

Maximizing Fidelity of Neuropsychology Assessments in Fully Remote Studies

Keera N Fishman, Paula M. McLaughlin, Brian Tan, Angela K. Troyer, Joseph B. Orange, Angela C. Roberts, Donna Kwan, Brian Levine, Natalie Rashkovan, Richard H. Swartz, HANDDS-ONT Investigators

Submitted to: Journal of Medical Internet Research
on: April 22, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 34

 Figures 35

 Figure 1..... 36

 Figure 2..... 37

 Figure 3..... 38

 Figure 4..... 39

Maximizing Fidelity of Neuropsychology Assessments in Fully Remote Studies

Keera N Fishman^{1, 2, 3} Ph.D.; Paula M. McLaughlin^{3, 4, 5, 6} Ph.D.; Brian Tan⁷ MSc; Angela K. Troyer^{1, 8} Ph.D.; Joseph B. Orange^{9, 10} PhD; Angela C. Roberts^{11, 10} PhD; Donna Kwan³ Ph.D.; Brian Levine⁷ Ph.D., C.Psych.; Natalie Rashkovan¹² MHSc; Richard H. Swartz^{12, 13, 14} PhD, MD; HANDDS-ONT Investigators¹²

¹ Neuropsychology & Cognitive Health Department Baycrest Hospital Toronto CA

² Ontario Shores Centre for Mental Health Sciences Whitby CA

³ Centre for Neuroscience Studies Queen's University Kingston CA

⁴ Nova Scotia Health Authority Halifax CA

⁵ Department of Medicine (Geriatrics) Dalhousie University Halifax CA

⁶ Department of Psychology & Neuroscience Dalhousie University Halifax CA

⁷ Rotman Research Institute, Baycrest Hospital Toronto CA

⁸ Psychology Department University of Toronto Toronto CA

⁹ Canadian Centre for Activity and Aging Western University London CA

¹⁰ School of Communication Sciences and Disorders Western University London CA

¹¹ Department of Computer Science Western University London CA

¹² Sunnybrook Health Science Centre Toronto CA

¹³ Sunnybrook Research Institute Toronto CA

¹⁴ Department of Medicine (Neurology) University of Toronto Toronto CA

Corresponding Author:

Keera N Fishman Ph.D.

Neuropsychology & Cognitive Health Department
Baycrest Hospital
3560 Bathurst Street
Toronto
CA

Abstract

Background: Remote, interdisciplinary, observational clinical studies and clinical trials are increasingly emerging in the scientific literature. Remote videoconference-based neuropsychological assessment offers numerous advantages, including improved access to care, reduced travel burden, and the ability to monitor patients over time. However, such methodologies present challenges, particularly regarding the fidelity of collected data. Data fidelity, defined as the accuracy, completeness, and consistency of data, can be compromised by technical issues, variability in testing environments, and risks of data loss. These challenges necessitate the development of robust, standardized protocols to ensure high-quality data collection and analysis.

Objective: This study aimed (1) to evaluate the data fidelity of a remote videoconference-administered neuropsychology protocol in the context of the Health in Aging, Neurodegenerative Diseases, and Dementias in Ontario (HANDDS-ONT) study, guided by the Ontario Neurodegenerative Disease Research Initiative (ONDRI); (2) to generate three roadmaps to support data fidelity procedures for future remote neuropsychological research; and (3) to characterize the sample and their neuropsychological assessment performance. Quality assurance and quality control procedures were implemented, and missing data, virtual environment-related errors, and outcomes of quality assurance and quality control measures were evaluated to assess data fidelity.

Methods: 148 participants (62% female; median age=67, mean=15.1 education years) completed the neuropsychology protocol. Data were analyzed as descriptive statistics.

Results: Implementing our quality assurance and control procedures, the average number of queries per participant was 5.59 during the data monitoring phase and 0.30 during the cleaning and curation pipeline phase, with only 0.34% of data missing.

Virtual environment factors, such as internet connectivity and distractions, had minimal impact on data quality as only 8 participants (5.4%) had 1-2 tasks impacted by the virtual environment resulting in missing data. The fidelity of the remote videoconference-administered neuropsychology protocol was comparable to that of in-person assessments, and our procedures effectively minimized missing data and reduced the need for data corrections. The study highlighted the feasibility of collecting reliable neuropsychology data remotely while identifying practical adaptations to mitigate potential challenges.

Conclusions: The current study highlights the potential of remote videoconference-administered neuropsychological assessment to deliver high-fidelity data in clinical and research settings. The protocol performed similarly to an in-person neuropsychology protocol as it pertained to quality control indices (i.e., missingness and number of data corrections needed), and few virtual environment-related factors impacted data missingness. Quality assurance and quality control measures were crucial for ensuring the data were collected robustly remotely. The proposed roadmaps offer a template for future studies, addressing common challenges in remote neuropsychological testing and enabling wider adoption of telehealth methodologies. By advancing standardized protocols, this research supports the ongoing evolution of remote, interdisciplinary approaches to studying and managing neurodegenerative diseases. Clinical Trial: CTO Project ID #3589

(JMIR Preprints 22/04/2025:76390)

DOI: <https://doi.org/10.2196/preprints.76390>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in https://preprints.jmir.org/preprint/76390 I will be able to make the full manuscript PDF available to all users.

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in https://preprints.jmir.org/preprint/76390 I will be able to make the full manuscript PDF available to all users.

Original Manuscript

Maximizing Fidelity of Neuropsychology Assessments in Fully Remote Studies

Dr. Keera N. Fishman, Ph.D., C.Psych.^{1,2,3}, Dr. Paula M. McLaughlin, Ph.D., C.Psych.^{3,4,5}, Brian Tan⁶, Dr. Angela K. Troyer, Ph.D., C.Psych.^{1,7}, Dr. Joseph B. Orange, Ph.D.^{8,9}, Dr. Angela C. Roberts, Ph.D.^{8,10}, Dr. Donna Kwan, Ph.D., C.Psych.³, Dr. Brian Levine, Ph.D., C.Psych.⁶, Natalie Rashkovan, MHSc¹¹, Dr. Richard H. Swartz, Ph.D., M.D.^{11,12,13}, & the HANDDDS-ONT Investigators

¹Neuropsychology & Cognitive Health, Baycrest Hospital, Toronto, Ontario, Canada

²Ontario Shores Centre for Mental Health Sciences, Whitby, Ontario, Canada

³Centre for Neuroscience Studies, Queen's University, Kingston, Ontario, Canada

⁴Nova Scotia Health Authority, Halifax, Nova Scotia, Canada

⁵Departments of Medicine (Geriatrics) and Psychology & Neuroscience, Dalhousie University, Halifax, Nova Scotia, Canada

⁶Rotman Research Institute, Baycrest Hospital, Toronto, Ontario, Canada

⁷Psychology Department, University of Toronto, Ontario, Canada

⁸School of Communication Sciences and Disorders, Western University, London, Ontario, Canada

⁹Canadian Centre for Activity and Aging, Western University, London, Canada

¹⁰Department of Computer Science, Western University, London, Ontario, Canada

¹¹Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

¹²Sunnybrook Research Institute, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

¹³Department of Medicine (Neurology), University of Toronto, Toronto, Ontario, Canada

Ethical Approval

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Sunnybrook Health Sciences Centre Research Ethics Board (CTO Project ID #3589, Clinical Trials #NCT05062512).

Consent to Participate

Informed consent was obtained from all individual participants and study partners included in the study.

Consent for Publication

The authors affirm that human research participants provided informed consent for publication of their data.

Contact information for the corresponding author: Dr. Keera Fishman, Neuropsychology & Cognitive Health, Baycrest Hospital, 3560 Bathurst Street, Toronto, ON, M6A 2E1, kfishman@baycrest.org.

Abstract

Background: Remote, interdisciplinary, observational clinical studies and clinical trials are increasingly emerging in the scientific literature. Remote videoconference-based neuropsychological assessment offers numerous advantages, including improved access to care, reduced travel burden, and the ability to monitor patients over time. However, such methodologies present challenges, particularly regarding the fidelity of collected data. Data fidelity, defined as the accuracy, completeness, and consistency of data, can be compromised by technical issues, variability in testing environments, and risks of data loss. These challenges necessitate the development of robust,

standardized protocols to ensure high-quality data collection and analysis.

Objective: This study aimed (1) to evaluate the data fidelity of a remote videoconference-administered neuropsychology protocol in the context of the Health in Aging, Neurodegenerative Diseases, and Dementias in Ontario (HANDDS-ONT) study, guided by the Ontario Neurodegenerative Disease Research Initiative (ONDRI); (2) to generate three roadmaps to support data fidelity procedures for future remote neuropsychological research; and (3) to characterize the sample and their neuropsychological assessment performance. Quality assurance and quality control procedures were implemented, and missing data, virtual environment-related errors, and outcomes of quality assurance and quality control measures were evaluated to assess data fidelity.

Methods: 148 participants (62% female; median age=67, mean=15.1 education years) completed the neuropsychology protocol in this observational clinical trial. Data were analyzed using descriptive statistics.

Results: Implementing our quality assurance and control procedures, the average number of queries per participant was 5.59 during the data monitoring phase and 0.30 during the cleaning and curation pipeline phase, with only 0.34% of data missing. Virtual environment factors, such as internet connectivity and distractions, had minimal impact on data quality as only 8 participants (5.4%) had 1-2 tasks impacted by the virtual environment resulting in missing data. The fidelity of the remote videoconference-administered neuropsychology protocol was comparable to that of in-person assessments, and our procedures effectively minimized missing data and reduced the need for data corrections. The study highlighted the feasibility of collecting reliable neuropsychology data remotely while identifying practical adaptations to mitigate potential challenges.

Conclusions: The current study highlights the potential of remote videoconference-administered neuropsychological assessment to deliver high-fidelity data in clinical and research settings. The protocol performed similarly to an in-person neuropsychology protocol as it pertained to quality control indices (i.e., missingness and number of data corrections needed), and few virtual environment-related factors impacted data missingness. Quality assurance and quality control measures were crucial for ensuring the data were collected robustly remotely. The proposed roadmaps offer a template for future studies, addressing common challenges in remote neuropsychological testing and enabling wider adoption of telehealth methodologies. By advancing standardized protocols, this research supports the ongoing evolution of remote, interdisciplinary approaches to studying and managing neurodegenerative diseases.

Keywords: Remote Neuropsychological Assessment; Neuropsychology; Observational Clinical Trials; Neurodegenerative diseases; Quality Assurance; Quality Control; Data Fidelity; Data Quality; e-Health

Introduction

Background

The use of remote videoconference technology to deliver health care assessment and treatment services from an individual's home continues to grow at an accelerated pace. The introduction of telehealth – defined as the remote delivery of care from a distance using technology such as telephone, videoconferencing, and emails - has paved the way for these advancements.

Recommendations for conducting remote videoconference-administered neuropsychological

assessments have been published by the Alzheimer Society of Canada [1], Ontario Psychological Association [2], American Psychological Association [3], and the Inter Organizational Practice Committee [4], along with others that predate the COVID-19 pandemic [e.g., 5,6]. The use of hybrid models of remote data collection may provide multiple novel benefits. These include opportunities to: 1) increase access to care for individuals living in remote areas (notwithstanding internet access for rural and remote regions), 2) increase access to care for individuals with physical mobility concerns, 3) improve convenience with respect to travel time, 4) reduce risk of physical exposure to illnesses, 5) eliminate the need for certain possible influencing factors on performance (e.g., mask, face shields, discomfort/anxiety in an unfamiliar environment), 6) support ongoing monitoring for more frequent assessments (enabling better tracking of disease progression and/or treatment efficacy), and 7) better support participants with hearing and vision difficulties, if resources are available [7-9]. However, standardized neuropsychological assessments often require modifications or can become increasingly difficult to administer via videoconferencing platforms; for example, tests requiring manipulation of materials or written responses, which pose possible test security issues [7,8]. Other important considerations include internet service reliability, connection speed, risks for distractibility, and the risk of data interception. When completed from the individual's home, videoconferencing administration also introduces a range of factors that may negatively influence the validity of the data collected. These include distractions in the room or on the device, availability of a quiet and private space, non-standardization of equipment (e.g., variation in the screen size, the audio/visual quality, and the processing speed on the device), variable internet access or quality of internet connection, technological glitches, and the risk of interception of electronic communications.

Studies using fully remote videoconference-administered supervised neuropsychological assessment have demonstrated the feasibility and reliability of several standard neuropsychological tasks [5,10,11]. However, less is known about the *fidelity* – the accuracy, completeness, and consistency – of the data produced in these studies, and in the context of individuals living with neurodegenerative diseases. Data fidelity procedures are essential to ensure that the data collected and used are of optimal quality (i.e., valid and reliable). Rigorous fidelity procedures are imperative to design and implement in order to ensure the data are accurate, precise, and are not missing due to systematic errors. Data fidelity is impacted by quality assurance (preventing errors) and quality control (identifying and correcting errors), which in turn can influence the validity of study findings [12]. Despite this observation, procedures ensuring methodological rigor in cohort studies that look at neuropsychological outcome data are often overlooked or seldom reported by researchers [for exceptions, see 12,13]. The absence of these robust, standardized protocols detailing test administration, scoring, data entry, and data curation (i.e., fidelity) can elevate the risk for generating low-quality data impacted by noise and systematic errors [14,15]. This may, in turn, lead to research that informs policy based on inaccurate or incomplete scientific conclusions. Given that many cohort studies investigate cognitive and behavioural data as outcomes in individuals living with neurodegenerative diseases, especially in repeated annual or semi-annual evaluations, minimizing the risks for noise and errors in the neuropsychology data are of paramount importance. Moreover, poor data quality may also contribute to the ongoing "replication crisis" in the psychology and life sciences research fields [16]. This issue has considerable financial repercussions given that many research projects are funded by public resources. In essence, the emergence of new hybrid and remote models of enquiry propel the need for further development and evaluation of protocols, or 'roadmaps', to navigate and to validate these novel avenues.

The current study, the Health in Aging, Neurodegenerative Diseases and Dementias in Ontario (HANDDS-ONT) study, was guided by the Ontario Neurodegenerative Disease Research Initiative [ONDRI; 17,18], which aims to improve the diagnosis, prognosis, care, and outcomes for Ontarians living with neurodegenerative diseases and those who support them. The initial foundational study of ONDRI (“ONDRI 1.0”) involved an in-person, prospective, longitudinal, observational study of individuals with neurodegenerative and cerebrovascular diseases across multiple research platforms (e.g., genomics, gait/balance, eye tracking, clinical, neuroimaging, neuropsychology, and SD-OCT imaging) and recruited experts across research, clinical, and neuroinformatics domains to develop data fidelity procedures [12,17,19].

The HANDDS-ONT study was inceptioned following the completion of ONDRI’s Remote Monitoring in Neurodegenerative Disease (ReMiNDD) Feasibility Study [20], which supported the feasibility and user acceptance of remote health and activity monitoring in individuals with neurodegenerative diseases using a multi-biosensor approach. The ReMiNDD protocol was a *hybrid* study that included synchronous, in-person visits (including the neuropsychology protocol) and asynchronous health and activity monitoring. HANDDS-ONT is a primarily remote, prospective, observational study that compares fluid biomarkers, clinical data (including remote wearable biosensors), and neuropsychology task performance among individuals with various neurodegenerative diseases as well as other community dwelling older adults without any neurodegenerative disease or neurologic disease diagnoses.

Goal of This Study

Given that data fidelity is equally as essential in the context of remote data collection, we adapted our in-person data fidelity protocol in ‘ONDRI 1.0’ to a remote videoconference format. The HANDDS-ONT protocol specifically accommodates training modalities and technical issues, and includes mitigating strategies to threats to validity that accompany the virtual testing environment. This manuscript describes a roadmap for our approach to quality assurance (i.e., the processes to prevent issues before they occur) and quality control (i.e., the processes to identify and address errors when they arise) to maximize data fidelity in a remote videoconference-administered neuropsychology protocol that can be used as a template for future studies. Interdisciplinary collaboration and data fidelity (quality assurance and quality control) procedures were instrumental to the organization, timely completion, and validity and reliability of data collection in successfully implementing our trial.

The aim of this manuscript is to describe the development of three roadmaps in the fully remote videoconference-administered neuropsychology protocol and to characterize the sample of participants who completed the neuropsychological tasks. Specifically, we address the:

- (1) Implementation of quality assurance and quality control procedures, including training protocols, use of assessment measures, monitoring procedure, and characterization of metrics including missing data and virtual environment errors, and;
- (2) The outcome of quality assurance and quality control metrics, including considerations for virtual neuropsychological testing, and the implementation of lessons learned and virtual adaptations from McLaughlin and colleagues’ [12] ONDRI 1.0 recommendations in the current protocol.

Supplementary Material A: Background

The Ontario Neurodegenerative Disease Research Initiative's (ONDRI) [17,18] Foundational study ("ONDRI 1.0") is a prospective, longitudinal, observational study of individuals with neurodegenerative and cerebrovascular diseases ("cross-cohort"). ONDRI 1.0 data were collected across multiple research platforms ("cross-platform") including clinical, eye tracking, gait/balance, genomics, molecular/neuropathology, neuroimaging, neuropsychology, and SD-OCT imaging. ONDRI 1.0 was supported by a Neuroinformatics and Biostatistics team, contributions from the aforementioned platforms, as well as the Patient and Community Advisory Committee. ONDRI infrastructure utilizes a secure neuroinformatics platform from the Ontario Brain Institute, Brain-CODE. The success of the ONDRI infrastructure lies in its advancement in neuroinformatics, data management procedures (including data collection, monitoring, cleaning, and curation), and characterization of cross-disease cohorts with longitudinal follow-up and linkages to long-term outcomes through administrative health record data. ONDRI researchers hold vested interests in developing and exploring new methodologies to continue contributing to the global scientific and medical community, including adapting these trial models to virtual, fully remote participant protocols for accruing neuropsychological and other interdisciplinary data.

HANDDS-ONT was designed to begin to explore these new methodologies. It is a "Master Observational Trial," connecting real-world clinical data with cellular, protein marker and genetic data, free-living behavioural data measured by wearable biosensors, cognitive performance measured by neuropsychological tasks, and longer-term health outcome data via linkage to administrative datasets through the Institute for Clinical Evaluative Sciences (ICES). The HANDDS-ONT study was inception following the completion of ONDRI's Remote Monitoring in Neurodegenerative Disease (ReMiNDD) Feasibility Study [20], which supported the feasibility and user acceptance of remote health and activity monitoring in individuals with neurodegenerative diseases using a multi-biosensor approach. The multi-biosensor approach was adopted given that free-living behaviours – collection of data such as sleep or activity, from people in their homes or community - can more accurately reflect disease burden, can provide more sensitive markers of impairment and decline, and have the potential to provide feedback that can guide self-care, caregiver decisions, or clinician interventions to improve function and independence.

The HANDDS-ONT neuropsychology protocol was implemented to address the study aims of linking free living behavioural data (e.g., sedentary behaviour, activity, sleep) with other behavioural, cognitive, and biomarker data, to characterize the cognitive profiles of the HANDDS-ONT cohort, to generate data that could be harmonized with other master observational trials, and to evaluate the feasibility of adapting a neuropsychology protocol to a remote testing environment with individuals with various self- and physician-reported neurodegenerative and vascular diseases. Eligible participants were Ontario residents, 18 years of age or older, with an English proficiency of 7/10 or higher on a modified language questionnaire (Language Experience and Proficiency Questionnaire, LEAP-Q; [21]). Participants were required to have telephone or internet access, be able to attend a Lifelabs facility for a blood sample, and be under the care of a primary care provider or specialist. All participants were living with a neurodegenerative disease, post-stroke, or were living independently with no diagnosis of a neurodegenerative disease. Exclusion criteria included participants with underlying conditions that may interfere with study participation or results, including contraindications to biosensors, a history of substance abuse within the past year or alcohol/drug abuse that could affect compliance, known brain tumors (e.g., glioblastoma), prior brain surgeries that may impact participation, difficulties with venous access or blood draws, and significant psychiatric disorders (e.g., untreated major depression or psychosis). The comprehensive

HANDDS-ONT protocol is in preparation.

Methods

Development of Protocol and Monitoring Protocol

The remote videoconference-administered neuropsychology protocol described in this manuscript was developed and implemented for robust characterization of cognitive abilities that would 1) be acceptable to participants with respect to time and study burden considerations, 2) assess a variety of cognitive domains, and 3) be validated for virtual use and testing. The neuropsychology measures incorporated in HANDDS-ONT (Table 1) consisted of 2 screening tools, 6 standard clinical neuropsychological tests, and 1 questionnaire. The measures were selected based on empirical support for their utility in a virtual testing environment [e.g., see 10,11 for reviews]. The measures included traditional clinical measures for which there are standardized data available. Participants were screened for hearing and vision and their proficiency in speaking and understanding English [21]. Finally, a self-report questionnaire regarding independence in instrumental activities of daily living (iADLs) was completed by participants to provide information regarding their perspective on their functional abilities.

The HANDDS-ONT Neuropsychology lead (K.F.) trained four study coordinators based at Sunnybrook Hospital in Toronto, Canada to administer the remote videoconference-administered neuropsychology protocol. The protocol was completed with all participants with a MoCA score of 14 or higher (max = 30) [12], which was completed at a prior virtual visit. The neuropsychology protocol, which took approximately 45 minutes to complete with participants, included a rigorous digital audio and video recording protocol via HIPAA-compliant Zoom for Healthcare for most cognitive tasks for data integrity and data monitoring purposes. Moreover, multiple quality assurance (QA) and quality control (QC) procedures were performed to maximize the consistency and accuracy of the data. The full training protocol is discussed in greater detail later in the manuscript.

Table 1. List of neuropsychology protocol screening and tests measures, domains, and empirical support for virtual testing in HANDDS-ONT.

Measure	Domain	Videoconference Testing Use
Hearing and Vision	Screening measures	Vision [22]; Hearing [23]
SDMT (Oral): Coding	Attention/Processing Speed	[24-26]
SDMT: Incidental Learning	Memory	--
Judgment of Line Orientation - Short Form (JLO)	Visual Perception	[27]
Test of Premorbid Function (TOPF)	Estimated Intelligence Quotient (IQ)	[8,28]
Rey Auditory Verbal Learning Test (RAVLT)	Memory	[25,29,30]
DKEFS: Color-Word Interference (CWI)	Attention/Processing Speed/Executive Function	[31]

DKEFS: Verbal Fluency (VF)	Language/Executive Function	[26,32]
Questionnaire: Instrumental Activities of Daily Living (iADLs)*	Independence in iADLs	--

*Participant self-reported.

Abbreviations: DKEFS = Delis-Kaplan Executive Function System [33]; JLO = Judgment of Line Orientation [34]; RAVLT = Rey Auditory Verbal Learning Test [35]; SDMT = Symbol-Digit Modalities Test [36]; VF = Verbal Fluency [33]; iADLs = Instrumental Activities of Daily Living [37].

Data Fidelity Approach in a Remote Virtual Neuropsychology Protocol: Quality Assurance (QA) and Quality Control (QC)

Using the QA/QC approach and recommendations from ONDRI 1.0 [12,19] as a framework, we developed and implemented multiple processes and procedures for the remote videoconference-administered neuropsychology protocol used in HANDDS-ONT (see Figure 1 for an overview). These included a variety of data fidelity best practices that were embedded across aspects of the research lifecycle to address different types of issues. QA processes serve to prevent issues before they occur (e.g., training, designing user-friendly data capture instruments) while QC processes serve to identify and address issues when they arise.

The neuropsychology protocol included (1) well-designed, standard paper and/or electronic data capture approaches (including digital audio/video recording of verbal responses whenever possible, and use of auto-calculation features); (2) development and maintenance of standard manuals of operating procedures for administration, scoring, and data entry; (3) a comprehensive standard training protocol; (4) central monitoring of the data which was (5) concurrent with data collection; and (6) a study-specific data collection cleaning and curation process to detect inaccuracies and anomalies in the data.

[Figure 1]

Quality Assurance

The purpose of the QA protocol was to increase the rigor and reproducibility of the data through standardizing and operationalizing data collection (developing administration, scoring, and data entry protocols), reporting, and storage procedures. Consistent with ONDRI 1.0, the aims at all stages of the QA protocol were to minimize errors, to optimize data accuracy, and to maximize consistency and accuracy across research coordinators. The tools and documentation that enabled quality assurance are described below:

Case Report Forms (CRFs) and Electronic Case Report Forms (eCRFs)

Data were collected using paper case report forms (CRFs) and were subsequently entered into a centralized online database system Brain-CODE [38] using the electronic data capture tool REDCap (Research Electronic Data Capture) [39]. Given this workflow, standardizing and harmonizing the design of the paper CRFs with the electronic CRFs (eCRFs) was imperative to minimize

transcription errors. To facilitate data entry, the data capture design also included the auto-calculation of summary scores, contrast scores, creation of summary sheets, minimizing the use of free text fields, and provision of warning messages for out-of-range values and missing data. Explanations for missing data (e.g., administrative error, verbal refusal, physical impairment, technical error, cognitive/behavioural impairment) were inputted into the applicable fields in REDCap. All tasks incorporated a standard set of fields including participant ID, date completed, and test version used. Individual item scores for most measures were included.

The source documents for the majority of neuropsychology protocol tasks (i.e., digital audio and video recordings) were uploaded to REDCap for monitoring purposes. Reporting (Data Reports) and data quality (Data Resolution Workflow) features in REDCap were used by the Neuropsychology lead (K.F.) once data collection commenced. Finally, an electronic CRF to capture neuropsychology-specific protocol deviations and data collection issues (e.g., equipment issues) and an electronic excel record for neuropsychology measure-specific monitoring for each participant was used so that the quality of data also could be monitored and tracked by the Neuropsychology lead. All electronic CRFs were beta tested and verified by the Neuropsychology lead (K.F.) and a Neuroinformatics and Biostatistics team member (B.T.) prior to any data entry.

Study Manuals

Consistent with ONDRI 1.0, three study manuals were provided to research coordinators prior to data collection including test administration, scoring, and data entry manuals. These documents provided standard test administration, scoring, recording, and data entry procedures, as well as information on study workflow, communication strategies, limits of confidentiality, calculating years of education, and procedures for incidental findings. Manuals were updated as needed in response to research coordinator queries and feedback. Proactively, the Neuropsychology Working Group developed a list of additional virtual environment-related guidelines for the research coordinators to follow prior to the administration of the neuropsychology protocol. For example, all participants were asked if they could see the full screen of the examiner when it was shared, all participants were asked to confirm they could see and hear the examiner (and also completed audibility and vision screenings prior to administering the neuropsychological tasks), and all participants were invited to use their mouse/finger to follow along for certain tasks (e.g., symbol-digit modalities test).

Hybrid Model Training Protocol Roadmap

A mandatory training and certification program for study coordinators responsible for administering the neuropsychology protocol was delivered using a hybrid (i.e., in person and virtual) training model (Figure 2). The training and certification program included a minimum of 14 hours of in-person training with the Neuropsychology lead (K.F.) as well as demonstration of proficiency in virtual test administration and scoring. As recommended by Sadler and colleagues [40], the training integrated didactic methods with experiential, practical, and applied learning components. Aspects of the training were repeated when proficiency was not demonstrated (e.g., data scoring errors on a given task for at least four consecutive participants). All coordinators were observed live, remotely by the Neuropsychology lead (K.F.) for their first five administrations to participants (via HIPAA-compliant Zoom for Healthcare). Neuropsychology research coordinators' administration, scoring, and data entry accuracy then were assessed regularly (i.e., weekly to bi-weekly) during the data collection phase of the project remotely via data monitoring, discussed in detail in the QC section.

A number of ongoing virtual activities and meetings were conducted with the neuropsychology research coordinators, including population-specific communication training and educational workshops. Virtual ad hoc “booster” training sessions (~1 hour) were also completed approximately every 2 months. These sessions included an overview of test administration, scoring, and data entry procedures, as well as common queries and errors made by coordinators that were either flagged during iterative data monitoring by the Neuropsychology lead or during the iterative cleaning and curation pipeline. Virtual, twice monthly meetings were held with the Neuropsychology Working Group (NWG; A.T., A.R., P.M., J.B.O) to review outstanding queries.

Over the course of the HANDDS-ONT study (18 months of data collection), four research coordinators completed the neuropsychology training protocol. Only one research coordinator required one additional live observation session of the neuropsychology protocol. Two research coordinators administered 70% ($n = 104$) of the neuropsychology protocol, with the third research coordinator administering 23% ($n = 34$) and the fourth research coordinator administering 7% ($n = 10$).

[Figure 2]

Standards and Data Conventions Policy

Standards and conventions policies were established for HANDDS-ONT based on the ONDRI 1.0 protocol [12], which were influenced by existing projects (e.g., Alzheimer’s Disease Neuroimaging Initiative), FAIR principles [41], and data management practices for statistical analyses [42-44]. In general, the standards and data convention policies involve developing consistency in file formatting (e.g., date inputs, variable naming conventions, non-numeric missing data codes) and data packaging (e.g., creation of a data dictionary, readme file, methods documents, and missing data descriptions). Additional information regarding standard and convention policies is included in McLaughlin and colleagues [12].

Quality Control

QC procedures were established to identify and to correct data anomalies or inaccuracies both *concurrently* with data collection (see recommendations from Keefe and Harvey [45]) and during the *post-collection* phase of the study. Notably, QC procedures occurred in phases and during all stages of the study. A roadmap for data collection through to data release workflow for the remote, videoconference-administered neuropsychology protocol is depicted in Figure 3. Once the cohort of data collected at the time were monitored, that cohort of data moved through to the cleaning and curation pipeline, standards check, and subsequently to the outliers analysis. Data were monitored on a weekly or biweekly basis, and the cleaning and curation pipeline, standards check, and outliers analysis were run approximately every 3 months.

[Figure 3]

Data Entry (and if needed, Data Entry Verification)

Based on learnings from ONDRI 1.0 [12], steps were taken to improve the accuracy and efficiencies of the quality control procedures. In the current protocol, the research coordinator administering the

test battery was responsible for *both* scoring and entering the data into REDCap. Neuropsychology research coordinators used the electronic query system in REDCap to flag the Neuropsychology lead of ambiguous scoring situations and they also used an ongoing excel spreadsheet to track all scoring and administration queries. All queries were extracted via an excel spreadsheet and were reviewed and verified by the Neuropsychology lead and/or Neuropsychology Working Group at twice monthly virtual meetings. Our protocol allowed for all errors to be addressed within a 2 to 3 week period, and in an iterative fashion.

Data Monitoring

Data were monitored centrally by the Neuropsychology lead (K.F.), in an iterative fashion, as the data were collected (Figure 3).

Table 2 outlines the monitoring procedure used in HANDDS-ONT. The monitoring procedure was developed based on: 1) the number of queries generated for each task in the ONDRI 1.0 protocol, as tasks that were more susceptible to error were monitored more frequently by the Neuropsychology lead and 2) the recommendations of McLaughlin and colleagues [12], including:

- a) Monitoring 100% of the participants' data whenever possible;
- b) Applying ongoing monitoring *concurrently* with data collection;
- c) Targeting tasks known to be vulnerable to administration, scoring, and/or data capture errors based on *early* monitoring;
- d) Providing participant cohort-specific training (e.g., training on communication with individuals with a diagnosis of ALS);
- and e) Implementing a pipeline that:
 - i) identifies missing data, scoring issues, out of range values, erroneous reporting, protocol deviations, and general anomalies within the dataset;
 - ii) takes raw scores entered into the database, and algorithmically generates standard scores (z scores, ss, T-scores, percentiles) based on normative data, and;
 - iii) generates data files that format the data according to study-specific standards and reporting practices.

As a crucial component to data monitoring, the automated pipeline is discussed in detail below.

Table 2. Virtual neuropsychology protocol data monitoring plan.

	Data Monitoring Plan	
	Participants 1-5: Live observation, joint scoring, REDCap monitoring for all NPSY Research Coordinators	Participant 6 and beyond: REDCap monitoring for all NPSY Research Coordinators
Neuropsychology measure		
DKEFS Verbal Fluency	100%	100%

Test of Premorbid Function (TOPF)	100%	100%
Rey Auditory Verbal Learning Test (RAVLT)	100%	100%
DKEFS Color-Word Interference (CWI)	100%	80%
Symbol Digit Modalities Test (SDMT)	100%	80%
Judgment of Line Orientation (JLO)	100%	80%

Abbreviations: DKEFS = Delis-Kaplan Executive Function System [33].

Querying protocol

Electronic data queries were generated in REDCap using the Data Resolution Workflow function. This function provides a formal process for documenting changes, missing data, or corrections to the data. In this protocol, the Neuropsychology lead (K.F.) opened a query in a data field and assigned it to the Neuropsychology lead. Each query included information on the suspected error and requested that the Neuropsychology lead (alone or with the Neuropsychology Working Group, who met on a bimonthly basis) verify the entered value using the study protocol and make the appropriate corrections. Once the data were confirmed or corrected by the Neuropsychology lead, the query was closed, and the issue was considered resolved in the eCRFs. The neuropsychology research coordinators corrected the data and made the appropriate documentation on the paper CRFs.

All CRF and eCRF questions and changes were logged in a comprehensive study log using a live, up-to-date, shared google excel spreadsheet which the Neuropsychology lead developed and maintained as the data were being monitored and verified. The spreadsheet contained all queries and all responses to the queries (with decision justifications, dates of decisions, and a list of Neuropsychology Working Group members participating in each decision). The majority of queries and other question entries consisted of those that were associated with ambiguous scoring situations, unusual participant responses, as well as standardized recommendations for HANDDDS-ONT.

Data Cleaning and Curation Pipeline

To supplement data monitoring activities performed by the Neuropsychology lead, a semi-automated data cleaning pipeline was developed in ONDRI 1.0. The data cleaning pipeline facilitates data monitoring for large batches of data that are too time-consuming and error-prone for the Neuropsychology lead to complete in a timely manner. In response to the need to monitor data in a consistent and rapid manner, the Neuropsychology Working Group and members of the Neuroinformatics & Biostatistics Platform working group (K.S., M.B.) identified relevant data quality issues such as missing data, scoring errors, out of range values, erroneous reporting, protocol deviations, and general anomalies within the data set. Then, reproducible R code was written and organized into a pipeline that processed the data under the following steps: 1) data ingestion from REDCap, 2) data processing and computing scores, 3) identifying errors in the data that warranted further investigation, 4) reporting of errors in a human-readable format, and 5) creation of a dataset in a standardized format. The pipeline served as a critical post-collection quality control procedure. In HANDDDS-ONT, this semi-automated cleaning and curation pipeline was optimized further based

on recommendations of McLaughlin and colleagues [12] by 1) incorporating more automated scores (e.g., contrast scores), 2) incorporating the use of additional missing codes in REDCap and 3) applying iterative monitoring and cleaning/curation concurrently with data collection.

After the data were scored, entered, and verified in REDCap (if applicable, in advance of monitoring), and all of the targeted files were monitored, the data entered the pipeline in an iterative fashion (Figure 3). The pipeline's results were verified manually against source documents (paper CRFs and video recordings) by the Neuropsychology lead, and the Neuropsychology lead resolved all queries on REDCap where an audit trail was recorded. As changing one value or score in the data could result in down-stream effects (e.g., correcting an item score will affect the total score and may affect the associated standard score), an iterative process was established so that the automated pipeline was rerun after the data were corrected and verified until no new errors were detected and the data were considered static.

Outlier Analyses

In ONDRI 1.0, the Neuroinformatics & Biostatistics team developed a study-wide QC protocol to identify and report on multivariate outliers. The bases of this protocol are resampling and principal components analysis-like methods to identify multivariate outliers, including the minimum covariance determinant [46] and the CorrMax transformation [47]. A general summary of the outlier analysis procedure can be found in Sunderland and colleagues [19] and McLaughlin and colleagues [12]. The outlier analysis procedure was similarly implemented in HANDDS-ONT, with the exception that the procedure was performed on all neuropsychology data collected together (i.e., across all cohorts). Further information can be found in Supplementary Material B.

Supplementary Material B: Outlier Analysis

In the outlier analysis, multivariate outlier observations (participants) are identified through the minimum covariance determinant and bootstrap resampling. Outlying individuals are those outside of a pre-specified cutoff for a bootstrap distribution of Mahalanobis distances. A CorrMax transformation is then used to identify the sources (variables) of outlying-ness for each observation. Both the outlying observations (participants) and their sources (variables) are then reported back to the Neuropsychology lead. All outlying data values identified are then verified against the source document (paper CRF and/or video recording) by the Neuropsychology lead to determine whether the outlier was an error or an anomaly (i.e., an unusual, but true data point). In keeping with GCP guidelines (ICH-E6, 2014), whenever evidence was uncovered that an error occurred, the Neuropsychology lead was queried to correct the data and make the appropriate documentation (see querying protocol). Data points verified as anomalies (unusual but true) were kept and released in the final data sets.

The foundation of the outlier protocol is described in Sunderland and colleagues [19], with expansions of the core techniques reported elsewhere [48]. Software was developed for the protocol, and a beta version has been made publicly available with examples (<https://github.com/derekbeaton/ours>) [49].

Participants

[Figure 4]

Figure 4 includes information regarding the eligible number of participants to complete the neuropsychology protocol after REB-approval was granted and the neuropsychology protocol was implemented. Of the individuals who were eligible to complete the NPSY protocol from the implementation date and who were eligible, the completion rate was 82.3%.

Measuring Data Fidelity in a Remote, Videoconference-Administered Neuropsychology Protocol

Efficacy of Approach

To investigate the effectiveness of our data fidelity approach in our remote videoconference-administered neuropsychology protocol, we examined the number of errors detected in the data (excluding questionnaires) during (1) the monitoring phase of the study, (2) as a direct result of the cleaning and curation pipeline, and (3) as a consequence of the outlier analyses. As an index of error, we used the number of electronic queries opened during each of the three stages. We excluded any data changes in the questionnaires and administrative fields (e.g., dates, neuropsychology protocol issues documentation).

Data Quality

As an index of data quality, we investigated (1) the number of missing data points across cohorts and tasks and the reason(s) for the missing data and (2) the number of missing data and other data issues related to the virtual testing environment.

Results

Demographics

Demographics information for the 148 participants who participated in the neuropsychology protocol can be found in Table 3. All diagnoses in HANDDS-ONT were self-reported, with 66% of diagnoses verified by family physicians/referring physicians. A subset of participants reported multiple diagnoses [e.g., AD and MCI (n=1), or MCI and CVD (n=2)]. Three participants who completed the neuropsychology battery self-reported diagnoses that conflicted with the clinician report or omitted pertinent medical history in their self-report (e.g., reporting MCI when they had a history of stroke); for these participants, the medical history and clinician-report were prioritized for categorization. For the purposes of the manuscript, AD and AD/MCI are grouped as “MCI”, and MCI and CVD are grouped as “CVD.”

Table 3. Self-reported demographics information of HANDDS-ONT neuropsychology protocol

participants.

	Cohort						
	MCI	CVD	ALS	FT D	PD**	Community dwelling	Total****
n	24	21	8	2	45	48	148 (n = 146 with reported statistics)
Age median years (range)	74 (45-91)	66 (37-80)	61 (32-74)	*	66 (51-82)	67 (30-91)	67 (30-91)
Sex Female (%) male (%)	13 (54.2) 11 (45.8)	16 (76.2) 5 (23.8)	2 (25.0) 6 (75.0)	*	19 (42.2) 25 (55.6)	40 (83.3) 8 (16.7)	90 (61.6) 55 (37.7)
Handedness (%) right (%) left (%) amb (%)	20 (83.3) 3 (12.5) 1 (4.2)	18 (85.7) 2 (9.5) 1 (4.8)	8 (100.0) -- --	*	41 (91.1) 3 (6.7) --	44 (91.6) 3 (6.3) 1 (2.1)	131 (89.7) 11 (7.5) 3 (2.1)
Ethnicity (%) European Other Don't know Prefer no answer	18 (75.0) 4 (16.7) 1 (4.2) 1 (4.2)	17 (81.0) 3 (14.3) 1 (4.8) -	6 (75.0) 2 (25.0) - -	*	36 (80.0) 7 (15.5) - 1 (2.2)	35 (72.9) 10 (20.8) 1 (2.1) 2 (4.2)	112 (76.7) 26 (17.8) 3 (2.1) 4 (2.7)
Education (yrs)*** Mean (S.D.)	15.2 (2.4)	15.1 (2.7)	14.1 (1.9)	*	15.5 (2.4)	14.8 (2.6)	15.1 (2.5)
Education description (%) less than high school GED/equiv. High school Some college, no degree Associate degree Bachelor's degree Master's	2 (8.3) 1 (4.2) 2 (8.3) 1 (4.2) 3 (12.5) 8 (33.3) 4 (16.7)	2 (9.6) -- 2 (9.5) 2 (9.5) 5 (23.8) 4 (19.0) 3 (14.3)	-- -- 1 (12.5) 2 (25.0) 2 (25.0) 2 (25.0) --	*	-- -- 3 (6.7) 5 (11.1) 6 (13.4) 18 (40.0) 8 (16.7)	-- -- 5 (10.3) 8 (16.7) 12 (25.0) 12 (25.0) 8 (16.7)	4 (2.7) 1 (0.7) 13 (8.9) 18 (12.3) 28 (19.2) 44 (30.1) 21 (14.4)

degree					6 (13.3)		
Doctoral	--	1 (4.8)	--			--	3 (2.1)
degree					2 (4.4)		
Professional	3 (12.5)	2 (9.5)	1 (12.5)			3 (6.3)	13 (8.9)
school degree					4 (8.9)		
Marital status (%)				*			
Married/domestic partner	17 (70.8)	10 (47.6)	6 (75.0)		36 (80.0)	28 (58.3)	97 (66.4)
Separated/divorced	3 (12.5)	5 (23.8)	1 (12.5)			13 (27.1)	26 (17.8)
Widowed	3 (12.5)	3 (14.3)	--		4 (8.9)	4 (8.3)	12 (8.2)
Never married	1 (4.2)	3 (14.3)	1 (12.5)		2 (4.4)	3 (6.3)	10 (6.8)
					2 (4.4)		
Employment				*			
Retired (%)	21 (87.5)	14 (66.7)	5 (62.5)		26 (57.8)	25 (52.1)	91 (62.3)
Income				*			
<50k (%)	6 (25.0)	6 (28.6)	--		1 (2.2)	10 (20.8)	23 (15.8)
50k-100k (%)	5 (20.8)	6 (28.6)	2 (25.0)		11 (24.5)	15 (31.3)	39 (26.7)
100k+ (%)	7 (29.2)	6 (28.6)	5 (62.5)			19 (39.6)	61 (41.7)
Prefer no answer	6 (25.0)	2 (9.5)	1 (12.5)		24 (53.3)	4 (8.3)	21 (14.4)
Don't know	--	1 (4.8)	--		8 (17.8)	--	1 (0.7)
					--		

S.D. = standard deviation.

*Information not reported given <5 participants in the cell.

** n = 1 did not complete the demographics information form.

*** = Education years based on Strauss and colleagues (2006) guidelines.

**** = Demographics information for participants with FTD (n=2) omitted from totals column.

Table 4. Descriptive statistics of neuropsychology performance, by cohort.

	Cohort					
	MCI	CVD	ALS	FTD*	PD	Community dwelling
Neuropsychology Test						
VF						
Letter Fluency	37.4 ^a (12.6)	44.1 (15.2)	41.9 ^a (7.7)		44.7 ^d (13.8)	42.7 (12.1)
Category Fluency	33.0 (7.6)	39.6 (11.0)	41.4 (15.3)		40.2 ^a (9.4)	41.8 (8.9)
RAVLT						
Immediate Recall	36.0 (11.7)	48.6 (12.0)	44.4 (10.1)		42.2 ^d (10.9)	47.6 ^a (10.8)
Long-delay Recall	5.1 (3.6)	9.0 ^a (2.6)	8.8 (3.2)		7.7 ^d (3.1)	9.3 ^a (3.6)
Recognition Discrimination	4.7 (6.3)	10.0 ^d (2.9)	10.3 (2.7)		8.3 ^d (4.4)	10.2 ^a (4.4)
TOPF						
TOPF score	47.3	48.1	47.0		47.3 ^a	47.6 ^a (11.7)

	(10.2)	(11.6)	(9.7)		(11.7)	
DKEFS CWI						
CWI Color Naming	40.4 ^a (8.0)	36.1 ^a (8.6)	36.4 (9.6)		35.6 ^a (7.6)	34.8 ^d (7.4)
CWI Word Reading	28.4 ^a (5.5)	27.7 ^d (7.8)	28.1 (11.1)		26.4 ^a (4.9)	26.1 (5.8)
CWI Inhibition	79.8 ^b (19.1)	66.5 ^a (18.8)	60.7 (13.2)		62.6 ^a (22.8)	60.9 (14.6)
CWI Inhibition/Switching	93.0 ^c (31.0)	65.4 ^b (11.0)	65.8 (20.1)		77.2 ^a (34.5)	72.2 (24.4)
SDMT						
SDMT score	29.8 ^d (8.2)	41.5 ^b (10.9)	43.9 ^a (12.8)		39.4 ^d (12.1)	46.4 ^d (12.0)
JLO						
JLO score	22.1 (5.6)	23.7 (6.9)	24.8 (4.7)		25.8 ^a (4.6)	23.9 (4.6)

Note. Data are presented as means (and standard deviations).

^an = 1 categorized as missing data.

^bn = 3 categorized as missing data.

^cn = 4 categorized as missing data.

^dn = 2 categorized as missing data.

Abbreviations: DKEFS = Delis-Kaplan Executive Function System [33]; JLO = Judgment of Line Orientation [34]; TOPF = Test of Premorbid Function [50]; CWI = Color-Word Interference [33]; VF = Verbal Fluency [33]; RAVLT = Rey Auditory Verbal Learning Test [35]; SDMT = Symbol-Digit Modalities Test [36]. MCI = amnesic mild cognitive impairment; CVD = cerebrovascular disease, with or without cognitive impairment; ALS = amyotrophic lateral sclerosis; FTD = frontotemporal lobar degeneration; PD = Parkinson's disease.

*2 participants only, data not reported.

Efficacy of Approach

Monitoring

A total of 715 queries were opened during the data monitoring phase (see Table 5). Twenty (n = 20; community dwelling n = 9; CVD = 3, MCI = 2, ALS = 1, PD = 5) remote videoconference-administered observations were excluded as files were co-scored and reviewed live with the Neuropsychology lead. The average number of queries per variable (for all overlapping measures [5 of 6 measures, as TOPF was not used in ONDRI 1.0]) were comparable to McLaughlin and colleagues' [12] in-person neuropsychological assessment protocol, supporting the rigor of the roadmap of our hybrid training model and remote videoconference neuropsychology protocol.

Table 5. Queries opened during the data monitoring phase across cohorts for each assessment (excluding questionnaires and spelling errors)

	Cohort		Variables
--	--------	--	-----------

	MC I	CV D	AL S	FT D	PD	Communit y Dwelling	Total	# variables per participant	Avg # queries per variable
n	22	18	7	2	40	39	128		
Neuropsychology Test									
VF	49	34	13	6	78	84	264	92	2.9
RAVLT	25	39	8	1	60	80	213	201	1.1
TOPF	25	24	14	2	56	40	161	12	13.4
CWI	15	4	1	1	5	16	42	63	0.7
SDMT	7	7	0	0	8	9	31	22	1.4
JLO	0	1	0	0	0	3	4	15	0.3
Total # of queries	121	109	36	10	207	232	715		
Avg # queries per participant	5.50	6.06	5.14	5.00	5.18	5.95	5.59		

Note. Queries excluded spelling errors and non-nominal data changes (e.g., adding a question mark to a stated word on the RAVLT).

Abbreviations: VF = Verbal Fluency; RAVLT = Rey Auditory Verbal Learning Test; TOPF = Test of Premorbid Function; CWI = Color-Word Interference (CWI); SDMT = Symbol Digit Modalities Test; JLO = Judgment of Line Orientation. MCI = amnesic mild cognitive impairment; CVD = cerebrovascular disease, with or without cognitive impairment; ALS = amyotrophic lateral sclerosis; FTD = frontotemporal lobar degeneration; PD = Parkinson's disease.

Cleaning and Curation Pipeline

A total of 45 queries were opened during the cleaning and curation pipeline phase (see Table 6). The average number of queries per variable (for all overlapping measures [5 of 6 measures, as TOPF was not utilized in ONDRI 1.0]) were comparable to McLaughlin and colleagues' [12] in-person neuropsychological assessment protocol, again supporting the rigor of the roadmap of our hybrid training model and remote videoconference-administered neuropsychology protocol given that the numbers were generally similar and that there were also few queries generated.

Table 6. Queries opened during the curation and cleaning pipeline phase across cohorts for each assessment (excluding questionnaires and spelling errors).

	Cohort							Variables	
	MCI	CV D	AL S	FT D	PD	Communit y Dwelling	Total	# variables per	Avg # queries per

								participan t	variabl e
n	24	21	8	2	45	48	148		
Neuropsychology Test									
VF	1	0	2	0	1	0	4	92	0.04
RAVLT	0	1	0	0	14	10	25	201	0.12
TOPF	0	0	0	0	0	0	0	12	0.00
CWI	0	0	0	0	4	2	6	63	0.10
SDMT	1	0	0	0	2	7	10	22	0.45
JLO	0	0	0	0	0	0	0	15	0.00
Total # of queries	2	1	2	0	21	19	45	--	--
Avg # queries per participant	0.08	0.05	0.25	0.00	0.47	0.40	0.30	--	--

Note. Queries excluded spelling errors and non-nominal data changes (i.e., adding a question mark to a stated word on the RAVLT).

Abbreviations: VF = Verbal Fluency; RAVLT = Rey Auditory Verbal Learning Test; TOPF = Test of Premorbid Function; CWI = Color-Word Interference; SDMT = Symbol Digit Modalities Test; JLO = Judgment of Line Orientation. MCI = amnesic mild cognitive impairment; CVD = cerebrovascular disease, with or without cognitive impairment; ALS = amyotrophic lateral sclerosis; FTD = frontotemporal lobar degeneration; PD = Parkinson's disease.

Outlier Analysis

A multivariate outlier detection process was performed including all participants across the six cohorts. Each identified participant-variable pair with a maximum proportion greater than 24.99 was verified in REDCap against the source data (i.e., paper CRF and audio/video file). In total, 36 participants were identified as having either 1 or 2 datapoints as outliers, corresponding to 37 outlying data points. All flagged outliers for neuropsychological measures were reviewed. All were deemed as anomalies in the data, and there were no errors detected (Table 7).

Table 7. Participants with flagged outlying data points across cohorts, by neuropsychology test.

	Cohort	
--	--------	--

	MCI	CVD	ALS	FTD	PD	Community Dwelling	Total
n	24	21	8	2	45	48	148
Number of Participants with Flagged Outlying Data points (%)	7 (29%)	7 (33%)	2 (25%)	0 (0%)	11 (24%)	9 (19%)	36 (24%)
Neuropsychology Task							
VF	0	1	1	0	3	1	6
RAVLT	2	0	0	0	0	1	3
TOPF	0	0	0	0	0	0	0
CWI	4	4	1	0	5	6	20
SDMT	0	0	0	0	2	2	4
JLO	1	2	0	0	1	0	4

Note. Data are presented as the number of outliers flagged.

Abbreviations: VF = Verbal Fluency; RAVLT = Rey Auditory Verbal Learning Test; TOPF = Test of Premorbid Function; CWI = Color-Word Interference; SDMT = Symbol Digit Modalities Test; JLO = Judgment of Line Orientation. MCI = amnesic mild cognitive impairment; CVD = cerebrovascular disease, with or without cognitive impairment; ALS = amyotrophic lateral sclerosis; FTD = frontotemporal lobar degeneration; PD = Parkinson's disease.

Data Quality

Missing Data

There was a low percentage of missing data across disease cohorts (0.34%), with 476 of 142,080 fields missing across 148 participants (Table 8). Again, these results are comparable to McLaughlin and colleagues' [12] in-person neuropsychological assessment across disease cohorts, which ranged from 0.8%-3.3% for each cohort. The largest percentage of missing data on a neuropsychology task was the DKEFS Color-Word Interference Task, which was primarily related to participants reaching the time limit to initiate the discontinue criterion for Conditions 3 or 4 or endorsing colour-blindness.

Table 8. Reasons for missing neuropsychology protocol data by cohort (excluding questionnaires).

	Cohort						
	MCI	CVD	ALS	FTD	PD	Community Dwelling	Total
n	24	21	8	2	45	48	148
Missing data type							
Administration error	27	32	3	11	5	28	106
Cognitive/behavioural impairment	45	12	0	0	0	0	57
Physical impairment	34	34	0	0	0	0	68
Technical error	10	9	0	0	34	0	53
Verbal refusal	0	0	0	0	0	0	0
Other	0	0	12	0	151	29	192
Number of potential missing variables per participant (i.e., the maximum number of potential missing variables that one participant could have on each task, if no data were recorded.)	160	160	160	160	160	160	960
Total # of missing values	116	87	15	11	190	57	476
% Missing	3.02	2.59	1.17	3.44	2.64	0.74	0.34

Note. administration error = coordinator made an error that invalidated the test and/or some aspect of the data; cognitive/behavioral impairment = participant was unable to understand the task requirements and/or perform the task as intended because of a cognitive and/or behavioral impairment; physical impairment = participant was unable to complete the task because of a physical deficit or limitation (e.g., reduced mobility in arms or hands, color blindness, etc.); technical error = technical or equipment issue occurred (e.g., laptop crashed midway through a task); verbal refusal =

participant refused to complete the measure after being read the task instructions; other = another issue resulted in missing or invalid data (e.g., test disrupted, test session discontinued early, etc.). % Missing = (Total number of missing values / [Number of potential missing variables per participant × Number of participants]) × 100.

Abbreviations: MCI = mild cognitive impairment; CVD = cerebrovascular disease, with or without cognitive impairment; ALS = amyotrophic lateral sclerosis; FTD = frontotemporal lobar degeneration; PD = Parkinson's disease.

Virtual Environment-Related Errors and Issues

A summary of missing data specifically related to administration in the virtual environment (e.g., audio, internet connection, Zoom platform) are included in Table 9, and include issues that resulted in and did not result in missing data. Only a few instances of both are described, with missing data related to the virtual environment involving 8 participants.

Table 9. Missing data and issues in neuropsychology protocol related to the virtual environment.

	Virtual testing issues resulting in missing data	Virtual testing issues not resulting in missing data
Virtual Testing Issue		
Audio cut out	1 participant, 1 task	N/A
Internet connection cut out	3 participants, 1 task each	11 participants, 1-2 task(s) each
Writing down list off camera (not caught)	1 participant, 1 task	N/A
Zoom	4 participants, 1 task each	4 participants, 1 task each
Total	8 participants (5.4% of participants), 1-2 tasks each	14 participants (9.4% of participants), 1-2 tasks each

Note. Data are presented as number of participants and number of tasks impacted.

Discussion

Remote, videoconference-administered, synchronous, interdisciplinary, observational clinical studies and clinical trials will continue to supply a wealth of data as research participants and scientists embrace the technologies that enable this approach. In the ongoing shift toward both cross-modality collaboration and use of solely virtual (or hybrid) models of data collection, it is critical that data fidelity remains a top priority for researchers. Therefore, the development of research-supported roadmaps are imperative to promote robust quality assurance and quality control procedures for accruing neuropsychology and clinical data. Our manuscript demonstrates that a remote videoconference-administered neuropsychology protocol performed similarly to an in-person neuropsychology protocol as it pertains to quality control indices (i.e., missingness and number of data corrections needed). It also highlights that there were few virtual environment-related factors that impacted data missingness.

Principal Results

In this manuscript, we outlined three primary roadmaps to illustrate our QA/QC procedures for a remote videoconference-administered neuropsychology protocol:

- 1) A schematic overview of the quality assurance (error prevention) and quality control (detection and correction of errors) roadmap;
- 2) A hybrid training roadmap (a component of the quality assurance procedures);
- 3) Data collection to data release workflow roadmap, including protocol administration, iterative data monitoring, cleaning/curation pipeline implementation, standards pipeline implementation, outlier analysis, and data release.

Of note, our roadmaps highlight 1) an *iterative, concurrent* approach to data collection, data monitoring, and data release to minimize recurrent errors, 2) a rigorous hybrid training procedure to ensure accuracy and consistency among coordinators, and 3) a thoughtful implementation of an iterative cleaning and curation pipeline. The cleaning and curation pipeline identifies missing data, scoring issues, out of range values, erroneous reporting, protocol deviations, general anomalies, optimizes time efficiencies via scoring automation, and extracts all of the neuropsychology data from the electronic CRF in REDCap to create a data set using a standard format (merging redundant variables and formatting the data according to these conventions). The semi-automated pipeline effectively minimizes the risk of human error and is an essential tenant of the design of effective data fidelity procedures.

During data monitoring, the TOPF had the highest average number of queries per variable, followed by verbal fluency. During the cleaning and curation pipeline, the RAVLT and the SDMT had the highest average number of queries per variable (which were both less than 1). The average number of queries for the community dwelling participants was also similar to the neurodegenerative/neurologic disease cohorts, supporting the acceptability of these measures across disease cohorts. With respect to missing data, it is noted that the most common reportable cause of missing data was due to administration error; nevertheless, administration errors accounted for <1% of missing data. Again, comparatively, a similar proportion of queries and errors were flagged across the in-person and remote videoconference-administered neuropsychology protocol [see 12]. Moreover, few virtual environment-related errors occurred, with 1-2 tasks impacted for 8 participants resulting in missing data. The approach to fidelity and resulting data demonstrate that researchers can obtain robust data in fully remote participant neuropsychology protocols using the same measures undertaken in the HANDDS-ONT study if they are designed in a rigorous, intentional, and collaborative way.

Limitations

The increasing use of remote videoconference assessments for research and clinical purposes also necessitates a critical reflection of factors influencing the accessibility and representativeness of the sample. First, the accessibility of reliable internet and appropriate devices to participate in a remote videoconference assessment must be considered [9]. Reliable internet access is not uniformly available across different geographical regions and socioeconomic groups, as urban areas and higher-income households typically have greater internet infrastructure compared to rural or lower-income areas. This disparity may lead to an underrepresentation of certain demographics (e.g., socioeconomic status) in videoconference-based studies. Second, there may be a participant bias as individuals who opt to participate in these studies are often more likely to have the digital literacy skills to navigate videoconference and other virtual platforms. Understandably, their ability to participate necessitates some familiarity with (and access to) technology. These factors limit the generalizability of our findings. To mitigate these issues, researchers may consider using strategies to

enhance accessibility and to reduce participant bias by collaborating with community organizations to facilitate access to technology and/or offering technical support and education to participants with lower proficiency with digital technology.

Conclusions

We acknowledge that our approach to neuropsychology assessment fidelity is rigorous and resource-intensive. However, we ask that research teams consider the returns. For example, reducing the occurrence of missing data and/or erroneous data will almost certainly facilitate high quality analyses. The neuropsychology protocol data used for this manuscript are scheduled to be shared openly with the scientific community and the data fidelity processes described here will enable the data to be used by external groups under an open science paradigm. Indeed, the work done to ensure and document high fidelity allows researchers to focus more on the scientific interpretation of the results and less time cleaning the data.

Importantly, our roadmap tools were designed collaboratively by a team of experts across neuroscience, neuroinformatics, statistics, cognition, language science, and neuropsychology fields. The effectiveness of the designed roadmaps used in the HANDDS-ONT study highlight the importance of establishing cross-disciplinary partnerships to increase and maximize the inclusivity, accessibility, and data fidelity of participants involved in fully remote protocols. More importantly, the results show that rigorous training, on-going, on-line, as well as post-hoc monitoring, errors analyses, and feedback loops back into the data collection protocol optimized the fidelity of data collection and validity of the neuropsychology dataset of the HANDDS-ONT study. These roadmaps suggest that in the current paradigm shift, it is both feasible and reasonable to consider remote videoconference-administered neuropsychological testing with good support.

Acknowledgments

This work was completed on behalf of the Health in Aging, Neurodegenerative Diseases and Dementias in Ontario study (HANDDS-ONT), conducted as part of the Ontario Neurodegenerative Disease Research Initiative (ONDRI). The authors and investigators would like to acknowledge the HANDDS-ONT's project management team, Natalie Rashkovan, Ashley Wilcox, Ruth Kruger, and Heather Hink for their input and administrative support of this project. The authors would also like to thank the HANDDS-ONT participants. In addition, the authors would like to acknowledge the ONDRI Founding Authors: Robert Bartha, Sandra E. Black, Michael Borrie, Dale Corbett, Elizabeth Finger, Morris Freedman, Barry Greenberg, David A. Grimes, Robert A. Hegele, Chris Hudson, Anthony E. Lang, Mario Masellis, William E. McIlroy, Paula M. McLaughlin, Manuel Montero-Odasso, David G. Munoz, Douglas P. Munoz, J. B. Orange, Michael J. Strong, Stephen C. Strother, Richard H. Swartz, Sean Symons, Maria Carmela Tartaglia, Angela K. Troyer, and Lorne Zinman.

This research was conducted with the support of the Ontario Brain Institute, an independent non-profit corporation, funded partially by the Ontario Government. The opinions, results, and conclusions are those of the authors and no endorsement by the Ontario Brain Institute is intended or should be inferred.

Conflicts of Interest

PMM received partial salary support from the Ontario Brain Institute. ACR receives salary funding from the Canada Research Chairs program. RHS is on the advisory board for Roche and has stock

ownership in FollowMD Inc.

Abbreviations

ALS = Amyotrophic lateral sclerosis

CVD = Cerebrovascular disease

CWI = Color-Word Interference

DKEFS = Delis-Kaplan Executive Function System

FTD = Frontotemporal lobar degeneration

HANDDDS-ONT = Health in Aging, Neurodegenerative Diseases and Dementias in Ontario

JLO = Judgment of Line Orientation

MCI = Mild cognitive impairment

NWG = Neuropsychology Working Group

ONDRI = Ontario Neurodegenerative Disease Research Initiative

PD = Parkinson's disease

RAVLT = Rey Auditory Verbal Learning Test

ReMiNDD = Remote Monitoring in Neurodegenerative Disease

SDMT = Symbol-Digit Modalities Test

VF = Verbal Fluency

Data Availability Statement

ONDRI data—including but not limited to the data in this article—will be accessible through the Ontario Brain Institute (<https://braininstitute.ca/>) to qualified researchers.

References

1. Geddes MR, O'Connell ME, Fisk JD, Gauthier S, Camicioli R, Ismail Z; Alzheimer Society of Canada Task Force on Dementia Care Best Practices for COVID-19. Remote cognitive and behavioral assessment: Report of the Alzheimer Society of Canada Task Force on dementia care best practices for COVID-19. *Alzheimers Dement (Amst)*. 2020;12(1):e12111. doi:10.1002/dad2.12111.
2. Ontario Psychological Association. Remote tele-assessment guidelines: Recommendations for neuropsychological assessments during the COVID-19 pandemic. Ontario Psychological Association. 2020;1-8. Available from: <https://www.psych.on.ca/getmedia/8d9c5cf3-e4f1-470b-8a51-3d603dbbc219/OPACAPDA-RemoteTele-Assessment-V9-revised.pdf>.
3. American Psychological Association. Telepsychology revisions. *Am Psychol*; 2024. Available from: <https://www.apa.org/about/policy/telepsychology-revisions>.
4. Bilder RM, Postal KS, Barisa M, et al. Inter organizational practice committee recommendations/guidance for teleneuropsychology in response to the COVID-19 pandemic. *Arch Clin Neuropsychol*. 2020;35(6):647-659. doi:10.1093/arclin/acia046. PMID:32666093.
5. Brearly TW, Liu H, Smith Z, et al. A review of remote neuropsychological assessment practices. *J Neuropsychol*. 2017;31(3):125-134. doi:10.1002/jnp.12345.
6. Grosch J, Simpson S, Ford J, et al. Assessment of cognitive function in telehealth

- interventions for older adults. *J Telemed Telecare*. 2011;17(3):145-151. doi:10.1258/jtt.2011.110507.
7. Brown T, Zakzanis KK. A review of the reliability of remote neuropsychological assessment. *Appl Neuropsychol Adult*. 2023;1-7. doi:10.1080/23279095.2023.2279208.
 8. Chapman JE, Gardner B, Ponsford J, Cadilhac DA, Stolwyk RJ. Comparing performance across in-person and videoconference-based administrations of common neuropsychological measures in community-based survivors of stroke. *J Int Neuropsychol Soc*. 2020;26(11):1089-1097. doi:10.1017/S1355617720000856.
 9. Støme LN, Wilhelmsen CR, Kværner KJ. Enabling guidelines for the adoption of eHealth solutions: scoping review. *JMIR Form Res*. 2021;5(4):e21357. doi:10.2196/21357.
 10. Marra DE, Hamlet KM, Bauer RM, Bowers D. Validity of teleneuropsychology for older adults in response to COVID-19: a systematic and critical review. *Appl Neuropsychol Adult*. 2020;27(6):1411-1452. doi:10.1080/13854046.2020.1769192.
 11. Parks AC, Davis J, Spreser CD, Stroescu I, Ecklund-Johnson E. Validity of in-home teleneuropsychological testing in the wake of COVID-19. *Arch Clin Neuropsychol*. 2021;36(6):887-896. doi:10.1093/arclin/acab002.
 12. McLaughlin PM, Sunderland KM, Beaton D, et al. The quality assurance and quality control protocol for neuropsychological data collection and curation in the Ontario Neurodegenerative Disease Research Initiative (ONDRI) study. *Assessment*. 2020;28(5):1267-1286. doi:10.1177/1073191120913933.
 13. Westervelt HJ, Bernier RA, Faust M, Gover M, Bockholt HJ, Zschiegner R, Long JD, Paulsen JS, PREDICT-HD Investigators, & Coordinators of the Huntington Study Group. Data quality assurance and control in cognitive research: Lessons learned from the PREDICT-HD study. *Int J Methods Psychiatr Res*. 2017;26(3):e1534. doi:10.1002/mpr.1534.
 14. Houston ML, Yu AP, Martin DA, Probst DY. Defining and developing a generic framework for monitoring data quality in clinical research. *AMIA Annu Symp Proc*. 2018;1300-1309.
 15. Whitney CW, Lind BK, Wahl PW. Quality assurance and quality control in longitudinal studies. *Epidemiol Rev*. 1998;20(1):71-80. doi:10.1093/oxfordjournals.epirev.a017973.
 16. Schooler JW. Metascience could rescue the “replication crisis.” *Nature*. 2014;515(7525):9. doi:10.1038/515009a.
 17. Farhan SMK, Bartha R, Black SE, Corbett D, Finger E, Freedman M, Greenberg B, Grimes DA, Hegele RA, Hudson C, Kleinstiver PW, Lang AE, Masellis M, McIlroy WE, McLaughlin PM, Montero-Odasso M, Munoz DG, Munoz DP, Strother S, Strong MJ. The Ontario Neurodegenerative Disease Research Initiative (ONDRI). *Can J Neurol Sci*. 2016;44(2):196-202. doi:10.1017/cjn.2016.415.
 18. Sunderland KM, Beaton D, Arnott SR, et al. The Ontario Neurodegenerative Disease Research Initiative. *medRxiv*. 2020. doi:10.1101/2020.07.30.20165456.
 19. Sunderland KM, Beaton D, Fraser J, et al. The utility of multivariate outlier detection

- techniques for data quality evaluation in large studies: an application within the ONDRI project. *BMC Med Res Methodol*. 2019;19(102). doi:10.1186/s12874-019-0729-9.
20. Godkin FE, Turner E, Demnati Y, et al. Feasibility of a continuous, multi-sensor remote health monitoring approach in persons living with neurodegenerative disease. *J Neurol*. 2022;269:2673–2686. doi:10.1007/s00415-021-10831-z.
 21. Marian V, Blumenfeld HK, Kaushanskaya M. The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *J Speech Lang Hear Res*. 2007;50(4):940–967. doi:10.1044/1092-4388.
 22. Calabrèse A, Cheong AMY, Cheung S, He Y, Kwon M, Mansfield JS, Subramanian A, Yu D; Legge, GE. Baseline MNREAD measures for normally sighted subjects from childhood to old age. *Invest Ophthalmol Vis Sci*. 2016;57(8):3836–3843. doi:10.1167/iovs.16-19580.
 23. Tillman TW, Carhart R. An expanded test for speech discrimination utilizing CNC monosyllabic words. *Northwestern University Auditory Test No. 6*. 1966; Brooks Air Force Base, TX: US Air Force School of Aerospace Medicine.
 24. Eilam-Stock TE, Shaw MT, Sherman K, Krupp LB, Charvet LE. Remote administration of the symbol digit modalities test to individuals with multiple sclerosis is reliable: A short report. *Mult Scler J Exp Transl Clin*. 2021;7(1):2055217321994853. doi:10.1177/2055217321994853.
 25. Tailby C, Collins AJ, Vaughan DN, Abbott DF, O'Shea M, Helmstaedter C, Jackson GD. Teleneuropsychology in the time of COVID-19: The experience of The Australian Epilepsy Project. *Seizure*. 2020;83:89-97. doi:10.1016/j.seizure.2020.06.011.
 26. Wadsworth HE, Galusha-Glasscock JM, Womack KB, Quiceno M, Weiner MF, Hynan LS, et al. Remote neuropsychological assessment in rural American Indians with and without cognitive impairment. *Arch Clin Neuropsychol*. 2016;31(5):420-425. doi:10.1093/arclin/acw030.
 27. Weintraub D, Gallagher J, Mamikonyan E, Xie S, Tran B, Shaw S. Validating virtual administration of neuropsychological testing in Parkinson disease: a pilot study. *Res Sq*; 2023. doi:10.21203/rs.3.rs-3.
 28. Kirkwood K, Peck DF, Bennie L. The consistency of neuropsychological assessments performed via telecommunication and face to face. *J Telemed Telecare*. 2000;6(3):147-151. doi:10.1258/13576330001935239.
 29. Hildebrand R, Chow H, Williams C, Nelson M, Wass P. Feasibility of neuropsychological testing of older adults via videoconference: implications for assessing the capacity for independent living. *Int J Telemed eHealth*. 2004;10(3):323-328. doi:10.1258/135763304323070751.
 30. Hunter MB, Jenkins N, Dolan C, Pullen H, Ritchie C, Muniz-Terrera G. Reliability of telephone and videoconference methods of cognitive assessment in older adults with and without dementia. *J Alzheimers Dis*. 2021;81(4):1625-1647. doi:10.3233/JAD-210088.
 31. Stine ZT. Videoconference Administration of the Delis-Kaplan Executive Function System.

- Doctoral dissertation, George Fox University. 2023;477. Retrieved from: <https://digitalcommons.georgefox.edu/psyd/477>.
32. Cullum CM, Hynan LS, Grosch M, Parikh M, Weiner MF. Teleneuropsychology: evidence for video teleconference-based neuropsychological assessment. *J Int Neuropsychol Soc.* 2014;20(10):1028–1033. doi:10.1017/S1355617714000873. PMID:25679535.
 33. Delis DC, Kaplan E, Kramer JH. Delis-Kaplan Executive Function System (D-KEFS) [Database record]. APA PsycTests. 2001. doi:10.1037/t15082-000.
 34. Benton AL, Varney NR, Hamsher KD. Judgment of Line Orientation [Database record]. APA PsycTests. 1975. doi:10.1037/t11036-000.
 35. Rey A. Rey auditory verbal learning test (RAVLT). *L'Examen clinique en psychologie [the clinical psychological examination]*. 1964. Paris: Presses Universitaires de France.
 36. Smith A. Symbol Digit Modalities Test: Manual. Western Psychological Services; 1982.
 37. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist.* 1969;9:179–86. doi:10.1093/geront/9.3_Part_1.179.
 38. Vaccarino AL, Dharsee M, Strother S, et al. Brain-CODE: A secure neuroinformatics platform for management, federation, sharing, and analysis of multi-dimensional neuroscience data. *Front Neuroinform.* 2018;12:28. doi:10.3389/fninf.2018.00028.
 39. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377–381. doi:10.1016/j.jbi.2008.08.010.
 40. Sadler ME, Yamamoto RT, Khurana L, Dallabrida SM. The impact of rater training on clinical outcomes assessment data: A literature review. *Int J Clin Trials.* 2017;4(3):101–110. doi:10.18203/2349-3259.ijct2017313.
 41. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data.* 2016;3:160018. doi:10.1038/sdata.2016.18.
 42. Broman KW, Woo KH. Data organization in spreadsheets. *Am Stat.* 2018;72(1):2–10. doi:10.1080/00031305.2017.1375989.
 43. Ellis SE, Leek JT. How to share data for collaboration. *Am Stat.* 2018;72(1):53–57. doi:10.1080/00031305.2017.1375987.
 44. Verde Arregoitia LD, Cooper N, D'Elía G. Good practices for sharing analysis-ready data in mammalogy and biodiversity research. *Hystrix: The Ital J Mammal.* 2018;29(2):155–161. doi:10.4404/hystrix-00133-2018.
 45. Keefe RSE, Harvey PD. Implementation considerations for multisite clinical trials with cognitive neuroscience tasks. *Schizophr Bull.* 2008;34(4):656–663. doi:10.1093/schbul/sbn042.
 46. Hubert M, Debruyne M. Minimum covariance determinant. *WIREs Comput Stat.*

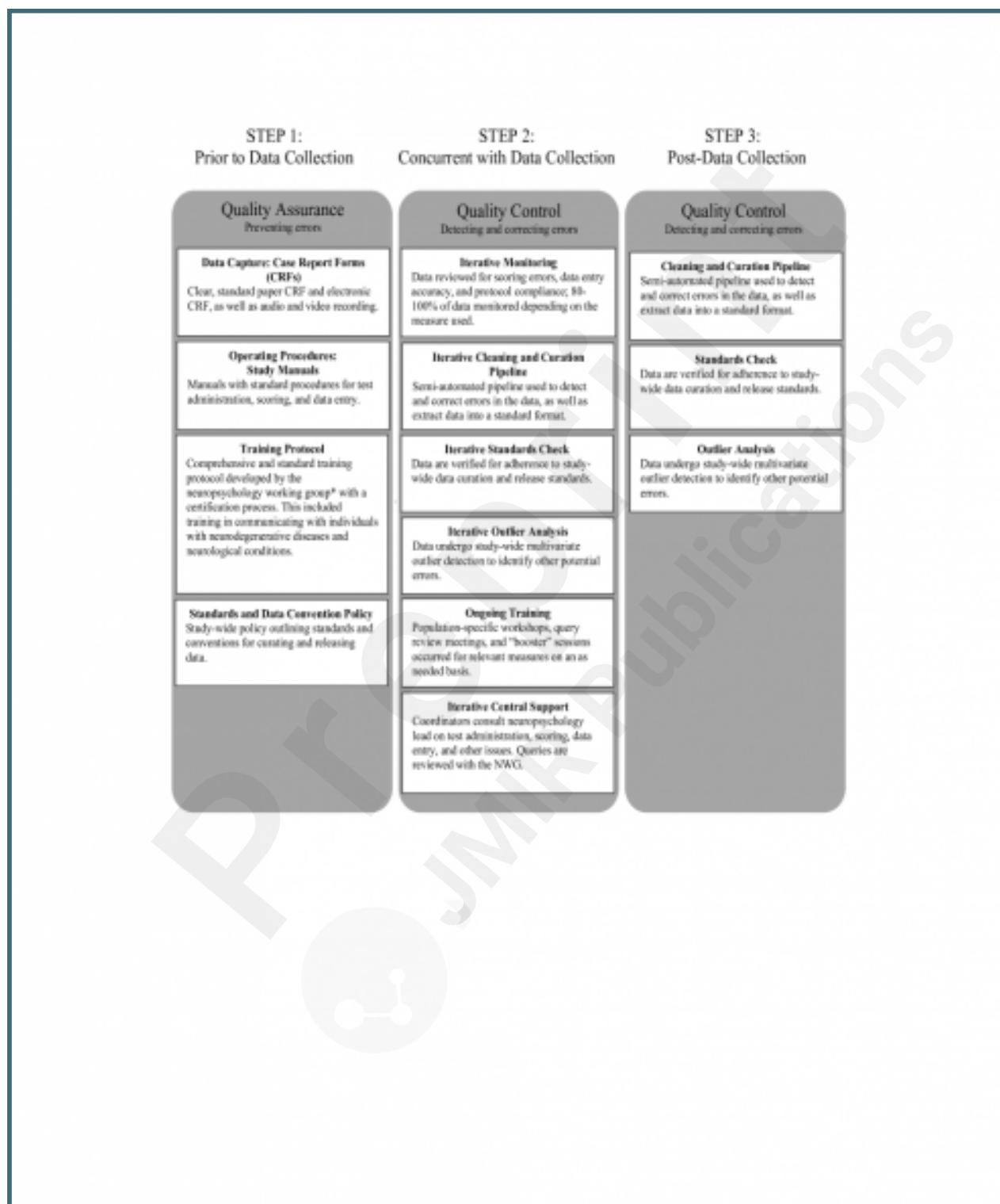
2010;2(1):36-43. doi:10.1002/wics.61.

47. Garthwaite PH, Koch I. Evaluation of the contributions of individual variables to a quadratic form. *Aust N Z J Stat.* 2016;58(1):99-119. doi:10.1111/anzs.12144.
48. Beaton D, Sunderland KM, Levine B, et al. Generalization of the minimum covariance determinant algorithm for categorical and mixed data types. *bioRxiv.* 2018. doi:10.1101/333005.
49. Beaton D, Sunderland KM. Outliers and robust structures: “OuRSs.” 2017. Available from: <https://github.com/derekbeaton/OuRS>.
50. Wechsler, D. Test of Premorbid Functioning: Advanced clinical solutions for use with WAIS-IV and WMS-IV. San Antonio, Texas: The Psychological Corporation; 2009.

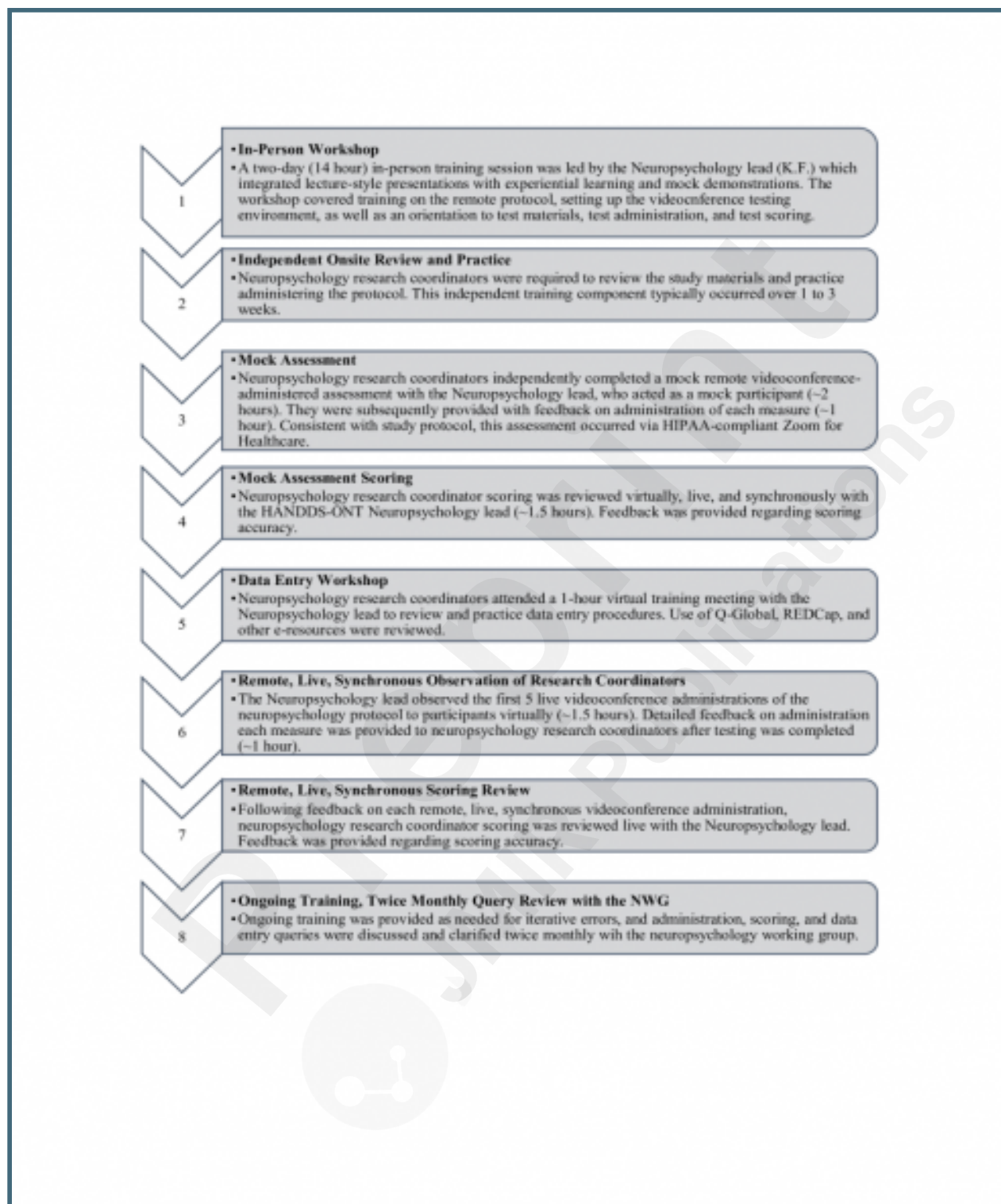
Supplementary Files

Figures

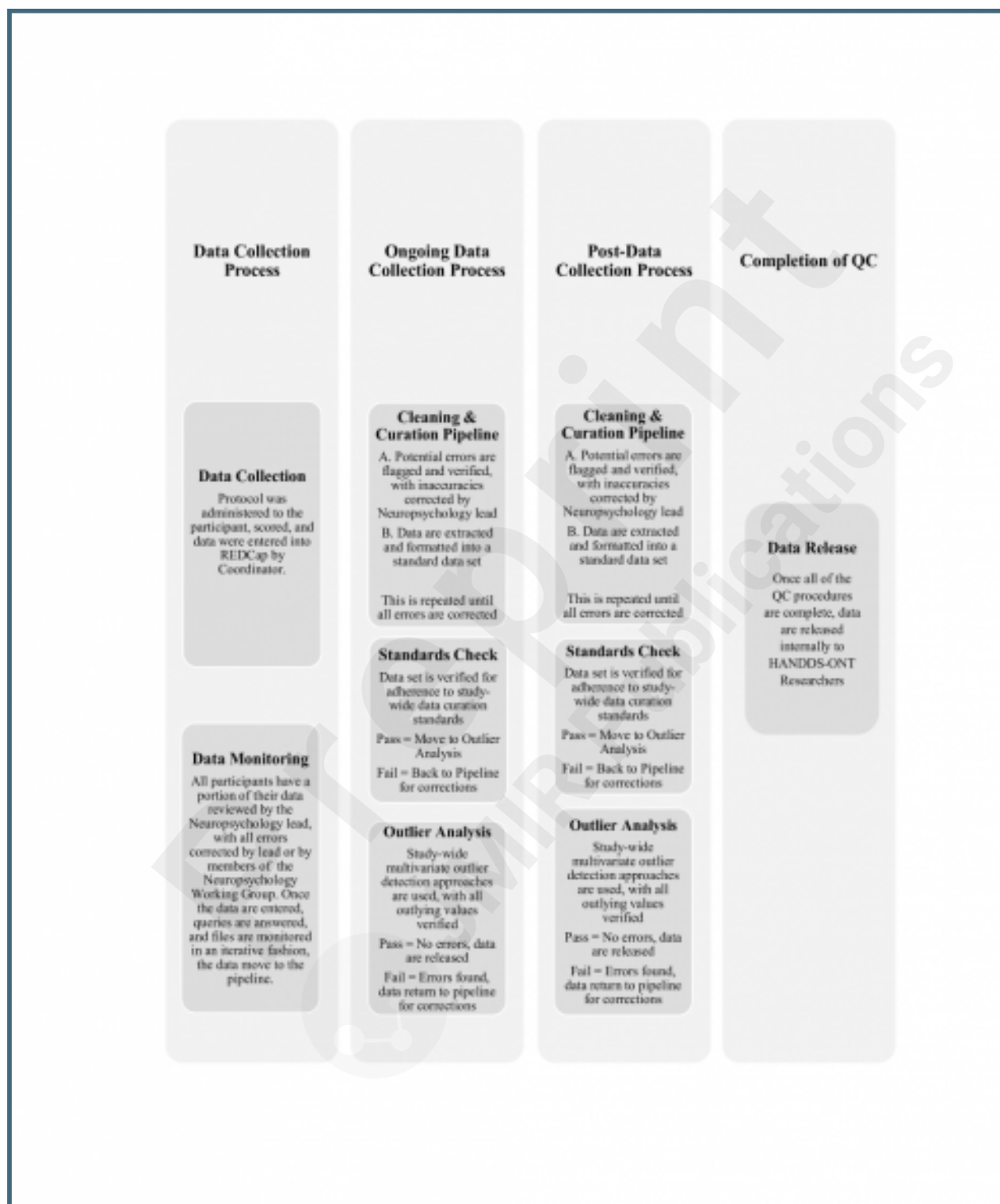
Roadmap of the schematic overview of the data fidelity (quality assurance and quality control) procedures developed and implemented in a remote, videoconference-administered, synchronous neuropsychology protocol. *Neuropsychology training roadmap further detailed in Figure 2.



Eight-step hybrid neuropsychology training protocol roadmap.



Roadmap for data collection through to data release workflow for the remote, synchronous, videoconference-administered neuropsychology protocol, adapted and revised from McLaughlin et al.'s (2020) in-person neuropsychology protocol.



Remote videoconference-administered neuropsychology protocol participant enrolment flow.

