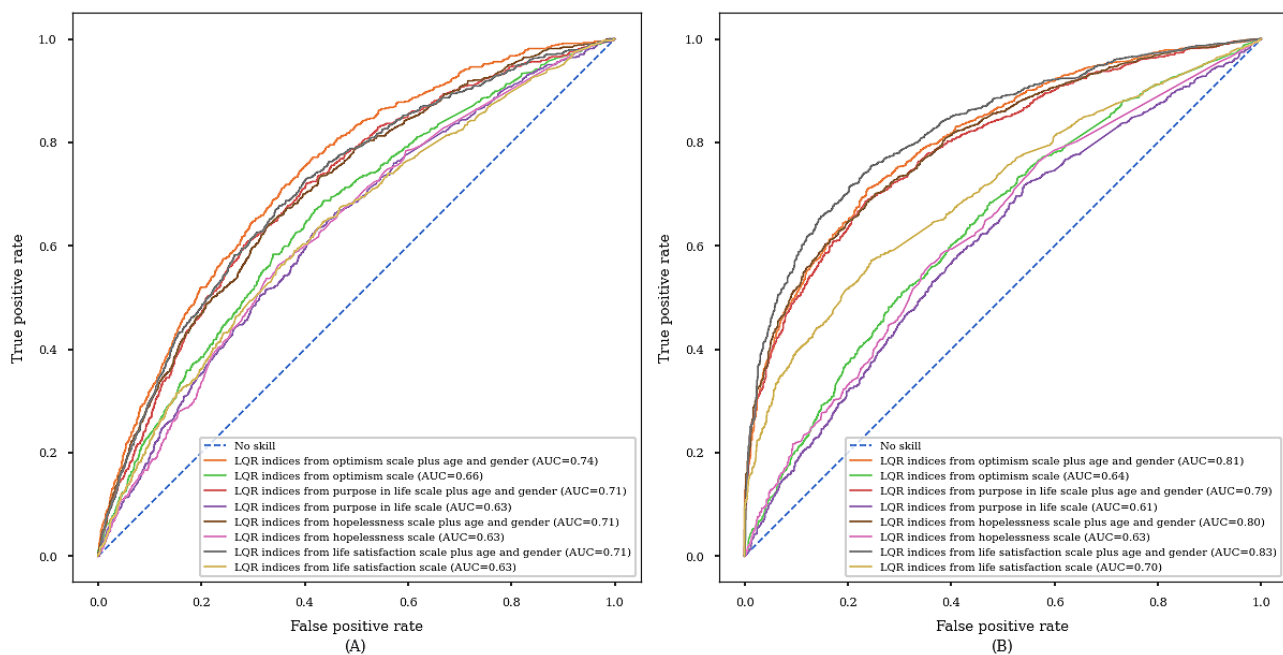
Predictive accuracy of the MLP models is summarized in [Multimedia Appendix 7](#) and AUC curves are shown in [Figure 2](#). The AUC values ranged from 0.63 to 0.66 across the 4 questionnaires when using LQR indices as predictors. The AUC values were improved to the range of 0.71 to 0.74 when age and gender were included as additional predictors. The AUCs were similar across the 4 questionnaires, and no evident signs of overfitting were observed during the training and validation phases ([Multimedia Appendix 8](#)). The LQR indices derived from the optimism scale performed the best among all 21 available questionnaires ([Multimedia Appendix 6](#)). The AUC for predicting dementia or mortality in the next ten years ranged from 0.61 to 0.70 and improved to 0.80 to 0.83 when age and gender were included as additional predictors.

Additionally, the evaluation of different thresholds revealed that as the cost ratio of underdiagnosis to overdiagnosis of cognitive impairment increased, the proportions of individuals referred for further assessment and the identification of cognitive

impairment cases also increased. The results demonstrated that the performance differences between models with and without age and gender as additional predictors became smaller as the cost ratio increased. For instance, with a cost ratio of underdiagnosis to overdiagnosis of two, the model using LQR indices from the optimism scale resulted in 3% (n=116) of all individuals being referred for further clinical cognitive impairment assessment and 7% (n=49) of all cognitive impairment cases being identified. On the other hand, the model using LQR indices from the same scale plus age and gender as predictors led to 10% (n=388) of individuals referred for further assessment and 24% (n=169) of cognitive impairment cases identified. When the cost ratio increased to four, the model

using LQR indices alone resulted in 39.99% (n=1533) of individuals being referred and 59.9% (n=422) of cognitive impairment cases being identified, and the model using LQR indices plus age and gender led to 35% (n=1359) of individuals being referred and 63.1% (n=444) of cognitive impairment identified. As mentioned above, the advancements in dementia treatment imply increased cost ratios of underdiagnosis to overdiagnosis, suggesting that the thresholds corresponding to larger cost ratios are more relevant for future practices. In light of this, a cost ratio of 4 was used as the default choice for the tool, because it provides a reasonable trade-off between underdiagnosis and overdiagnosis.

Figure 2. (A) AUC curves for predicting current cognitive impairment and (B) AUC curves for predicting dementia or mortality in the next 10 years. AUC: area under the curve.



Compared to rule-based screening strategies, the cognitive impairment assessment tool showed greater efficiency in identifying individuals at high risk of cognitive impairment. As shown in Figure 3, the efficiency curves corresponding to the four questionnaires were all located to the left of rule-based strategies, indicating higher efficiency as fewer resources were required to achieve the same desired outcome. Moreover, the tool offered greater flexibility as users could choose thresholds depending on the needs of their practices. We selected model thresholds at a default cost ratio of 4 for comparison between the two strategies, as shown in Table 2. Except for the rule “high blood pressure,” where the screening efficiency is relatively balanced, the screening efficiency under the other rules is generally polarized, making it difficult to reach a reasonable trade-off. For instance, although the rule “age ≥ 65 ” years leads to 84.9% (n=598) of this study’s sample with cognitive impairment being identified, it also consumes the most resources (n=2757, 71% of the sample will need to receive follow-up assessments), almost twice as much as the machine learning-based screening strategies. The machine learning methods are more efficient in identifying individuals with cognitive impairment. For instance, in the model based on the

optimism scale, 35% (n=1359) of the sample needs to receive follow-up assessments to ensure that 63.1% (n=444) of individuals with cognitive impairment are identified. The proportion receiving follow-up assessments is approximately 51% and 41% less compared to the “age ≥ 65 ” years and “high blood pressure” rules, respectively, yet it achieves an output (ie, the proportion of individuals with cognitive impairment identified) that is approximately 75% and 91% of theirs. This indicates that machine learning approaches are more efficient per resource usage and capable of achieving a higher identification ratio with less resource input.

Moreover, we explored the performance of LQR indices as predictors across different age groups, and the results are presented in Table 3. Overall, the cognitive impairment assessment tool demonstrated greater predictive accuracy and efficiency in the 50- to 59-year and 60- to 69-year age groups. Notably, in the optimism scale, the LQR indices consistently exhibited the best predictive performance (AUC=0.72) in the 50- to 59-year age group, regardless of whether age and gender were included as predictors. The tool performed better in predicting cognitive impairment among relatively younger older

adults, suggesting that LQR indices may be more sensitive to early cognitive deficits.

Finally, the cognitive impairment assessment tool was deployed to a web portal for free public access (URL available in the reference list [86]). As shown in Figure 4, the tool allows users to enter responses from one of the four questionnaires examined

in this study, set up thresholds by choosing a cost ratio (default is 4), predict risk scores of cognitive impairment, and generate recommendations for whether follow-up cognitive health assessment would be suggested. Users can choose a different cost ratio based on their circumstances. The tool can also be expanded using the same modelling approach described in this paper to incorporate additional questionnaires in the future.

Figure 3. Comparison of machine-learning-based and rule-based identification efficiency for high-risk individuals with cognitive impairment.

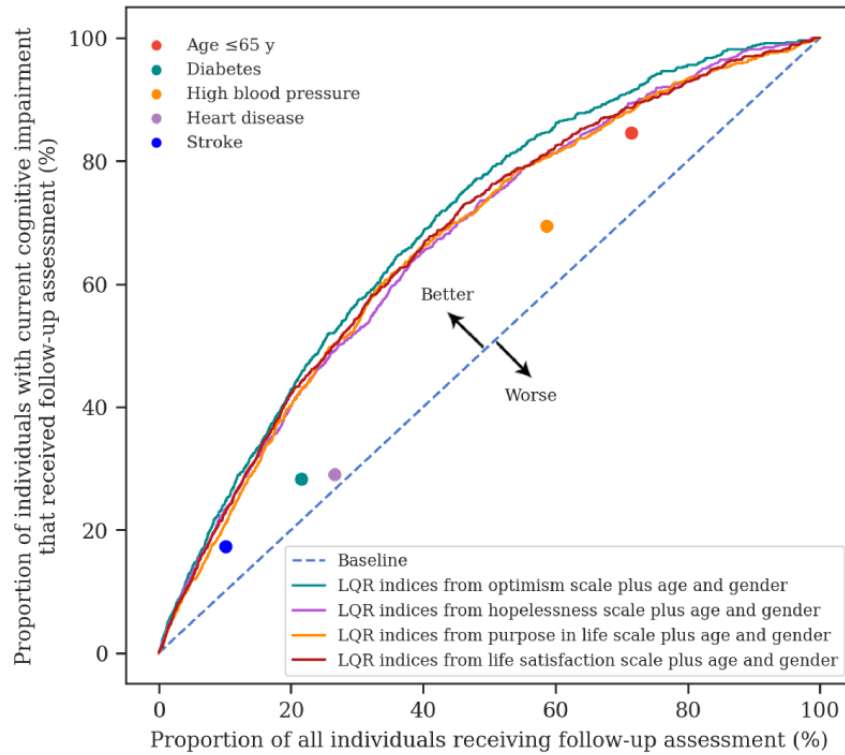


Table 2. Comparison of identification efficiency for high-risk individuals with cognitive impairment between machine learning-based and rule-based approaches under a cost ratio of 4.

Rule or model name	Proportion of all individuals receiving follow-up assessment, n (%)	Proportion of individuals with current cognitive impairment that received follow-up assessment, n (%)
Age ≥65 years	2757 (71)	598 (84.9)
Diabetes	854 (22)	197 (28)
High blood pressure	2291 (59)	486 (69)
Heart disease	1048 (26.99)	204 (29)
Stroke	388 (10)	120 (17)
LQR ^a indices from optimism scale plus age and gender	1359 (35)	444 (63.1)
LQR indices from purpose in life scale plus age and gender	1281 (32.99)	422 (59.9)
LQR indices from hopelessness scale plus age and gender	1553 (39.99)	458 (65.1)
LQR indices from life satisfaction scale plus age and gender	1243 (32.01)	408 (58)

^aLQR: low-quality response.

Table 3. Comparison of performance and efficiency of the LQR^a indices as predictors for predicting current cognitive impairment across scales and age groups under a cost ratio of 4.

Models and age groups (years)	AUC ^b	Sample, n (%)	Sample with cognitive impairment, n (%)	Individuals receiving follow-up assessment, n (%)	Individuals with current cognitive impairment that received follow-up assessment, n (%)
LQR indices from optimism scale plus age and gender					
50-59	0.72	790 (20.4)	73 (9)	65 (8)	22 (30)
60-69	0.7	1268 (32.66)	154 (12.2)	214 (16.9)	51 (33)
70-79	0.66	1256 (32.35)	257 (20.5)	588 (46.8)	170 (66.2)
80 and above	0.62	569 (14.7)	220 (38.7)	529 (93)	210 (95.5)
LQR indices from purpose in life scale plus age and gender					
50-59	0.64	790 (20.4)	73 (9)	41 (5)	9 (12)
60-69	0.66	1268 (32.66)	154 (12.2)	167 (13.2)	48 (31)
70-79	0.61	1256 (32.35)	257 (20.5)	562 (44.8)	153 (59.5)
80 and above	0.56	569 (14.7)	220 (38.7)	533 (93.7)	212 (96.4)
LQR indices from hopelessness scale plus age and gender					
50-59	0.62	790 (20.4)	73 (9)	87 (11)	13 (18)
60-69	0.65	1268 (32.66)	154 (12.2)	193 (15.2)	42 (27)
70-79	0.62	1256 (32.35)	257 (20.5)	693 (55.2)	184 (71.6)
80 and above	0.61	569 (14.7)	220 (38.7)	566 (99.5)	218 (99.1)
LQR indices from life satisfaction scale plus age and gender					
50-59	0.63	790 (20.4)	73 (9)	37 (5)	9 (12)
60-69	0.64	1268 (32.66)	154 (12.2)	151 (11.9)	33 (21)
70-79	0.66	1256 (32.35)	257 (20.5)	516 (41.1)	156 (60.7)
80 and above	0.57	569 (14.7)	220 (38.7)	549 (96.4)	212 (96.4)
LQR indices from optimism scale					
50-59	0.72	790 (20.4)	73 (9)	309 (39.1)	53 (73)
60-69	0.71	1268 (32.66)	154 (12.2)	489 (38.6)	104 (67.5)
70-79	0.62	1256 (32.35)	257 (20.5)	527 (42.0)	139 (54.1)
80 and above	0.59	569 (14.7)	220 (38.7)	277 (48.7)	128 (58.2)
LQR indices from purpose in life scale					
50-59	0.67	790 (20.4)	73 (9)	321 (40.6)	47 (64)
60-69	0.65	1268 (32.66)	154 (12.2)	508 (40.1)	91 (59)
70-79	0.6	1256 (32.35)	257 (20.5)	568 (45.2)	151 (58.8)
80 and above	0.54	569 (14.7)	220 (38.7)	346 (60.8)	145 (65.9)
LQR indices from hopelessness scale					
50-59	0.65	790 (20.4)	73 (9)	261 (33)	38 (52)
60-69	0.67	1268 (32.66)	154 (12.2)	404 (31.9)	83 (54)
70-79	0.6	1256 (32.35)	257 (20.5)	433 (34.5)	120 (46.7)
80 and above	0.6	569 (14.7)	220 (38.7)	245 (43.1)	116 (52.7)
LQR indices from life satisfaction scale					
50-59	0.66	790 (20.4)	73 (9)	291 (36.8)	42 (58)
60-69	0.64	1268 (32.66)	154 (12.2)	507 (40)	89 (58)
70-79	0.65	1256 (32.35)	257 (20.5)	541 (43.1)	159 (61.9)

Models and age groups (years)	AUC ^b	Sample, n (%)	Sample with cognitive impairment, n (%)	Individuals receiving follow-up assessment, n (%)	Individuals with current cognitive impairment that received follow-up assessment, n (%)
80 and above	0.58	569 (14.7)	220 (38.7)	276 (48.5)	121 (55)

^aLQR: low-quality response.

^bAUC: area under the curve.

Figure 4. A web portal allowing free public use of the machine-learning tool for cognitive impairment assessment based on subtle inconsistencies in questionnaire responses.

Discussion

Principal Findings

In this study, we developed a novel machine learning tool for predicting cognitive impairment based on LQR indices derived from subtle inconsistencies in questionnaire responses. One unique advantage of the tool is its reliance solely on the brief questionnaire that does not directly assess cognitive impairment. Of all the questionnaires evaluated, the best performer was the optimism scale, which comprises just six questions. Crucially, the content of these questionnaires does not need to be directly related to cognitive impairment. This flexibility allows health professionals, such as CHSWs, to select questionnaires that resonate more with their practical emphasis, such as focusing on overall well-being aspects such as optimism, hopelessness, life purpose, and life satisfaction [53-56]. Powered by our machine learning models, these questionnaires now enable a “passive or backend” cognitive impairment assessment with acceptable accuracy (AUC around 0.7) for prescreening and screening purposes. This approach alleviates the need for health professionals such as CHSWs to undergo specialized cognitive impairment assessment training and addresses potential stigmatization concerns. Residents can engage without fearing labels such as “potentially cognitively impaired” or being seen as “likely to develop dementia” from receiving cognitive impairment evaluations [12]. In essence, our tool streamlines

the cognitive impairment assessment process, making it more adaptable and less burdensome for both professionals and the community.

Compared with existing machine learning models for cognitive impairment prediction, our machine learning model is developed from a large epidemiological sample without using class-balancing techniques. This ensures that the proportion of cognitive impairment in both training and testing sets mirrors the actual prevalence of cognitive impairment in communities. As mentioned above, our model demands fewer efforts in gathering data on predictors compared to its counterparts. The tool also allows for flexible threshold selection based on the ratio of costs related to underdiagnosis to the costs related to overdiagnosis. Such flexibility facilitates end users to navigate the trade-off between underdiagnosis and overdiagnosis when using a risk assessment tool.

Compared to clinical studies using clinical biomarkers as predictors for dementia [26,29], our tool still has a gap in predictive performance. However, our goal does not intend to replace existing clinical cognitive impairment assessment tests or tools. Instead, we hope that the tool we have developed will lower the threshold for the usage of cognitive impairment risk assessment, thereby adapting it to the needs of community settings. The tool may establish an effective bridge between the community end and the clinical end, allowing the clinical end

to align resources more productively, resulting in fewer unnecessary inputs and greater efficiency in the health care system. In addition, by comparing to recent studies that use similar HRS data to predict cognitive impairment [43,87,88], our tool demonstrated predictive performance close to those studies, even though other studies used clinical or cognitive impairment-related risk factors as additional predictors. For instance, a study achieved an AUC of 0.78 but used 13 additional cognitive impairment risk factors (eg, race, stroke history, and glycated hemoglobin) alongside questionnaire data [87].

Finally, our approach expands the boundary of current research by innovatively integrating knowledge and techniques from data sciences and psychometrics to enhance the cognitive health of older adults. Central to our methodology is the machine learning tool rooted in LQR indices that are designed to detect latent cognitive deficits through modelling subtle inconsistencies in questionnaire responses. Unlike clinical markers of cognitive impairment [29-31], these psychometric indices are not only cheaper but also simpler to obtain, making them ideal for community settings. More importantly, the psychometric method allows for the evaluation of cognitive impairment independently of the questionnaire's content. Further, while conventional research methodologies typically handle low-quality data by either eliminating or statistically adjusting them by methods such as weighting [89-92], our approach presents a new perspective: leveraging, rather than dismissing, low-quality data. This strategy underscores the potential of our methodology to enhance data utility for future research.

In addition to the development and validation of this tool, we further explored the possibility of improving predictive performance to provide insights for future research. We experimented with two approaches: one was to incorporate education level as an additional predictor, and the other was to combine four selected scales. The results showed that both methods could enhance the prediction performance of current cognitive impairment using LQR as predictors (Multimedia Appendix 9). It is worth pointing out that the number of items in each questionnaire had no direct impact on the predictive effect (Multimedia Appendix 6), but combining different

questionnaires did improve the discriminative ability of the model. The application of these two approaches may increase the burden of assessment as more predictors are added to the model. Furthermore, incorporating other indicators of LQR, such as prolonged response time, may increase the predictive accuracy [93].

This study is constrained by the questionnaires available in the HRS, potentially limiting its generalizability to other questionnaire response data. However, an ongoing comprehensive meta-analysis of longitudinal aging surveys from ten countries indicates that the relationships between LQR indices and cognitive functioning remain consistent across various questionnaires and nations [94]. Additionally, ascertainment of cognitive impairment in this study is based on empirically established thresholds rather than detailed clinical evaluations. Clinical diagnosis of cognitive impairment is intricate, often requiring a mix of cognitive assessments, health assessments, laboratory tests, and brain imaging. Given this complexity, large aging survey studies such as the HRS find it impractical to clinically evaluate every participant. Consequently, these studies typically perform clinical evaluations for a select group and then use empirically derived standards, such as the Langa et al [61], to determine cognitive impairment in more extensive samples. Lastly, this study is limited by its nature as a secondary data analysis. Future research should implement the tool in real-life settings and conduct external validation and impact assessment to ensure its effectiveness.

Conclusions

The machine learning tool developed in this study provides a novel yet practical solution for tackling the challenges of early identifying cognitive impairment in community environments. The approach adopted in this study innovatively integrates psychometric methods with data science techniques and large questionnaire response data, resulting in a risk assessment tool that can facilitate health professionals working in community environments to conduct "passive or backend" cognitive impairment assessment and therefore better collaborate with medical systems to promote early identification and treatment of mild cognitive impairment and dementia.

Acknowledgments

This study is funded by the National Institute on Aging under the project titled Testing Early Markers of Cognitive Decline and Dementia Derived From Survey Response Behaviors (R01AG068190).

Data Availability

The data and code that support the findings of this study are available from the corresponding authors on request.

Conflicts of Interest

AS is a senior scientist at Gallup Organization, a consultant at Lore Contagious, and a consultant at Aztra-Zenca, Inc.

Multimedia Appendix 1

TRIPOD checklist: prediction model development and validation. TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis.

[DOCX File, 25 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Psychosocial and lifestyle questionnaire scales in the health and retirement study.

[\[DOCX File , 14 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Description of missing values.

[\[DOCX File , 15 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Training hyperparameter settings.

[\[DOCX File , 16 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Sample characteristics for the training and testing datasets.

[\[DOCX File , 16 KB-Multimedia Appendix 5\]](#)

Multimedia Appendix 6

Comparison of AUC values between the seven candidate models in predicting current cognitive impairments. AUC: area under curve.

[\[DOCX File , 19 KB-Multimedia Appendix 6\]](#)

Multimedia Appendix 7

Predictive accuracy and performance of the assessment tool under different cost ratios of underdiagnosis to overdiagnosis.

[\[DOCX File , 16 KB-Multimedia Appendix 7\]](#)

Multimedia Appendix 8

Training and validation performance of the # MLP models in 10-fold cross-validation. MLP: multilayer perceptron.

[\[DOCX File , 1015 KB-Multimedia Appendix 8\]](#)

Multimedia Appendix 9

Comparison of machine learning performance metrics after scale combination and inclusion of education level variable.

[\[DOCX File , 14 KB-Multimedia Appendix 9\]](#)

References

1. Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, et al. Mild cognitive impairment: ten years later. *Arch Neurol*. 2009;66(12):1447-1455. [[FREE Full text](#)] [doi: [10.1001/archneurol.2009.266](https://doi.org/10.1001/archneurol.2009.266)] [Medline: [20008648](https://pubmed.ncbi.nlm.nih.gov/20008648/)]
2. Hugo J, Ganguli M. Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. *Clin Geriatr Med*. 2014;30(3):421-442. [[FREE Full text](#)] [doi: [10.1016/j.cger.2014.04.001](https://doi.org/10.1016/j.cger.2014.04.001)] [Medline: [25037289](https://pubmed.ncbi.nlm.nih.gov/25037289/)]
3. van Dyck CH, Swanson CJ, Aisen P, Bateman RJ, Chen C, Gee M, et al. Lecanemab in early Alzheimer's disease. *N Engl J Med*. 2023;388(1):9-21. [doi: [10.1056/NEJMoa2212948](https://doi.org/10.1056/NEJMoa2212948)] [Medline: [36449413](https://pubmed.ncbi.nlm.nih.gov/36449413/)]
4. Sims JR, Zimmer JA, Evans CD, Lu M, Ardayfio P, Sparks J, et al. TRAILBLAZER-ALZ 2 Investigators. Donanemab in early symptomatic alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. *JAMA*. 2023;330(6):512-527. [[FREE Full text](#)] [doi: [10.1001/jama.2023.13239](https://doi.org/10.1001/jama.2023.13239)] [Medline: [37459141](https://pubmed.ncbi.nlm.nih.gov/37459141/)]
5. Cognitive assessment & care plan services. URL: <https://www.cms.gov/cognitive> [accessed 2023-11-12]
6. Association A. 2019 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*. 2019;15(3):321-387. [doi: [10.1016/j.jalz.2019.01.010](https://doi.org/10.1016/j.jalz.2019.01.010)]
7. Chung S, Romanelli RJ, Stults CD, Luft HS. Preventive visit among older adults with Medicare's introduction of annual wellness visit: closing gaps in underutilization. *Prev Med*. 2018;115:110-118. [[FREE Full text](#)] [doi: [10.1016/j.ypmed.2018.08.018](https://doi.org/10.1016/j.ypmed.2018.08.018)] [Medline: [30145346](https://pubmed.ncbi.nlm.nih.gov/30145346/)]
8. UK national screening committee. Dementia. 2019. URL: <https://view-health-screening-recommendations.service.gov.uk/dementia/> [accessed 2023-11-12]
9. World Health Organization. Global action plan on the public health response to dementia 2017-2025. 2017. URL: <https://tinyurl.com/33793pj8> [accessed 2024-10-12]

10. Connolly A, Gaehl E, Martin H, Morris J, Purandare N. Underdiagnosis of dementia in primary care: variations in the observed prevalence and comparisons to the expected prevalence. *Aging Ment Health*. 2011;15(8):978-984. [doi: [10.1080/13607863.2011.596805](https://doi.org/10.1080/13607863.2011.596805)] [Medline: [21777080](https://pubmed.ncbi.nlm.nih.gov/21777080/)]
11. Alzheimer's society. Having a cognitive assessment. 2023. URL: <https://tinyurl.com/3xkx7saa> [accessed 2023-09-12]
12. Herrmann LK, Welter E, Leverenz J, Lerner AJ, Udelson N, Kanetsky C, et al. A systematic review of dementia-related stigma research: can we move the stigma dial? *Am J Geriatr Psychiatry*. 2018;26(3):316-331. [doi: [10.1016/j.jagp.2017.09.006](https://doi.org/10.1016/j.jagp.2017.09.006)] [Medline: [29426607](https://pubmed.ncbi.nlm.nih.gov/29426607/)]
13. Meade AW, Craig SB. Identifying careless responses in survey data. *Psychol Methods*. 2012;17(3):437-455. [doi: [10.1037/a0028085](https://doi.org/10.1037/a0028085)] [Medline: [22506584](https://pubmed.ncbi.nlm.nih.gov/22506584/)]
14. Krosnick JA. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Appl Cogn Psychol*. 1991;5(3):213-236. [doi: [10.1002/acp.2350050305](https://doi.org/10.1002/acp.2350050305)]
15. Johnson JA. Ascertaining the validity of individual protocols from Web-based personality inventories. *J Res Pers*. 2005;39(1):103-129. [doi: [10.1016/j.jrp.2004.09.009](https://doi.org/10.1016/j.jrp.2004.09.009)]
16. Conijn JM, Emons WHM, De Jong K, Sijtsma K. Detecting and explaining aberrant responding to the outcome questionnaire-45. *Assessment*. 2015;22(4):513-524. [doi: [10.1177/1073191114560882](https://doi.org/10.1177/1073191114560882)] [Medline: [25520211](https://pubmed.ncbi.nlm.nih.gov/25520211/)]
17. Schneider S, Junghaenel D, Meijer E, Zelinski EM, Jin H, Lee PJ, et al. Quality of survey responses at older ages predicts cognitive decline and mortality risk. *Innov Aging*. 2022;6(3):igac027. [FREE Full text] [doi: [10.1093/geroni/igac027](https://doi.org/10.1093/geroni/igac027)] [Medline: [35663275](https://pubmed.ncbi.nlm.nih.gov/35663275/)]
18. Buchanan EM, Scofield JE. Methods to detect low quality data and its implication for psychological research. *Behav Res Methods*. 2018;50(6):2586-2596. [doi: [10.3758/s13428-018-1035-6](https://doi.org/10.3758/s13428-018-1035-6)] [Medline: [29542063](https://pubmed.ncbi.nlm.nih.gov/29542063/)]
19. Kutschar P, Weichbold M, Osterbrink J. Effects of age and cognitive function on data quality of standardized surveys in nursing home populations. *BMC Geriatr*. 2019;19(1):244. [FREE Full text] [doi: [10.1186/s12877-019-1258-0](https://doi.org/10.1186/s12877-019-1258-0)] [Medline: [31481012](https://pubmed.ncbi.nlm.nih.gov/31481012/)]
20. Colsher PL, Wallace RB. Data quality and age: health and psychobehavioral correlates of item nonresponse and inconsistent responses. *J Gerontol*. 1989;44(2):P45-P52. [doi: [10.1093/geronj/44.2.p45](https://doi.org/10.1093/geronj/44.2.p45)] [Medline: [2921475](https://pubmed.ncbi.nlm.nih.gov/2921475/)]
21. Lechner CM, Rammstedt B. Cognitive ability, acquiescence, and the structure of personality in a sample of older adults. *Psychol Assess*. 2015;27(4):1301-1311. [doi: [10.1037/pas0000151](https://doi.org/10.1037/pas0000151)] [Medline: [26011482](https://pubmed.ncbi.nlm.nih.gov/26011482/)]
22. Tourangeau R, Rips LJ, Rasinski K. *The Psychology of Survey Response*. Cambridge, England. Cambridge University Press; 2000.
23. Schneider S, Junghaenel DU, Zelinski EM, Meijer E, Stone AA, Langa KM, et al. Subtle mistakes in self-report surveys predict future transition to dementia. *Alzheimers Dement (Amst)*. 2021;13(1):e12252. [FREE Full text] [doi: [10.1002/dad2.12252](https://doi.org/10.1002/dad2.12252)] [Medline: [34934800](https://pubmed.ncbi.nlm.nih.gov/34934800/)]
24. Fuchs M. Item-Nonresponse in einer Befragung von Alten und Hochbetagten. In: Weichbold M, Bacher J, Wolf C, editors. *Umfrageforschung*. Wiesbaden. VS Verlag für Sozialwissenschaften; 2009.
25. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeblerlein SB, et al. Contributors. NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14(4):535-562. [FREE Full text] [doi: [10.1016/j.jalz.2018.02.018](https://doi.org/10.1016/j.jalz.2018.02.018)] [Medline: [29653606](https://pubmed.ncbi.nlm.nih.gov/29653606/)]
26. Tanveer M, Richhariya B, Khan RU, Rashid AH, Khanna P, Prasad M, et al. Machine learning techniques for the diagnosis of alzheimer's disease. 2020. Presented at: ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM); 2020 April 17:1-35; New York, NY, United States. [doi: [10.1145/3344998](https://doi.org/10.1145/3344998)]
27. Kim WJ, Sung JM, Sung D, Chae MH, An SK, Namkoong K, et al. Cox proportional hazard regression versus a deep learning algorithm in the prediction of dementia: an analysis based on periodic health examination. *JMIR Med Inform*. 2019;7(3):e13139. [FREE Full text] [doi: [10.2196/13139](https://doi.org/10.2196/13139)] [Medline: [31471957](https://pubmed.ncbi.nlm.nih.gov/31471957/)]
28. Soroski T, da Cunha Vasco T, Newton-Mason S, Granby S, Lewis C, Harisinghani A, et al. Evaluating web-based automatic transcription for alzheimer speech data: transcript comparison and machine learning analysis. *JMIR Aging*. 2022;5(3):e33460. [FREE Full text] [doi: [10.2196/33460](https://doi.org/10.2196/33460)] [Medline: [36129754](https://pubmed.ncbi.nlm.nih.gov/36129754/)]
29. Aljović A, Badnjević A, Gurbeta L. Artificial neural networks in the discrimination of alzheimer's disease using biomarkers data. 2016. Presented at: 2016 5th Mediterranean Conference on Embedded Computing (MECO); 2016 June 12-16:286-289; Bar, Montenegro.
30. Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, et al. Alzheimer's Disease Neuroimaging Initiative. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *Neuroimage Clin*. 2019;21:101645. [FREE Full text] [doi: [10.1016/j.nicl.2018.101645](https://doi.org/10.1016/j.nicl.2018.101645)] [Medline: [30584016](https://pubmed.ncbi.nlm.nih.gov/30584016/)]
31. Bellomo G, Indaco A, Chiasserini D, Maderna E, Paolini Paoletti F, Gaetani L, et al. Machine learning driven profiling of cerebrospinal fluid core biomarkers in alzheimer's disease and other neurological disorders. *Front Neurosci*. 2021;15:647783. [FREE Full text] [doi: [10.3389/fnins.2021.647783](https://doi.org/10.3389/fnins.2021.647783)] [Medline: [33867925](https://pubmed.ncbi.nlm.nih.gov/33867925/)]
32. Lee G, Nho K, Kang B, Sohn K, Kim D, for Alzheimer's Disease Neuroimaging Initiative. Predicting Alzheimer's disease progression using multi-modal deep learning approach. *Sci Rep*. 2019;9(1):1952. [FREE Full text] [doi: [10.1038/s41598-018-37769-z](https://doi.org/10.1038/s41598-018-37769-z)] [Medline: [30760848](https://pubmed.ncbi.nlm.nih.gov/30760848/)]

33. El-Sappagh S, Saleh H, Ali F, Amer E, Abuhmed T. Two-stage deep learning model for Alzheimer's disease detection and prediction of the mild cognitive impairment time. *Neural Comput Appl*. 2022;34(17):14487-14509. [doi: [10.1007/s00521-022-07263-9](https://doi.org/10.1007/s00521-022-07263-9)]
34. Grassi M, Rouleaux N, Caldirola D, Loewenstein D, Schruers K, Perna G, et al. Alzheimer's Disease Neuroimaging Initiative. A novel ensemble-based machine learning algorithm to predict the conversion from mild cognitive impairment to alzheimer's disease using socio-demographic characteristics, clinical information, and neuropsychological measures. *Front Neurol*. 2019;10:756. [FREE Full text] [doi: [10.3389/fneur.2019.00756](https://doi.org/10.3389/fneur.2019.00756)] [Medline: [31379711](https://pubmed.ncbi.nlm.nih.gov/31379711/)]
35. Bai W, Chen P, Cai H, Zhang Q, Su Z, Cheung T, et al. Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: a meta-analysis and systematic review of epidemiology studies. *Age Ageing*. 2022;51(8):afac173. [doi: [10.1093/ageing/afac173](https://doi.org/10.1093/ageing/afac173)] [Medline: [35977150](https://pubmed.ncbi.nlm.nih.gov/35977150/)]
36. Wei Q, Dunbrack RL. The role of balanced training and testing data sets for binary classifiers in bioinformatics. *PLoS One*. 2013;8(7):e67863. [FREE Full text] [doi: [10.1371/journal.pone.0067863](https://doi.org/10.1371/journal.pone.0067863)] [Medline: [23874456](https://pubmed.ncbi.nlm.nih.gov/23874456/)]
37. Junsomboon N, Phienthrakul T. Combining over-sampling/under-sampling techniques for imbalance dataset. New York, NY, USA. Association for Computing Machinery; 2017. Presented at: Proceedings of the 9th International Conference on Machine Learning and Computing; 2017 February 24:243-247; Singapore.
38. van den Goorbergh R, van Smeden M, Timmerman D, Van Calster B. The harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression. *J Am Med Inform Assoc*. 2022;29(9):1525-1534. [FREE Full text] [doi: [10.1093/jamia/ocac093](https://doi.org/10.1093/jamia/ocac093)] [Medline: [35686364](https://pubmed.ncbi.nlm.nih.gov/35686364/)]
39. Krawczyk B. Learning from imbalanced data: open challenges and future directions. *Prog Artif Intell*. 2016;5(4):221-232. [doi: [10.1007/s13748-016-0094-0](https://doi.org/10.1007/s13748-016-0094-0)]
40. Pu L, Pan D, Wang H, He X, Zhang X, Yu Z, et al. A predictive model for the risk of cognitive impairment in community middle-aged and older adults. *Asian J Psychiatr*. 2023;79:103380. [doi: [10.1016/j.ajp.2022.103380](https://doi.org/10.1016/j.ajp.2022.103380)] [Medline: [36495830](https://pubmed.ncbi.nlm.nih.gov/36495830/)]
41. Cong L, Ren Y, Wang Y, Hou T, Dong Y, Han X, et al. Mild cognitive impairment among rural-dwelling older adults in China: a community-based study. *Alzheimers Dement*. 2023;19(1):56-66. [FREE Full text] [doi: [10.1002/alz.12629](https://doi.org/10.1002/alz.12629)] [Medline: [35262288](https://pubmed.ncbi.nlm.nih.gov/35262288/)]
42. Liu H, Zhang X, Liu H, Chong ST. Using machine learning to predict cognitive impairment among middle-aged and older chinese: a longitudinal study. *Int J Public Health*. 2023;68:1605322. [FREE Full text] [doi: [10.3389/ijph.2023.1605322](https://doi.org/10.3389/ijph.2023.1605322)] [Medline: [36798738](https://pubmed.ncbi.nlm.nih.gov/36798738/)]
43. Tan WY, Hargreaves C, Chen C, Hilal S. A machine learning approach for early diagnosis of cognitive impairment using population-based data. *J Alzheimers Dis*. 2023;91(1):449-461. [FREE Full text] [doi: [10.3233/JAD-220776](https://doi.org/10.3233/JAD-220776)] [Medline: [36442196](https://pubmed.ncbi.nlm.nih.gov/36442196/)]
44. Jin H, Junghaenel DU, Orriens B, Lee PJ, Schneider S. Developing early markers of cognitive decline and dementia derived from survey response behaviors: protocol for analyses of preexisting large-scale longitudinal data. *JMIR Res Protoc*. 2023;12:e44627. [doi: [10.2196/44627](https://doi.org/10.2196/44627)] [Medline: [36809337](https://pubmed.ncbi.nlm.nih.gov/36809337/)]
45. Wagner M, Kuppler M, Rietz C, Kaspar R. Non-response in surveys of very old people. *Eur J Ageing*. 2019;16(2):249-258. [FREE Full text] [doi: [10.1007/s10433-018-0488-x](https://doi.org/10.1007/s10433-018-0488-x)] [Medline: [31139038](https://pubmed.ncbi.nlm.nih.gov/31139038/)]
46. Vigil-Colet A, Morales-Vives F, Lorenzo-Seva U. How social desirability and acquiescence affect the age-personality relationship. *Psicothema*. 2013;25(3):342-348. [doi: [10.7334/psicothema2012.297](https://doi.org/10.7334/psicothema2012.297)] [Medline: [23910749](https://pubmed.ncbi.nlm.nih.gov/23910749/)]
47. James G, Witten D, Hastie T. An introduction to statistical learning. Berlin. Springer; 2013.
48. Hartzler AL, Tuzzio L, Hsu C, Wagner EH. Roles and functions of community health workers in primary care. *Ann Fam Med*. 2018;16(3):240-245. [FREE Full text] [doi: [10.1370/afm.2208](https://doi.org/10.1370/afm.2208)] [Medline: [29760028](https://pubmed.ncbi.nlm.nih.gov/29760028/)]
49. Collins GS, Reitsma JB, Altman DG, Moons K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med*. 2015;13:1. [FREE Full text] [doi: [10.1186/s12916-014-0241-z](https://doi.org/10.1186/s12916-014-0241-z)] [Medline: [25563062](https://pubmed.ncbi.nlm.nih.gov/25563062/)]
50. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort profile: the health and retirement study (HRS). *Int J Epidemiol*. 2014;43(2):576-585. [FREE Full text] [doi: [10.1093/ije/dyu067](https://doi.org/10.1093/ije/dyu067)] [Medline: [24671021](https://pubmed.ncbi.nlm.nih.gov/24671021/)]
51. Smith J, Fisher G, Ryan L. Health and Retirement Study Psychosocial and Lifestyle Questionnaire 2006-2010: Documentation report. United States. University of Michigan; 2013. URL: <https://hrsonline.isr.umich.edu/sitedocs/USERG/HRS2006-2010SAQdoc.pdf>
52. Schneider S, Hernandez R, Junghaenel DU, Jin H, Lee P, Gao H, et al. Can you tell people's cognitive ability level from their response patterns in questionnaires? *Behav Res Methods*. 2024;56(7):6741-6758. [doi: [10.3758/s13428-024-02388-2](https://doi.org/10.3758/s13428-024-02388-2)] [Medline: [38528247](https://pubmed.ncbi.nlm.nih.gov/38528247/)]
53. Boylan JM, Tompkins JL, Krueger PM. Psychological well-being, education, and mortality. *Health Psychol*. 2022;41(3):225-234. [FREE Full text] [doi: [10.1037/hea0001159](https://doi.org/10.1037/hea0001159)] [Medline: [35157480](https://pubmed.ncbi.nlm.nih.gov/35157480/)]
54. Kim ES, Park N, Peterson C. Dispositional optimism protects older adults from stroke: the health and retirement study. *Stroke*. 2011;42(10):2855-2859. [doi: [10.1161/STROKEAHA.111.613448](https://doi.org/10.1161/STROKEAHA.111.613448)] [Medline: [21778446](https://pubmed.ncbi.nlm.nih.gov/21778446/)]
55. Matthews KA, Räikkönen K, Sutton-Tyrrell K, Kuller LH. Optimistic attitudes protect against progression of carotid atherosclerosis in healthy middle-aged women. *Psychosom Med*. 2004;66(5):640-644. [doi: [10.1097/01.psy.0000139999.99756.a5](https://doi.org/10.1097/01.psy.0000139999.99756.a5)] [Medline: [15385685](https://pubmed.ncbi.nlm.nih.gov/15385685/)]

56. Uchino BN, Cribbet M, de Grey RGK, Cronan S, Trettevik R, Smith TW. Dispositional optimism and sleep quality: a test of mediating pathways. *J Behav Med.* 2017;40(2):360-365. [FREE Full text] [doi: [10.1007/s10865-016-9792-0](https://doi.org/10.1007/s10865-016-9792-0)] [Medline: [27592128](https://pubmed.ncbi.nlm.nih.gov/27592128/)]
57. Samejima F. Graded response models. In: *Handbook of item response theory, volume one.* United Kingdom. Chapman and Hall/CRC; 2016:123-136.
58. Yu X, Cheng Y. A change-point analysis procedure based on weighted residuals to detect back random responding. *Psychol Methods.* 2019;24(5):658-674. [doi: [10.1037/met0000212](https://doi.org/10.1037/met0000212)] [Medline: [30762378](https://pubmed.ncbi.nlm.nih.gov/30762378/)]
59. Chalmers RP. Mirt: a multidimensional item response theory package for the R environment. *J Stat Soft.* 2012;48(6):1-29. [doi: [10.18637/jss.v048.i06](https://doi.org/10.18637/jss.v048.i06)]
60. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. *J Open Source Softw.* 2019;4(43):1686. [doi: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686)]
61. Langa K, Weir D, Kabeto M. Langa-weir classification of cognitive function (1995 Onward). Survey Research Center Institute for Social Research, University of Michigan. URL: <https://tinyurl.com/43m397an> [accessed 2024-10-12]
62. Ramchoun H, Ghanou Y, Ettaouil M, Idrissi MAJ. Multilayer perceptron: architecture optimization and training. *Int J Interact Multimed.* 2016;4(1):26-30. [doi: [10.9781/ijimai.2016.4.15](https://doi.org/10.9781/ijimai.2016.4.15)] [Medline: [34887180](https://pubmed.ncbi.nlm.nih.gov/34887180/)]
63. Popescu MC, Balas VE, Perescu-Popescu L, Mastorakis N. Multilayer perceptron and neural networks. *WSEAS Trans Circuits Syst.* 2009;8:579-588. [FREE Full text]
64. Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput Sci.* 2021;2(6):420. [FREE Full text] [doi: [10.1007/s42979-021-00815-1](https://doi.org/10.1007/s42979-021-00815-1)] [Medline: [34426802](https://pubmed.ncbi.nlm.nih.gov/34426802/)]
65. 6.4. Imputation of missing values. scikit-learn. URL: <https://scikit-learn.org/stable/modules/impute.html> [accessed 2023-09-12]
66. TensorFlow v2.13.0 - python. TensorFlow. URL: https://www.tensorflow.org/api_docs/python/tf [accessed 2023-09-12]
67. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. 2015. Presented at: Proceedings of the 32nd International Conference on International Conference on Machine Learning; 2015 July 06:448-456; Lille, France.
68. Sharma S, Sharma S, Athaiya A. Activation functions in neural networks. *Int J Eng Appl Sci Technol.* 2020;4(12):310-316. [doi: [10.33564/IJEAST.2020.v04i12.054](https://doi.org/10.33564/IJEAST.2020.v04i12.054)]
69. Baldi P, Sadowski PJ. Understanding dropout. In: Burges CJ, Bottou L, Welling M, editors. *Advances in Neural Information Processing Systems.* New York. Curran Associates, Inc; 2013.
70. Glorot GX, Bordes A, Bengio Y. Deep sparse rectifier neural networks. Fort Lauderdale, FL, USA. PMLR; 2011. Presented at: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics; 2011 April 11-13:315-323; FL, USA.
71. Rodríguez JD, Pérez A, Lozano J. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans Pattern Anal Mach Intell.* 2010;32(3):569-575. [doi: [10.1109/TPAMI.2009.187](https://doi.org/10.1109/TPAMI.2009.187)] [Medline: [20075479](https://pubmed.ncbi.nlm.nih.gov/20075479/)]
72. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression.* New Jersey. John Wiley & Sons Inc; 2013.
73. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv:1412.6980. Preprint posted online on January 29, 2017
74. Malley O, Bursztein E, Long J. KerasTuner. 2019. URL: <https://github.com/keras-team/keras-tuner> [accessed 2024-10-12]
75. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. arXiv:1603.04467. Preprint posted online on March 16, 2016
76. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 2011;12:2825-2830. [FREE Full text]
77. Berkson J. Application of the logistic function to bio-assay. *J Am Stat Assoc.* 1944;39(227):357-365. [doi: [10.1080/01621459.1944.10500699](https://doi.org/10.1080/01621459.1944.10500699)]
78. Song YY, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry.* 2015;27(2):130-135. [FREE Full text] [doi: [10.11919/j.issn.1002-0829.215044](https://doi.org/10.11919/j.issn.1002-0829.215044)] [Medline: [26120265](https://pubmed.ncbi.nlm.nih.gov/26120265/)]
79. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. New York, NY, USA. Association for Computing Machinery; 2016. Presented at: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016:785-794; California, San Francisco, USA.
80. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: a highly efficient gradient boosting decision tree. Red Hook, NY, United States. Curran Associates Inc; 2017. Presented at: Proceedings of the 31st International Conference on Neural Information Processing Systems; 2017 December 04:3149-3157; Red Hook, NY, United States.
81. Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. Doha, Qatar. Association for Computational Linguistics; 2014. Presented at: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP); 2014 October 01:1724-1734; Doha, Qatar.
82. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* 1989;1(4):541-551. [doi: [10.1162/neco.1989.1.4.541](https://doi.org/10.1162/neco.1989.1.4.541)]
83. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* 1997;9(8):1735-1780. [doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735)] [Medline: [9377276](https://pubmed.ncbi.nlm.nih.gov/9377276/)]

84. Freeman EA, Moisen G. PresenceAbsence: an R package for presence absence analysis. *J Stat Soft.* 2008;23(11):1-31. [doi: [10.18637/jss.v023.i11](https://doi.org/10.18637/jss.v023.i11)]
85. Jin H, Wu S, Di Capua P. Development of a clinical forecasting model to predict comorbid depression among diabetes patients and an application in depression screening policy making. *Prev Chronic Dis.* 2015;12:E142. [FREE Full text] [doi: [10.5888/pcd12.150047](https://doi.org/10.5888/pcd12.150047)] [Medline: [26334714](https://pubmed.ncbi.nlm.nih.gov/26334714/)]
86. Community-based cognitive impairment risk assessment tool. 2023. URL: <https://papernew-08e37b67fd73.herokuapp.com> [accessed 2024-10-12]
87. Li W, Zeng L, Yuan S, Shang Y, Zhuang W, Chen Z, et al. Machine learning for the prediction of cognitive impairment in older adults. *Front Neurosci.* 2023;17:1158141. [FREE Full text] [doi: [10.3389/fnins.2023.1158141](https://doi.org/10.3389/fnins.2023.1158141)] [Medline: [37179565](https://pubmed.ncbi.nlm.nih.gov/37179565/)]
88. Hu M, Shu X, Yu G, Wu X, Välimäki M, Feng H. A risk prediction model based on machine learning for cognitive impairment among chinese community-dwelling elderly people with normal cognition: development and validation study. *J Med Internet Res.* 2021;23(2):e20298. [FREE Full text] [doi: [10.2196/20298](https://doi.org/10.2196/20298)] [Medline: [33625369](https://pubmed.ncbi.nlm.nih.gov/33625369/)]
89. Greszki R, Meyer M, Schoen H. Exploring the effects of removing “Too Fast” responses and respondents from web surveys. *Public Opin Q.* 2015;79(2):471-503. [doi: [10.1093/poq/nfu058](https://doi.org/10.1093/poq/nfu058)]
90. Saris WE, Revilla M. Correction for measurement errors in survey research: necessary and possible. *Soc Indic Res.* 2015;127(3):1005-1020. [doi: [10.1007/s11205-015-1002-x](https://doi.org/10.1007/s11205-015-1002-x)]
91. Schneider S. Extracting response style bias from measures of positive and negative affect in aging research. *J Gerontol B Psychol Sci Soc Sci.* 2017;73(1):64-74. [FREE Full text] [doi: [10.1093/geronb/gbw103](https://doi.org/10.1093/geronb/gbw103)] [Medline: [27543081](https://pubmed.ncbi.nlm.nih.gov/27543081/)]
92. DeCastellarnau A. A classification of response scale characteristics that affect data quality: a literature review. *Qual Quant.* 2018;52(4):1523-1559. [FREE Full text] [doi: [10.1007/s11135-017-0533-4](https://doi.org/10.1007/s11135-017-0533-4)] [Medline: [29937582](https://pubmed.ncbi.nlm.nih.gov/29937582/)]
93. Schneider S, Junghaenel D, Meijer E, Stone AA, Orriens B, Jin H, et al. Using item response times in online questionnaires to detect mild cognitive impairment. *J Gerontol: Ser B.* 2023;78(8):1278-1283. [FREE Full text] [doi: [10.1093/geronb/gbad043](https://doi.org/10.1093/geronb/gbad043)] [Medline: [36879431](https://pubmed.ncbi.nlm.nih.gov/36879431/)]
94. Schneider S, Lee P, Hernandez R, Junghaenel DU, Stone AA, Meijer E, et al. Cognitive functioning and the quality of survey responses: an individual participant data meta-analysis of 10 epidemiological studies of aging. *J Gerontol B Psychol Sci Soc Sci.* 2024;79(5):gbae030. [doi: [10.1093/geronb/gbae030](https://doi.org/10.1093/geronb/gbae030)] [Medline: [38460115](https://pubmed.ncbi.nlm.nih.gov/38460115/)]

Abbreviations

AUC: area under the curve

CHSW: community health and social worker

HRS: Health and Retirement Study

LQR: low-quality response

MCI: mild cognitive impairment

MLP: multilayer perceptron

TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

Edited by A Mavragani; submitted 06.11.23; peer-reviewed by M Mardini, Y Cheng, K Miyata; comments to author 12.03.24; revised version received 18.06.24; accepted 23.09.24; published 13.11.24

Please cite as:

Gao H, Schneider S, Hernandez R, Harris J, Maupin D, Junghaenel DU, Kapteyn A, Stone A, Zelinski E, Meijer E, Lee P-J, Orriens B, Jin H

Early Identification of Cognitive Impairment in Community Environments Through Modeling Subtle Inconsistencies in Questionnaire Responses: Machine Learning Model Development and Validation

JMIR Form Res 2024;8:e54335

URL: <https://formative.jmir.org/2024/1/e54335>

doi: [10.2196/54335](https://doi.org/10.2196/54335)

PMID:

©Hongxin Gao, Stefan Schneider, Raymond Hernandez, Jenny Harris, Danny Maupin, Doerte U Junghaenel, Arie Kapteyn, Arthur Stone, Elizabeth Zelinski, Erik Meijer, Pey-Juan Lee, Bart Orriens, Haomiao Jin. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 13.11.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.