Updated Criteria for the Diagnostic Procedure for Parkinson's Disease Dementia on Level I and their Validity in Deep Brain Stimulation Cohort

Martina Mana, Josef Mana, Tereza Uhrova, Robert Jech, and Ondrej Bezdicek

Department of Neurology and Centre of Clinical Neuroscience, First Faculty of Medicine and General University Hospital in Prague, Charles University, Czech Republic

Author Note

Martina Mana https://orcid.org/0009-0007-4665-3946

Josef Mana https://orcid.org/0000-0002-7817-3978

Robert Jech https://orcid.org/0000-0002-9732-8947

Ondrej Bezdicek https://orcid.org/0000-0002-5108-0181

Author roles were classified using the Contributor Role Taxonomy (CRediT; https://credit.niso.org/) as follows: Martina Mana: Conceptualization, Data curation, Writing - original draft; Josef Mana: Conceptualization, Data curation, Investigation, Formal analysis, Software, Methodology, Project administration, Validation, Writing - original draft; Tereza Uhrova: Investigation; Robert Jech: Funding acquisition, Resources, Writing - review & editing; Ondrej Bezdicek: Investigation, Data curation, Funding acquisition, Conceptualization, Project administration, Supervision, Writing - original draft

Correspondence concerning this article should be addressed to Ondrej Bezdicek, Email: ondrej.bezdicek@gmail.com

Updated Criteria for the Diagnostic Procedure for Parkinson's Disease Dementia on Level I and their Validity in Deep Brain Stimulation Cohort

Introduction

Parkinson's disease (PD) is a neurodegenerative disorder characterised by a progressive onset of motor symptoms, including rigidity, bradykinesia, postural instability and resting tremor. Beyond motor symptoms, patients routinely suffer from non-motor impairments (Postuma et al., 2015), such as cognitive decline. This decline might peak in Parkinson's disease dementia (PDD) in a subset of patients (Meireles & Massano, 2012).

Despite its clinical relevance, the diagnosis of PDD remains complex. The original diagnostic criteria (Dubois et al., 2007), first formalized in 2006, were heavily influenced by frameworks established primarily for Alzheimer's disease (AD). Although these criteria provided a valuable initial foundation, they lacked the specificity required to capture the distinct pathophysiological and cognitive features of PD-related dementia (Emre et al., 2007; Yamashita et al., 2023).

One notable feature of the original criteria was the provision of an algorithm that allowed for flexibility in test selection (Dubois et al., 2007). Specifically, clinicians could choose between months reversed or seven backwards for attention assessment, lexical fluency or clock drawing for executive function evaluation. Moreover, the critaria included MMSE pentagons for evaluation of visuospatial ability, and three-word recall for memory assessment. Agreement across different criteria allows for the parallel computation of inter-rater reliability , which, in turn, facilitates the calculation of construct validity (Conway et al., 1995) further strengthening the diagnostic framework for PDD.

Subsequent steps have introduced the concept of Parkinson's disease mild cognitive impairment (PD-MCI), refining the understanding of cognitive dysfunction in PD (Litvan et al., 2012). The PD-MCI criteria have propelled progress in diverse areas, enabling improvements in clinical characterization, identification of genetic correlates, therapeutic interventions, clinical trial design, and the assessment of progression risk to PDD (Aarsland et al., 2021; Hoogland et al., 2017, 2019).

Currently, efforts are focused on refining the PDD diagnostic framework to improve its

consistency and applicability in both research and clinical contexts (Kulisevsky et al., 2024). Our prior investigations have underscored the importance of employing rigorous psychometric methodologies to differentiate PD-MCI from PDD (Bezdicek et al., 2016; Deuschl et al., 2006).

According to a recent metanalysis, the epidemiological estimates of PDD prevalence among individuals with PD vary widely, ranging from 14% to 55%, depending on methodological criteria employed (Sousa et al., 2022). Moreover, factors such as patients' sex (Cereda et al., 2016), age and disease duration appear to modulate the risk of cognitive decline (Oh et al., 2016; Rana et al., 2011).

The present study evaluates the diagnostic concordance between the original Level I PDD criteria, as established by the Movement Disorder Society (MDS) Task Force (Dubois et al., 2007; Emre et al., 2007) and criteria inspired by the recent call for change (Kulisevsky et al., 2024) within a PD cohort selected for DBS. Furthermore, both sets of criteria are compared to PDD diagnosed on Level II. The study aims to address following research objectives (RO): 1) to estimate the prevalence of PDD among patients considered for DBS (RO1), 2) to assess variability in PDD diagnosis depending on the diagnostic criteria applied (RO2), 3) to evaluate the diagnostic concordance between different sets of PDD criteria (RO3), and 4) to identify diagnostic components that contribute to variability in PDD classification across criteria (RO4). By addressing these objectives, this study seeks to validate the revised PDD criteria and evaluate their relevance in the context of DBS eligibility, thereby contributing to the refinement of cognitive assessment protocols in PD.

Suggestion: (MK) **(RO1) To estimate the prevalence of Parkinson's disease dementia (PDD) and evaluate the diagnostic variability and concordance across different PDD criteria. (RO2) To identify specific diagnostic components that contribute to variability in PDD classification across the applied criteria.**

Methods

Participants

This study retrospectively analyzed clinical data from a cohort of patients with PD at the General University Hospital in Prague. All patients were diagnosed with idiopathic PD by

a movement disorder specialist according to the MDS Clinical Diagnostic Criteria for PD (Postuma et al., 2015). Clinical records spanning August 2014 to February 2025 were examined. All participants underwent neuropsychological evaluation conducted by a trained clinical psychologist (OB) as part of standard preoperative cognitive assessments for DBS eligibility at the General University Hospital in Prague.

Ethical approval for the study protocol was obtained from the Ethics Committee of the General University Hospital in Prague. Informed consent was secured from all patients prior to their neuropsychological assessments, in adherence to ethical research guidelines.

Neuropsychological Assessment

Cognitive performance was evaluated at both Level I (abbreviated assessment) and Level II (comprehensive assessment) according to the standard MDS neuropsychology battery for Parkinson's Disease Mild Cognitive Impairment (PD-MCI) Bezdicek et al. (2017). Level I was assessed using the Mini-Mental State Examination (MMSE) (Folstein et al., 1975; Stepankova et al., 2015) and the Montreal Cognitive Assessment (MoCA) (Kopecek et al., 2016; Nasreddine et al., 2005), both of which provide measures of global cognitive functioning. The neuropsychological assessment at Level II covered five cognitive domains, each evaluated through specific tests as follows: attention and working memory assessed using Trail Making Test Part A (TMT-A) (Bezdicek et al., 2012; Reitan, 2004), and WAIS Digit Span Backward (WAIS DSB) (Wechsler, 1997); executive function evaluated via Categorical Verbal Fluency (CF) (Benton et al., 1989), and subtest from the Prague Stroop Test – Colors (PST-C) (Bezdíček et al., 2021); language measured with the WAIS Similarities subtest (Wechsler, 1997), and the Boston Naming Test (BNT-60) (Kaplan et al., 1983); memory examined using the Rey Auditory Verbal Learning Test (RAVLT) (Frydrychová et al., 2018; Rey, 1964) for delayed recall, and the Brief Visual Memory Test–Revised (BVMTR) (Benedict, 1997; Havlík et al., 2020) for delayed recall, or WAIS Family Pictures subtest (Wechsler, 1997) including delayed recall and forced choice recognition; visuospatial function assessed through the Judgment of Line Orientation Test (JoLO) (Benton et al., 1983), and Clock Drawing Test (CLOX) (Royall et al., 1998).

In addition to the core cognitive assessments, tasks such as the Clock Drawing Test

(CDT) and Letter Fluency tasks were included to capture domain-specific impairments. The classification of Parkinson's Disease Dementia (PDD) based on Level I criteria was determined using established scoring thresholds from the original criteria (Dubois et al., 2007) with corresponding MoCA equivalents.

To assess functional impairment, the Functional Activities Questionnaire (FAQ)(Bezdicek et al., 2016; Pfeffer et al., 1982) was administered. Additionally, neuropsychiatric status was evaluated using the Beck Depression Inventory-II (BDI) (Beck et al., 1996; Ciharova et al., 2020) and State-Trait Anxiety Inventory (STAI) (Mullner et al., 1980; Spielberger et al., 1983). Psychotic symptoms were assessed through structured psychiatric interviews conducted by a trained psychiatrist.

Theoretical and Empirical Estimands

Following the framework proposed by Lundberg et al. (2021), this section connects our research objectives and their corresponding theoretical (i.e., targets of inference) and empirical (i.e., data-driven) estimands. The theoretical estimand refers to a unit-specific quantity defined over a target population and represents the ideal quantity that would address the research question under optimal conditions, such as access to complete population data or perfect experimental control. In contrast, the empirical estimand corresponds to the quantity that is actually computable using the available dataset, given real world constraints. In our study, the target population for all theoretical estimands is defined as individuals diagnosed with Parkinson's disease (PD) who are potential candidates for deep brain stimulation (DBS) treatment.

The theoretical estimand corresponding to the RO1 is the true prevalence of probable PDD in this population. Empirically, this is estimated by calculating the proportion of patients classified as having probable PDD according to each of the diagnostic criteria under consideration.

For the RO2, the theoretical estimand is the variance in prevalence attributable to the diagnostic process itself—that is, the extent to which different sets of criteria yield differing prevalence estimates for the same population. This variability is empirically quantified by comparing the distribution of PDD classifications across criteria within the cohort.

The RO3 concerns diagnostic concordance. Here, the theoretical estimand is a set of population-level contingency tables comparing PDD classifications assigned by each pair of diagnostic criteria, along with derived metrics such as sensitivity, specificity, and kappa coefficients. The corresponding empirical estimands are represented by matrices generated through pairwise Receiver Operating Characteristic (ROC) curve analyses that evaluate the discriminatory performance of each diagnostic set.

Finally, for the RO4, the theoretical estimand is defined as the set of diagnostic components whose variation systematically alters the probability of a probable PDD diagnosis. This aspect of the study is exploratory in nature. Empirically, we assess the contribution of each diagnostic feature by examining how variations in operational definitions (e.g., domain-specific thresholds, criteria for functional impairment) influence the empirical estimands derived for the first three objectives. This allows us to identify the diagnostic elements most responsible for inter-criterion discrepancies. We approach this RO from an exploratory point of view and evaluate the importance of each PDD criteria component by observing change in empirical estimands for RO1-3 when stratified by different oprationalization decisions.

Operationalization of Parkinson's Disease Dementia

In this study, we applied three distinct sets of diagnostic criteria for **probable**[^1] PDD at Level I. The first set was based on the original framework (Dubois et al., 2007), which utilized the Mini-Mental State Examination (MMSE) as a global cognitive screening tool, supplemented by assessments of attention, executive function, visuospatial abilities, and memory. The second set of criteria was drawn from the recent call for change of dementia diagnostic guidelines (Kulisevsky et al., 2024), which advocate for more sensitive cognitive domain assessments in the context of Parkinson's disease (PD). This updated approach incorporated specific items from the Montreal Cognitive Assessment (MoCA) to better detect PD-related dementia. The third approach applied the Czech version of the shortened Montreal Cognitive Assessment (sMoCA) (Bezdicek et al., 2020), a time-efficient modification designed to ascertain whether equivalent cognitive impairments could be reliably identified using a reduced testing protocol. Lastly, the fourth approach followed the Level II battery

protocol, which is commonly used in the evaluation of PD-MCI. **The Level II methodology, including the use of a regression-based normative scoring approach, has been detailed in a prior study (Bezdicek et al., 2017). Refer to Table 1**[1] to a summary of the components and scoring thresholds of each diagnostic criterion. All non-cognitive criteria of probable PDD (i.e., diagnosis of PD that developed before dementia and absence of Major Depression, delirium or other abnormalities that obscure diagnosis were established by an independent neurological and psychiatric assessments and held true for all patients in the sample).

«««< HEAD [^1]: **JM: We should unite the terminlogy. Most importantly, I reckon, we should differentiate between "probable PDD" (what is implied by meeting all the criteria within an operationalization) and "PDD" (the latent state of patient's cognition); and to use consistently the following trifecta - "criterion," "operationalization" and their "component".** ======= [2]: **JM: We should unite the terminology. Most importantly, I reckon, we should differentiate between "probable PDD" (what is implied by meeting all the criteria within an operationalization) and "PDD" (the latent state of patient's cognition); and to use consistently the following trifecta - "criterion," "operationalization" and their "component".** »»»> 0d33cc0b0961c3776e2225e6d367cfc52b9a2e48

[Insert Table 1 here]

For each of these diagnostic approaches, we applied two operationalization strategies based on deficits in Instrumental Activities of Daily Living (IADL). First, we utilized FAQ item 9, which approximates the pill questionnaire from the original criteria (Dubois et al., 2007) employing a cut-off score of 2 points or higher. Second, we applied the entire Functional Activities Questionnaire (FAQ) as suggested in the call for change (Kulisevsky et al., 2024), employing a cut-off score of 7 points or higher based on Czech normative data (Bezdíček et al., 2011). These methodologies resulted in a total of 68 operationalizations, which were distributed across different diagnostic criteria: 4 MMSE-based, 60 MoCA-based,

[1] **JM: Placeholder until we make the table proper.**

[2] **JM: Need to add Level II operationalization and allow the `stopifnot()` test code above. My job.**

2 sMoCA-based, and 2 based on the Level II battery[3].

Statistical Analyses

For sample description, we summarized continuous variables using mean, standard deviation, median, minimum, and maximum values. Categorical variables were summarized by the number of patients in each category. To address study objectives, we started by repeatedly assigning each patient the diagnosis of probable PDD based on each PDD operationalization listed in Table 1 resulting in a 204 (patients) × 68 (operationalizations) matrix where each cell indicates whether a patient (row) meets criteria for probable PDD according to an operationalization (column). PDD prevalence estimates were then computed as $\frac{N_{PDD}}{N_{total}}$ separately for each operationalization to address RO1.

To address RO2-4, a set of two class cross-tabulations with associated statistics was computed for each pair of operationalizations via the `confusionMatrix()` function from the R package *caret* (Kuhn, 2008). For each pair of operationalizations, the analysis was repeated twice such that each variables of the pair served once as the reference and once as the predictor. Following measures were used to evaluate pairwise concordance between different operationalizations of PDD criteria: 1) Cohen's κ with its 95% confidence interval (CI) computed via the `cohen.kappa()` function from the R package *psych* (William Revelle, 2024); 2) Accuracy (i.e., the proportion of correct predictions, both true positives and true negatives, among the total number of cases) with its 95% CI; 3) Sensitivity/Recall (i.e., the proportion of true positives); and 4) Specificity (i.e., the proportion of true negatives).[4]

Finally, the No Information Rate (NIR) was calculated for each pair of operationalizations. NIR is the accuracy that could be obtained by always predicting the majority class and in our case it is equivalent to the complement of the PDD prevalence

---

[3] **JM: Need to add Level II operationalization and allow the `stopifnot()` test code above. My job.**

[4] Unlike Cohen's κ, Accuracy, Sensitivity and Specificity are not symmetrical, i.e., their value depend on which variable is considered reference and which is considered predictor. Consequently, we report these values twice for each pair of operationalizations. Note that the Sensitivity of a reference/predictor pair corresponds to the Positive Predictive Value if their roles were reversed. The same relationship holds true between the Specificity and the Negative Predictive Value.

estimate according to the reference operationalization. Accuracy of prediction was subsequently compared to the NIR via a one-sided Exact Binomial Test as implemented by the `binom.test()` R stats function. Reference/predictor pairs associated with $p < .05$ were considered to show significantly better accuracy than NIR. In other word, for reference/predictor pairs associated with $p < .05$, we conclude that knowing the probable PDD status according to the predictor operationalization helps to estimate the probable PDD status according to the reference operationalization and the two opeartionalizations thus show substantial concordance.

Data wrangling and visualizations were done in the *tidyverse* package (Wickham et al., 2019) and tables were formatted in the *gt* package (Iannone et al., 2024). All analyses were conducted with the R (version 4.4.1) software environment for statistical computing (R Core Team, 2024). The software code supporting this article is available at https://github.com/josefmana/DemCr1t.git.[5]

Results

Sample Description

A total of 204 patients were included. Demographic and clinical characteristics of the sample are summarized in Table 2. **...**[6]

[Insert Table 2 here]

Prevalence Estimates

Operationalization-wise prevalence estimates are presented in Table 3. On average, estimated prevalence was 6.03% (SD = 3.46, Md = 3.94, range 2.00-16.75). Notably, the prevalence estimate was substantially lower when FAQ item 9 was used as a criterion of IADL deficit (M = 3.11%, SD = 0.48, range 2.00-3.94) compared to using total FAQ score criterion (M = 8.95%, SD = 2.54, range 3.50-16.75) as demonstrated also in Figure 1.[7]

---

[5] **JM: Do not forget to make it public before submitted!**

[6] **JM: Give some brief impressions from the Table 2 here. E.g., evaluate sex proportion, age, (we are missing education for some reason) etc. with general PD population (e.g., from some meta-analysis). Maybe say a word or two about mean cognitive profile or its spread.**

[7] **JM: For oneself - work on the code here, it ain't good.**

[Insert Table 3 and Figure 1[8] here]

Criteria Concordance

Results of the analyses of prediction Accuracy, Cohen's κ, Sensitivity and Specificity are presented in Figure 2, Figure A1, Figure A2 and Figure A3 respectively. Numerical results are available … [9]

In this section, we need to discuss among ourselves what and how to report such that the results give enough information, the most important information and are readable as well.

[Insert Figure 2[10] here]

Discussion

This study evaluated the applicability and validity of diagnostic frameworks for diagnosing probable PDD within a cohort of patients considered for deep brain stimulation (DBS). Our results demonstrate that diagnostic outcomes are markedly influenced by the chosen type of operationalization, particularly in relation to the assessment of the cognitive domains and IADLs.

Variability in Prevalence Estimates

One of the key findings of this study was the broad range of estimated PDD rate depending on the operationalization strategy. As seen in Table 3, the rate of PDD presence estimates varied from 2.00% to 16.75%, with the highest being derived from a combination of the sMoCA and the total FAQ score combination. When only FAQ item 9 was used to determine IADL deficits, the rate estimates were significantly lower (M = 3.09%, SD = 0.48), underscoring significant influence of functional assessment choice on diagnostic outcomes.

These findings are lower in comparison with previous research demonstrating wide variability in reported PDD rate among PD patients. For instance, retrospective study reported a PDD rate of 19.7% (Rana et al., 2011), while a clinical investigation by Aarsland et al. found

---

[8] **JM: For thyself - add description.**

[9] **JM: … as some kind of Supplementary Table, ideally html but Excel file would work as well.**

[10] **JM: For thyself - add description.**

even higher rate of around 30% (**?**). A recent complex meta-analysis synthesizing global data placed the pooled rate at 26.30% (Sousa et al., 2022). Compared to these estimates, our study reports generally lower prevalence rates, likely reflecting differences in sample characteristics, diagnostic criteria, and methodology.

Notably, our cohort consisted of patients being evaluated for DBS, a procedure typically reserved for individuals with relatively preserved cognitive function, which inherently biases against higher PDD prevalence rate. These comparisons emphasize that the observed lower prevalence in our sample is likely attributable to the preselection of cognitively intact individuals for DBS consideration, as well as to methodological variations such as the use of brief screening tool and differing IADL operationalizations.[11]

Cognitive and Clinical Context [12]

The relatively low prevalence of PDD across operationalizations may also reflect the relatively preserved cognitive profile of the DBS cohort, as evident in Table 2. The mean MoCA (M = 24.07, SD = 3.48) and MMSE (M = 26.69, SD = 2.22) scores suggest that on average the patients were functioning at a globally intact level. This likely reflects the pre-selection of cognitively preserved individuals for DBS, in line with standard eligibility criteria (**?**).

Memory performance, such as RAVLT and BVMTR delayed recall scores, also pointed to only mild deficits, particularly in domains central to the PDD vs. PD-MCI distinction. This aligns with previous findings that Level I criteria may overestimate dementia in patients with subtle impairments unless operational thresholds are rigorously defined (Aarsland et al., 2021; Bezdicek et al., 2016).

Implications for DBS Eligibility and Clinical Practice

Our findings also bear direct implications for clinical decision-making, especially in the context of surgical candidacy for DBS. Given that PDD remains a contraindication for DBS, the observed diagnostic instability could result in disparate treatment decisions based solely on which cognitive criteria are employed. Thus, harmonization of assessment

[11] **Keep it?**

[12] **Keep this section?**

procedures and operational cutoffs is essential to ensure equitable access to surgical treatment while maintaining diagnostic rigor.

Further, our data support the use of multidimensional assessment strategies, such as Level II neuropsychological batteries or regression-based normative comparisons, as confirmatory tools in diagnostically ambiguous cases. These approaches may enhance diagnostic precision and mitigate the risks of both false positives and unjustified treatment exclusion.

Constraints on generalizability

This study's generalizability is limited by its retrospective design and the homogeneity of the DBS candidate cohort, which may not reflect broader PD populations with varying cognitive profiles. Additionally, reliance on Czech normative data may restrict international applicability, although it represents real-world clinical standards within the region (**?**).

Future Directions

Future research should focus on prospective validation of updated Level I criteria against longitudinal functional outcomes and neurobiological markers. Additionally, the development of adaptive diagnostic algorithms that can integrate performance across cognitive, functional, and psychiatric domains may enhance diagnostic sensitivity while reducing the risk of over-classification.

Conclusions

This study systematically investigated the application of multiple Level I diagnostic criteria for Parkinson's disease dementia (PDD) within a cohort of patients considered for deep brain stimulation (DBS). The findings reveal substantial variability in prevalence estimates, strongly influenced by the choice of cognitive screening instruments and the operationalization of functional impairment. The divergence observed across operationalizations demonstrates the sensitivity of diagnostic outcomes to seemingly minor methodological choices.

The proposed revisions to the diagnostic framework (Kulisevsky et al., 2024) criteria offer enhanced sensitivity by leveraging MoCA-based components and broader IADL assessments, the use must be cautiously calibrated to prevent over-diagnosis in populations with mild or borderline cognitive deficits. Conversely, overly conservative criteria, such as

reliance on pill questionnaire (i.e. FAQ item 9 equivalence) may fail to detect meaningful functional decline and thus under-identify true cases of probable PDD.

Ultimately, this study contributes to the improving landscape of PDD diagnostics by offering empirical evidence for the refinement of Level I criteria and reinforcing the value of psychometric rigor in clinical neuropsychology. Future work should extend this validation to longitudinal trajectories and integrate neurobiological correlates, ensuring that cognitive criteria remain both scientifically grounded and clinically actionable.

References

Aarsland, D., Batzu, L., Halliday, G. M., Geurtsen, G. J., Ballard, C., Ray Chaudhuri, K., & Weintraub, D. (2021). Parkinson disease-associated cognitive impairment. *Nature Reviews Disease Primers*, *7*(1). https://doi.org/10.1038/s41572-021-00280-3

Beck, A. T., Steer, R. A., & Brown, G. (1996). *Beck depression inventory–II*. American Psychological Association (APA). https://doi.org/10.1037/t00742-000

Benedict, R. H. B. (1997). *Brief visuospatial memory test revised: Professional manual*. Psychological Assessment Resources.

Benton, A. L., Hamsher, K. D., & Sivan, A. B. (1989). *Multilingual aphasia examination*. AJA Associates.

Benton, A. L., Varney, N. R., & Hamsher, K. (1983). Visuospatial judgment: A clinical test. *Archives of Neurology*, *40*(3), 429–432.

Bezdicek, O., M., Č., Moore, T. M., Georgi, H. S., Sulc, Z., Wolk, D. A., Weintraub, D. A., Moberg, P. J., Jech, R., Kopecek, M., & Roalf, D. R. (2020). Determining a short form Montreal Cognitive Assessment (s-MoCA) Czech version: Validity in mild cognitive impairment Parkinson's disease and cross-cultural comparison. *Assessment*, *27*(8), 1960–1970. https://doi.org/10.1177/1073191118778896

Bezdicek, O., Motak, L., Axelrod, B. N., Preiss, M., Nikolai, T., Vyhnalek, M., Poreh, A., & Ruzicka, E. (2012). Czech Version of the Trail Making Test: Normative Data and Clinical Utility. *Archives of Clinical Neuropsychology*, *27*(8), 906–914. https://doi.org/10.1093/arclin/acs084

Bezdicek, O., Stepankova, H., Martinec Novakova, L., & Kopecek, M. (2016). Toward the

processing speed theory of activities of daily living in healthy aging: Normative data of the functional activities questionnaire. *Aging Clinical and Experimental Research*, *28*, 239–247.

Bezdicek, O., Sulc, Z., Nikolai, T., Stepankova, H., Kopecek, M., Jech, R., & Růžička, E. (2017). A parsimonious scoring and normative calculator for the Parkinson's disease mild cognitive impairment battery. *The Clinical Neuropsychologist*, *31*(6-7), 1231–1247. https://doi.org/10.1080/13854046.2017.1293161

Bezdíček, O., Georgi, H., Nikolai, T., & Kopeček, M. (2021). *Pražská verze stroopova testu*. Karolinum. https://karolinum.cz/en/books/bezdicek-prazska-verze-stroopova-testu-25158

Bezdíček, O., Lukavský, J., & Preiss, M. (2011). Functional activities questionnaire, czech version – a validation study. *Česká a Slovenská Neurologie a Neurochirurgie*, *74*(107), 36–42. https://www.csnn.eu/casopisy/ceska-slovenska-neurologie/2011-1/validizacni-studie-ceske-verze-dotazniku-faq-34140

Cereda, E., Cilia, R., Klersy, C., Siri, C., Pozzi, B., Reali, E., Colombo, A., Zecchinelli, A. L., Mariani, C. B., Tesei, S., Canesi, M., Sacilotto, G., Meucci, N., Zini, M., Isaias, I. U., Barichella, M., Cassani, E., Goldwurm, S., & Pezzoli, G. (2016). Dementia in parkinson's disease: Is male gender a risk factor? *Parkinsonism & Related Disorders*, *26*, 67–72. https://doi.org/10.1016/j.parkreldis.2016.02.024

Ciharova, M., Cígler, H., Dostálová, V., Šivicová, G., & Bezdicek, O. (2020). Beck depression inventory, second edition, Czech version: demographic correlates, factor structure and comparison with foreign data. *International Journal of Psychiatry in Clinical Practice*, *24*(4), 371–379. https://doi.org/10.1080/13651501.2020.1775854

Conway, J. M., Jako, R. A., & Goodman, D. F. (1995). A Meta-Analysis of Interrater and Internal Consistency Reliability of Selection Interviews. *Journal of Applied Psychology*, *80*(5), 565. https://doi.org/10.1037/0021-9010.80.5.565

Deuschl, G., Schade-Brittinger, C., Krack, P., Volkmann, J., Schäfer, H., Bötzel, K., Daniels, C., Deutschländer, A., Dillmann, U., Eisner, W., Gruber, D., Hamel, W., Herzog, J., Hilker, R., Klebe, S., Kloß, M., Koy, J., Krause, M., Kupsch, A., … Voges, J. (2006). A Randomized Trial of Deep-Brain Stimulation for Parkinson's Disease. *New England*

Journal of Medicine, *355*(9), 896–908. https://doi.org/10.1056/nejmoa060281

Dubois, B., Burn, D., Goetz, C., Aarsland, D., Brown, R. G., Broe, G. A., Dickson, D., Duyckaerts, C., Cummings, J., Gauthier, S., Korczyn, A., Lees, A., Levy, R., Litvan, I., Mizuno, Y., McKeith, I. G., Olanow, C. W., Poewe, W., Sampaio, C., … Emre, M. (2007). Diagnostic procedures for Parkinson's disease dementia: Recommendations from the movement disorder society task force. *Movement Disorders*, *22*(16), 2314–2324. https://doi.org/10.1002/mds.21844

Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., Broe, G. A., Cummings, J., Dickson, D. W., Gauthier, S., Goldman, J., Goetz, C., Korczyn, A., Lees, A., Levy, R., Litvan, I., McKeith, I., Olanow, W., Poewe, W., … Dubois, B. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Movement Disorders*, *22*(12), 1689–1707. https://doi.org/10.1002/mds.21507

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Frydrychová, Z., Kopeček, M., Bezdíček, O., & Štěpánková, J. (2018). České normy pro revidovaný reyův auditorně-verbální test učení (RAVLT) pro populaci starších osob. *Československá Psychologie*, *62*(4), 330–349. https://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.02da6923-6394-4d8d-b018-dd6e3b2503be

Havlík, F., Mana, J., Dušek, P., Jech, R., Ržička, E., Kopeček, M., Georgi, H., & Bezdicek, O. (2020). Brief visuospatial memory test-revised: Normative data and clinical utility of learning indices in parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, *42*(10), 1099–1110.

Hoogland, J., Boel, J. A., Bie, R. M. A. de, Geskus, R. B., Schmand, B. A., Dalrymple-Alford, J. C., Marras, C., Adler, C. H., Goldman, J. G., Tröster, A. I., Burn, D. J., Litvan, I., & Geurtsen, G. J. (2017). Mild cognitive impairment as a risk factor for Parkinson's disease dementia. *Movement Disorders*, *32*(7), 1056–1065. https://doi.org/10.1002/mds.27002

Hoogland, J., Boel, J. A., Bie, R. M. A. de, Schmand, B. A., Geskus, R. B., Dalrymple-Alford, J. C., Marras, C., Adler, C. H., Weintraub, D., Junque, C., Pedersen, K. F., Mollenhauer, B., Goldman, J. G., Tröster, A. I., Burn, D. J., Litvan, I., & Geurtsen, G. J. (2019). Risk of

Parkinson's disease dementia related to level I MDS PD-MCI. *Movement Disorders*, *34*(3), 430–435. https://doi.org/10.1002/mds.27617

Iannone, R., Cheng, J., Schloerke, B., Hughes, E., Lauer, A., Seo, J., Brevoort, K., & Roy, O. (2024). *Gt: Easily create presentation-ready display tables*. https://CRAN.R-project.org/package=gt

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston naming test*. Lea & Febiger.

Kopecek, M., Stepankova, H., Lukavsky, J., Ripova, D., Nikolai, T., & Bezdicek, O. (2016). Montreal cognitive assessment (MoCA): Normative data for old and very old Czech adults. *Applied Neuropsychology: Adult*, *24*(1), 23–29. https://doi.org/10.1080/23279095.2015.1065261

Kuhn, M. (2008). Building predictive models in r using the caret package. *Journal of Statistical Software*, *28*(5), 1–26. https://doi.org/10.18637/jss.v028.i05

Kulisevsky, J., Litvan, I., Weintraub, D., Goldman, J. G., Tröster, A. I., Lewis, S. J., Parkinson, I., Group, M. D. S. P.-M. S., et al. (2024). A call for change: Updating the operational definition for dementia in parkinson's disease. *Movement Disorders Clinical Practice*.

Litvan, I., Goldman, J. G., Tröster, A. I., Schmand, B. A., Weintraub, D., Petersen, R. C., Mollenhauer, B., Adler, C. H., Marder, K., Williams-Gray, C. H., Aarsland, D., Kulisevsky, J., Rodriguez-Oroz, M. C., Burn, D. J., Barker, R. A., & Emre, M. (2012). Diagnostic criteria for mild cognitive impairment in Parkinson's disease: *Movement* Disorder Society Task Force guidelines. *Movement Disorders*, *27*(3), 349–356. https://doi.org/10.1002/mds.24893

Meireles, J., & Massano, J. (2012). Cognitive impairment and dementia in parkinsons disease: Clinical features, diagnosis, and management. *Frontiers in Neurology*, *3*. https://doi.org/10.3389/fneur.2012.00088

Mullner, J., Ruisl, I., & Farkas, G. (1980). *Dotaznik na meranie uzkosti a uzkostlivosti - STAI*. Bratislava: Psychodiagnostické a didaktické testy.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics*

Society, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Oh, Y.-S., Kim, J.-S., Park, I.-S., Shim, Y.-S., Song, I.-U., Park, J.-W., Lee, P.-H., Lyoo, C.-H., Ahn, T.-B., Ma, H.-I., et al. (2016). Prevalence and treatment pattern of p arkinson's disease dementia in k orea. *Geriatrics & Gerontology International*, *16*(2), 230–236.

Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Chance, J. M., & Filos, S. (1982). Measurement of Functional Activities in Older Adults in the Community. *Journal of Gerontology*, *37*(3), 323–329. https://doi.org/10.1093/geronj/37.3.323

Postuma, R. B., Berg, D., Stern, M., Poewe, W., Olanow, C. W., Oertel, W., Obeso, J., Marek, K., Litvan, I., Lang, A. E., Halliday, G., Goetz, C. G., Gasser, T., Dubois, B., Chan, P., Bloem, B. R., Adler, C. H., & Deuschl, G. (2015). MDS clinical diagnostic criteria for Parkinson's disease. *Movement Disorders*, *30*(12), 1591–1601. https://doi.org/10.1002/mds.26424

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rana, A. Q., Yousuf, M. S., Naz, S., & Qa'aty, N. (2011). Prevalence and relation of dementia to various factors in parkinson's disease. *Psychiatry and Clinical Neurosciences*, *65*(4), 317–321. https://doi.org/10.1111/j.1440-1819.2011.02291.x

Reitan, R. (2004). The Trail Making Test as an initial screening procedure for neuropsychological impairment in older children. *Archives of Clinical Neuropsychology*, *19*(2), 281–288. https://doi.org/10.1016/s0887-6177(03)00042-8

Rey, A. (1964). *L'examen clinique en psychologie (the clinical psychological examination)*. Presses Universitaires de France.

Royall, D. R., Cordes, J. A., & Polk, M. (1998). CLOX: an executive clock drawing task. *Journal of Neurology, Neurosurgery & Psychiatry*, *64*(5), 588–594. https://doi.org/10.1136/jnnp.64.5.588

Sousa, C. S. e, Alarcão, J., Martins, I. P., & Ferreira, J. J. (2022). Frequency of dementia in parkinson's disease: A systematic review and meta-analysis. *Journal of the Neurological Sciences*, *432*, 120077.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. &. (1983).

Manual for the state-trait anxiety inventory. Palo Alto, CA: Consulting Psychologists Press.

Stepankova, H., Nikolai, T., Lukavsky, J., Bezdicek, O., Vrajova, M., & Kopecek, M. (2015). Mini-Mental State Examination – česká normativní studie. *Ceska a slovenska neurologie a neurochirurgie*, *78*(111), 57–63.

Wechsler, D. (1997). *Wechsler adult intelligence scale—third edition (WAIS-III)*. Psychological Corporation. https://books.google.cz/books/about/Wais_III_Wechsler_Adult_Intelligence_Sca.html?id=qTCuGQAACAAJ

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

William Revelle. (2024). *Psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych

Yamashita, K. Y., Bhoopatiraju, S., & Silverglate, G. T., Bret D. aGrossberg. (2023). Biomarkers in Parkinson's disease: A state of the art review. *Biomarkers in Neuropsychiatry*, *9*, 100074. https://doi.org/10.1016/j.bionps.2023.100074

Table 1

group	type	glob	glob_t	atte	atte_t	exec	exec_t	memo	memo_t
mmse	mmse1	mmse	26	mmse_7	4	cloc	2	mmse_3words	3
mmse	mmse2	mmse	26	mmse_7	4	cloc	2	mmse_3words	3
mmse	mmse3	mmse	26	mmse_7	4	vf_s	10	mmse_3words	3
mmse	mmse4	mmse	26	mmse_7	4	vf_s	10	mmse_3words	3
moca	moca1	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	1
moca	moca2	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	1
moca	moca3	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	1
moca	moca4	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	1
moca	moca5	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	2
moca	moca6	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	2
moca	moca7	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	2
moca	moca8	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	2
moca	moca9	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	3
moca	moca10	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	3
moca	moca11	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	3
moca	moca12	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	3
moca	moca13	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	4
moca	moca14	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	4
moca	moca15	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	4
moca	moca16	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	4
moca	moca17	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	5
moca	moca18	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	5
moca	moca19	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	5
moca	moca20	moca_total	26	moca_7	3	moca_cloc	2	moca_5words	5
moca	moca21	moca_total	26	moca_7	3	vf_k	11	moca_5words	1
moca	moca22	moca_total	26	moca_7	3	vf_k	11	moca_5words	1
moca	moca23	moca_total	26	moca_7	3	vf_k	11	moca_5words	1
moca	moca24	moca_total	26	moca_7	3	vf_k	11	moca_5words	1
moca	moca25	moca_total	26	moca_7	3	vf_k	11	moca_5words	2
moca	moca26	moca_total	26	moca_7	3	vf_k	11	moca_5words	2

Table 2

Table 2</br> Sample description. Demographic, clinical and cognitive characteristics of the sample.

	N	Md	Min-max	M
Demographics				
Sex of Participants (Males)	126 (62%)	-	-	-
Age of Participants (in Years)	203	60	34-73	58.96
Education (in Years)	204	13	8-24	13.75
Clinical				
Type of PD (Akinetic-rigid, Tremor-dominant, Axial)	32/107	-	-	-
Hoehn Yahr stage (Range 0-5)	2/8/66/33/14/2	-	-	-
PD duration (in Years)	162	10	1-25	10.71
L-DOPA (in Miligrams)	138	1607	0-4138	1697.58
UPDRS III off state (Range 0-132)	135	35	10-81	37.43
UPDRS III on state (Range 0-132)	135	14	2-45	15.97
MMSE				
Total score (Range 0-30)	203	27	15-30	26.69
Sevens (Range 0-5)	1/2/8/20/34/139	-	-	-
VF S (Number of Words per Minute)	202	15	1-34	14.95
Clock Drawing (Range 0-2)	26/91/86	-	-	-
Pentagons (Range 0-1)	187 (92%)	-	-	-
Three words (Range 0-3)	5/14/60/124	-	-	-
MoCA				
Total score (Range 0-30)	203	24	9-30	24.07
sMoCA total score (Range 0-16)	203	11	1-16	11.26
Sevens (Range 0-3)	1/2/29/171	-	-	-
VF K (Number of Words per Minute)	204	16	0-29	15.50
Clock drawing (Range 0-3)	24/83/96	-	-	-
Cube drawing (Range 0-1)	164 (81%)	-	-	-

Table 3

Table 3</br> Estimates of prevalence. Estimates of the prevalence of probable PD-D in the sample.

	Impaired cognition				
Global deficit	Attention	Executive function	Construction	Memory	Lan
sMoCA total score < 13	-	-	-	-	
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 4	Abstra
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 5	Abstra
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 2	Abstra
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 3	Abstra
Total score < 26	Sevens < 3	VF K < 11	Cube drawing < 1	Five words < 4	Abstra
Total score < 26	Sevens < 3	VF K < 11	Cube drawing < 1	Five words < 5	Abstra
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 5	Animal r
character(0) < 1	-	-	-	-	
Total score < 26	Sevens < 3	Clock drawing < 2	Cube drawing < 1	Five words < 4	Abstra
Total score < 26	Sevens < 3	Clock drawing < 2	Cube drawing < 1	Five words < 5	Abstra
Total score < 26	Sevens < 3	VF K < 11	Cube drawing < 1	Five words < 3	Abstra
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 4	Animal r
Total score < 26	Sevens < 3	Clock drawing < 2	Cube drawing < 1	Five words < 3	Abstra
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 1	Abstra
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 3	Animal r
Total score < 26	Sevens < 3	Clock drawing < 2	Cube drawing < 1	Five words < 2	Abstra
Total score < 26	Sevens < 3	VF K < 11	Cube drawing < 1	Five words < 2	Abstra
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 2	Animal r
Total score < 26	Sevens < 3	Clock drawing < 2	Cube drawing < 1	Five words < 1	Abstra
Total score < 26	Sevens < 3	Clock drawing < 2	Cube drawing < 1	Five words < 5	Animal r
Total score < 26	Sevens < 3	Clock drawing < 2	Cube drawing < 1	Five words < 3	Animal r
Total score < 26	Sevens < 3	Clock drawing < 2	Cube drawing < 1	Five words < 4	Animal r
Total score < 26	Sevens < 3	VF K < 11	Cube drawing < 1	Five words < 1	Abstra
Total score < 26	Sevens < 3	VF K < 11	Cube drawing < 1	Five words < 5	Animal r
Total score < 26	Sevens < 3	Clock drawing < 3	Cube drawing < 1	Five words < 1	Animal r

Figure 1

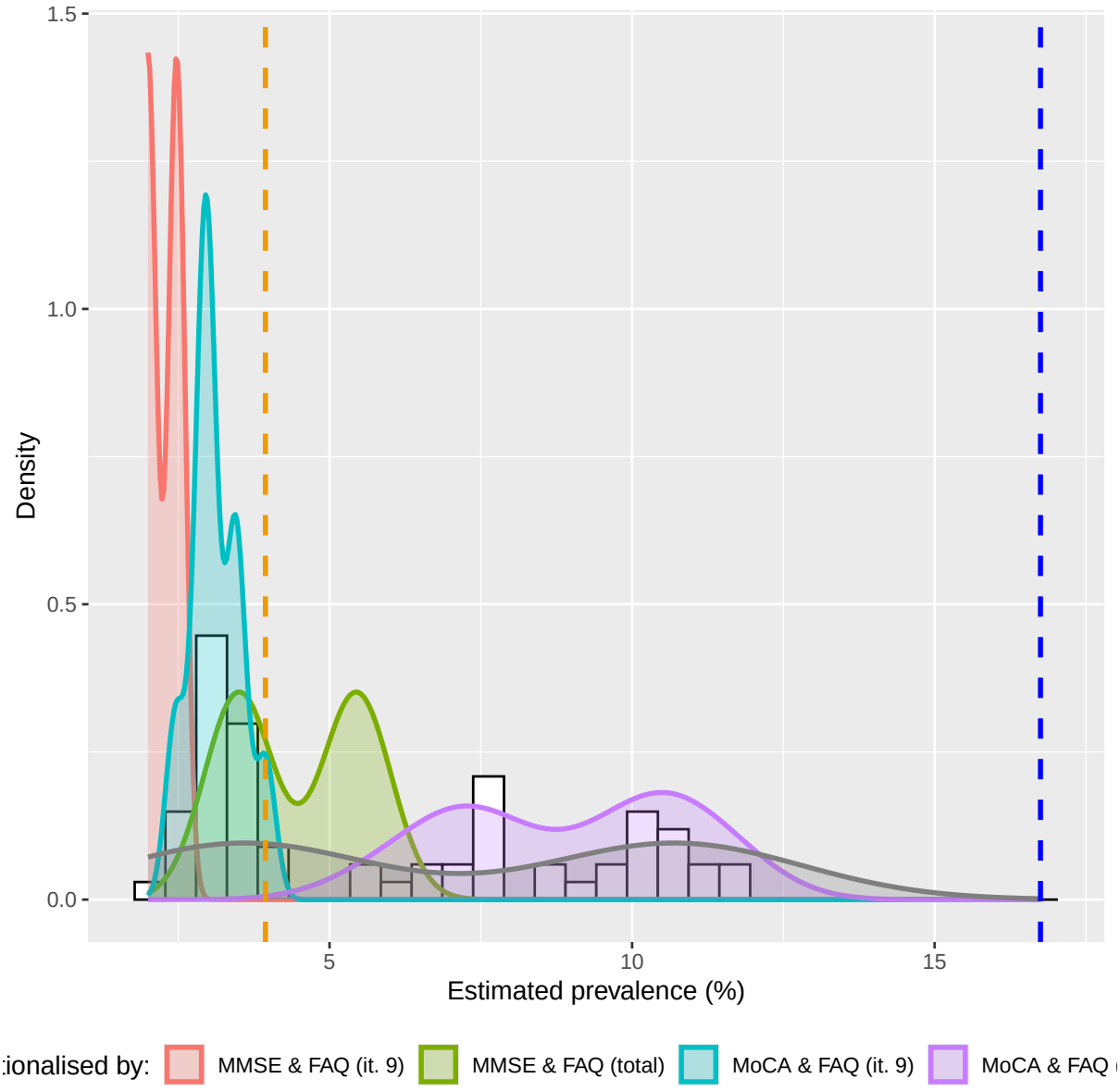
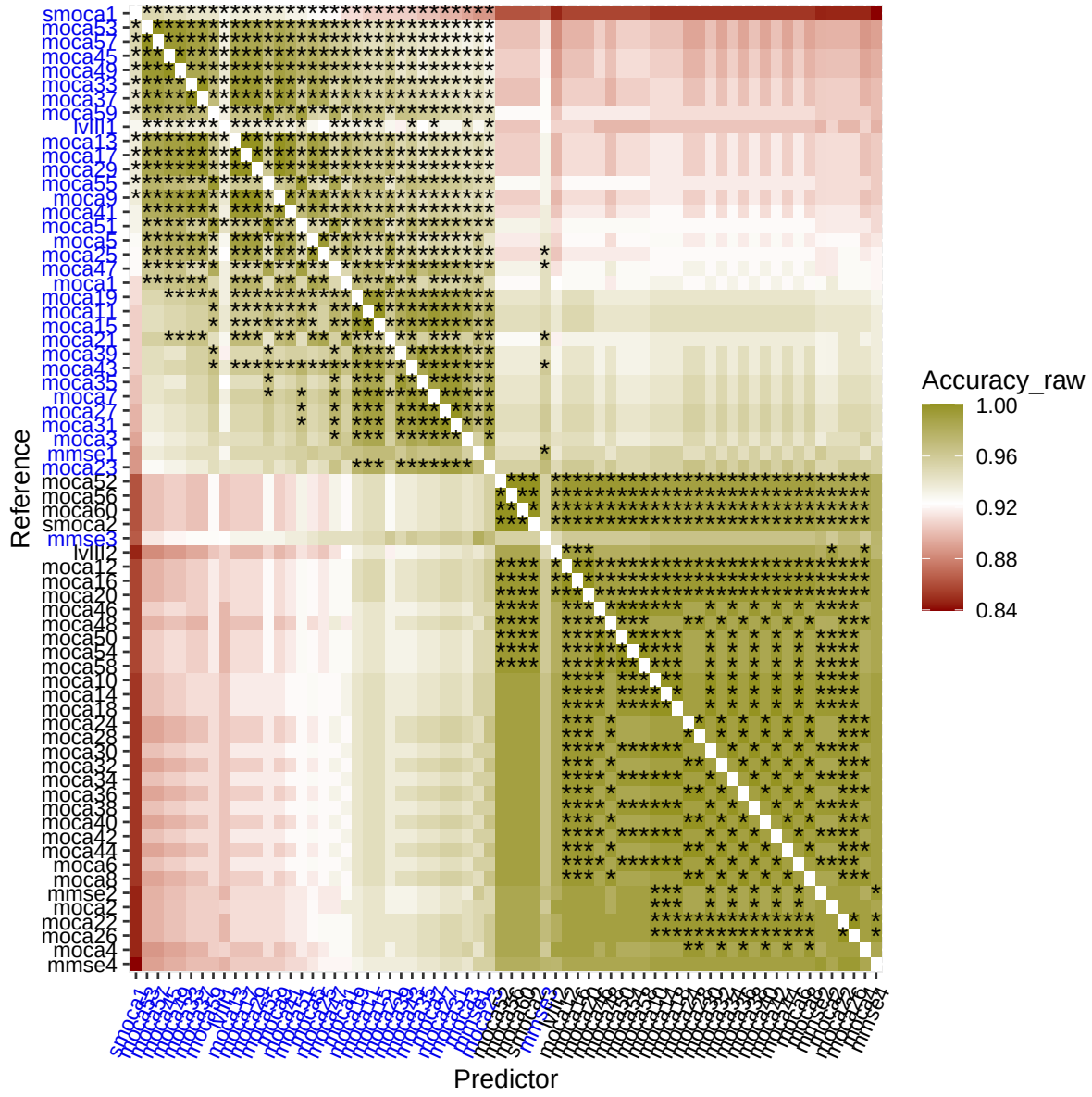

1.5
1.0
0.5
0.0
Density
5
10
15
Estimated prevalence (%)
:ionalised by:
MMSE & FAQ (it. 9)
MMSE & FAQ (total)
MoCA & FAQ (it. 9)
MoCA & FAQ

Figure 2



Appendix

Figure A1

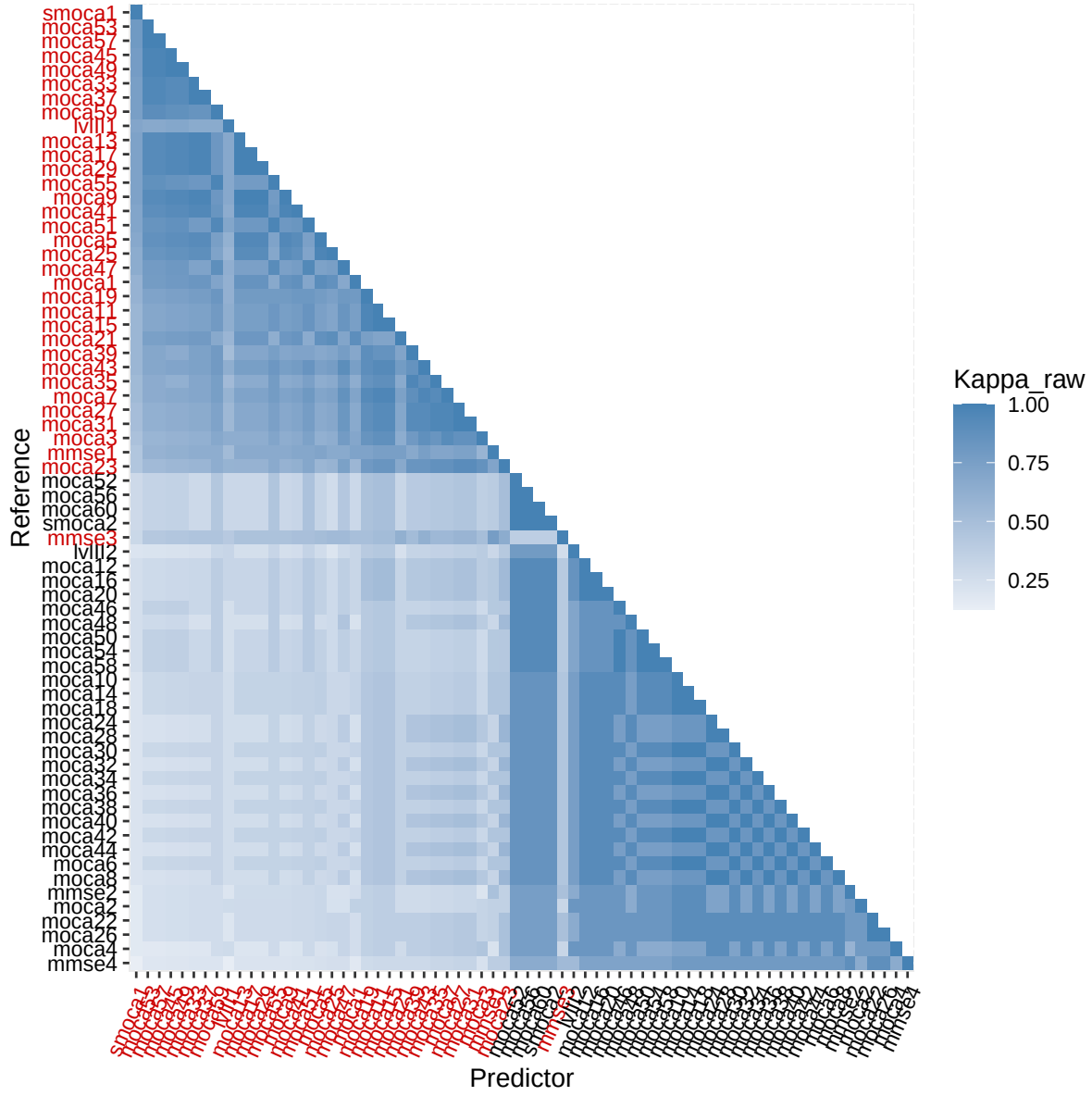


Figure A2

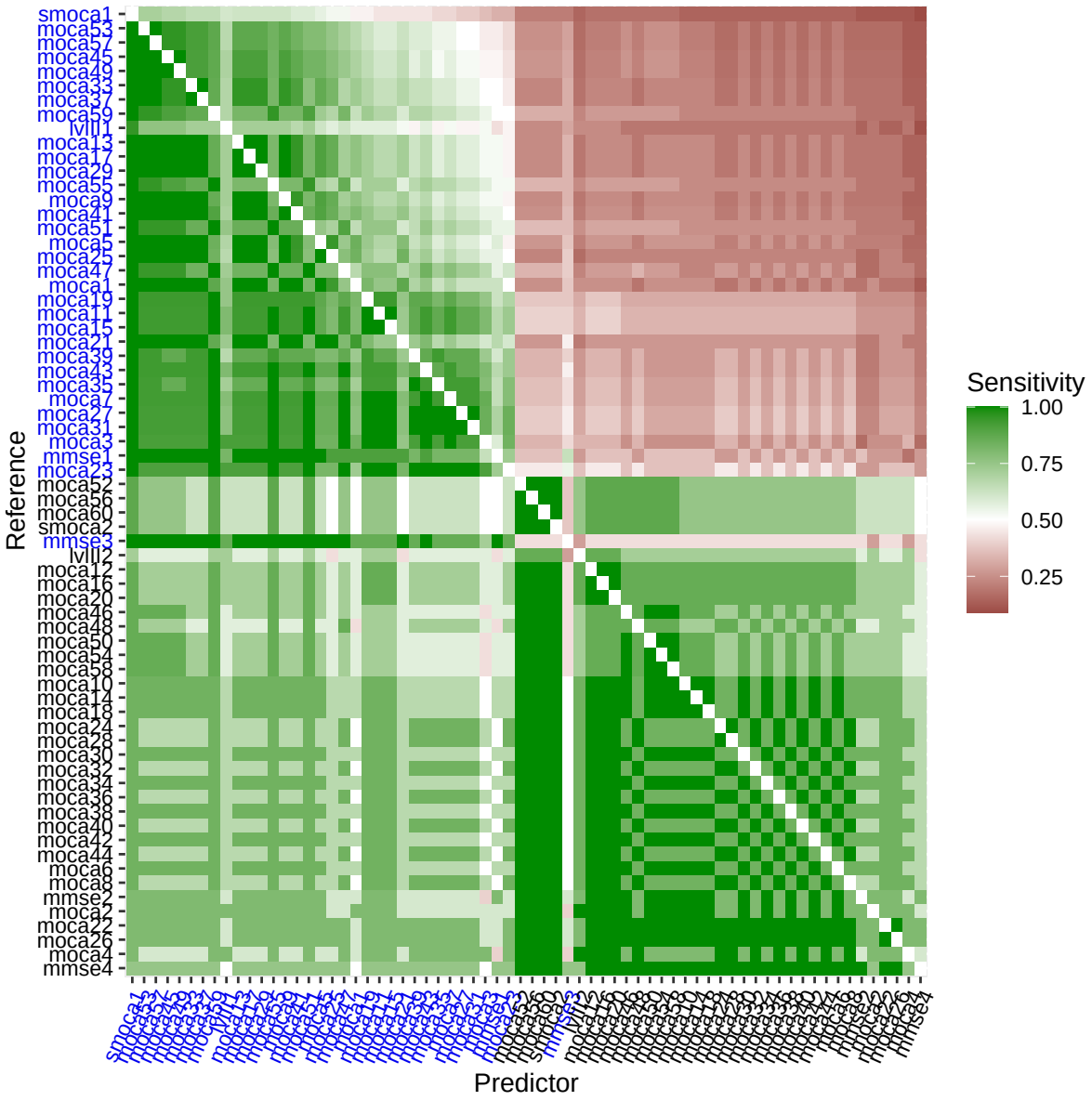


Figure A3