

Remote Neuropsychological Assessments in Research: Present and Future Needs

National Institute of Neurological Disorders and Stroke

November 2-3, 2023

Virtual Meeting

Draft December 1, 2023



This meeting summary was prepared by Rose Li and Associates, Inc., under contract to the National Institute of Neurological Disorders and Stroke (NINDS). The views expressed in this document reflect both individual and collective opinions of the meeting participants and not necessarily those of NINDS. Review of earlier versions of this meeting summary by the following individuals is gratefully acknowledged: Martin Sabandal, Maya Sapiurka, Nancy Tuvevson

Acronym Definitions

DMSP	data management and sharing plan
EMA	ecological momentary assessments
IRB	Institutional Review Board
LMT	Language and Memory Test
MS	multiple sclerosis
NHLBI	National Heart, Lung, and Blood Institute
NIA	National Institute on Aging
NICHD	<i>Eunice Kennedy Shriver</i> National Institute of Child Health and Human Development
NIDCD	National Institute on Deafness and Other Communication Disorders
NIH	National Institutes of Health
NIMH	National Institute of Mental Health
NINDS	National Institute of Neurological Disorders and Stroke
NP	neuropsychology
PANDABox	Parent Administered NeuroDevelopmental Assessment Box
TabCAT	Tablet-based Cognitive Assessment Tool
TeleNP	teleneuropsychology

Table of Contents

Acronym Definitions	ii
Meeting Summary	4
Opening Remarks	4
Setting the Clinical Landscape: A Patient Perspective	4
How Do We Move from “Worked in a Pinch” to a Valid, Reliable, and Equitable Remote Assessment Approach?	5
Metrics of Equity Assurance and Success	5
Accessibility Gaps in Remote Assessments with Special Populations	6
Gaps and Needs in Remote Assessment of Cognition	6
NIH Data Sharing Policy: How This Impacts You	6
Ethics in Remote Neuropsychological Assessment: Research Applications and Clinical Implications	7
Budgeting for Remote Neuropsychological Assessments	7
Remote Assessment: Understanding Implementation and Comparability Evidence	7
Summary of Day 1 Breakouts: Special Topics (Gaps/Needs)	8
Validation of Tools into Different Languages and with Diverse Populations	8
Ensuring Methodological Fidelity and Integrity with Multi-Site Trials.....	8
Identify Safety and Reporting Issues.....	9
Identify Equivalence Versus Validity and Reliability (Comparison Between Remote and In-Person Administration)	9
Day 2: Welcome Remarks	10
Live Data Blast.....	10
Remote Assessment of Cognition: Future Directions and Opportunities	12
Novel Emerging Methodologies and Proposed Framework for Developing Standards.....	12
Accessible Designs.....	12
Discussion	12
Summary of Day 2 Breakouts: Special Topics in Optimal Development and Delivery of Remote Assessment Tools	13
Planning Remote Assessment in Natural History Studies	13
Planning Remote Assessment in Clinical Trials.....	14
Clinical Unity of Remote Research Tools.....	14
Use and Development of Apps/Computerized Measures.....	15
Conclusions and Synthesis	15
State of the Science.....	15
Opportunities in Remote Assessment.....	15
Challenges in Remote Assessment.....	15
Validation and Standardization.....	16

Meeting Summary

Opening Remarks

Walter Koroshetz, MD, National Institute for Neurological Disorders and Stroke (NINDS)

Kristi Hardy, PhD, NINDS

On November 2-3, 2023, the National Institute for Neurological Disorders and Stroke (NINDS) held a workshop titled Remote Neuropsychological Assessments: Present and Future Needs. The workshop was planned in partnership with multiple NIH Institutes and Centers including the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD), National Heart, Lung, and Blood Institute (NHLBI), National Institute on Aging (NIA), National Institute of Mental Health (NIMH), and National Institute on Deafness and Other Communication Disorders (NIDCD). Dr. Koroshetz welcomed attendees to Day 1 of the workshop, which focused on gaps, needs, and opportunities in remote assessment of cognition.

The COVID-19 pandemic necessitated a rapid adaptation of assessment tools designed for in-person use for remote use. Although remote assessments enable testing of natural behaviors in real-world environments, they must be validated against well-established in-person tools (i.e., “gold standards”) before being adapted to clinical research and practice. The workshop objectives were to (1) identify gaps in knowledge and potential pitfalls in existing remote assessments of cognition and neurodevelopment across the lifespan, (2) discuss novel or emerging methodology for assessing cognition and neurodevelopmental changes, (3) identify metrics of equity assurance and success in remote monitoring, evaluation, and visit technologies, and (4) facilitate collaboration and communication among researchers who develop and validate remote assessment measures and those who seek to utilize them for clinical research.

Setting the Clinical Landscape: A Patient Perspective

Anna Williams

Ms. Williams shared the story of her daughter, Lily, who has been diagnosed with a variety of disorders, including autism spectrum disorder, dyspraxia, Lennox-Gastaut syndrome, and Phelan-McDermid syndrome. Lily is supported by a network of medical and caretaking specialists, many of whom are only available for in-person visits. Lily experiences severe car sickness, making the hour-long drive to the medical provider’s office a stressful and painful experience for her and her family. Remote assessments eliminate these stressors, minimize disruption to everyday schedules and routines, and provide access to the home environment’s supportive devices and set-up. Ms. Williams highlighted areas of improvement for assessment developers, particularly adapting parent surveys and patient assessments to reflect the use of augmentative and alternative communication (AAC) devices.

How Do We Move from “Worked in a Pinch” to a Valid, Reliable, and Equitable Remote Assessment Approach?

Munro Cullum, PhD, University of Texas Southwestern Medical Center

Telemedicine, which was growing in popularity before 2020, increased in use by 683 percent during the COVID-19 pandemic. Telemedicine provided opportunities for clinicians to continue to provide care that was safe, effective, patient-centered, timely, efficient, and equitable. Teleneuropsychology (TeleNP) services concomitantly increased in popularity, although the feasibility, reliability, validity, and acceptability of TeleNP assessments across different populations should be explored. Dr. Cullum’s team has found no difference in the reliability, validity, feasibility, or acceptability of most tasks in common adult neuropsychological assessments when administered remotely rather than face-to-face, although concerns remain about the effect of distractions or deviations from testing norms under remote conditions. Future work should focus on validating assessments in children and populations affected by health disparities, as well as providing education and training in administering TeleNP. TeleNP tool developers should capitalize on advances in testing platforms, machine learning analyses, real-time performance feedback, and data repositories that facilitate data sharing.

Metrics of Equity Assurance and Success

Elena Tsoy, PhD, UCSF

Digital neuropsychology (NP) tools include digitalization of current tests as well as the development of new digital tests and batteries. Digital NP tools can improve the precision, administration, and scoring of assessments while decreasing costs and enabling the incorporation of multiple languages. However, digital NP tools are accompanied by technical issues that include variability in hardware and software, data and privacy concerns, and variability of user internet access and speed. Digital NP tool validation studies have also found variability in psychometric and normative data and have rarely reported on the demographic characteristics of the normative samples. Dr. Tsoy recommended that development of digital NP tools include community feedback and participation, enhance inclusion of populations with disparities in technology access, and provide user-friendly educational guidelines and materials. She shared the Tablet-based Cognitive Assessment Tool (TabCAT) as an example of a digital NP tool that has minimal user literacy and numeracy demands. TabCAT includes racially, ethnically, and culturally diverse stimuli, and recognizes and respects within-language variation. TabCAT has been adapted into 17 languages and is being used in 20 countries, with high feasibility and acceptability scores across cultural groups that indicate the broad success of TabCAT’s development model.

Accessibility Gaps in Remote Assessments with Special Populations

Moderator: Adam Hartman, MD, NINDS

Panelists: Nikki Stricker, PhD, LP, Mayo Clinic

Bridgette Kelleher, PhD, Purdue University

Audra Sterling, PhD, University of Wisconsin-Madison

Madison Berl, PhD, Children's National Hospital

Special populations in which remote assessments might be of particular use include older adults, patients with limited verbal or motor skills, and children, particularly children with intellectual or developmental disabilities or rare disorders. Remote assessments in older adults are feasible and valid, although participation rates and longitudinal challenges in this population lead to selection bias. For studies of children, especially those without verbal capabilities due to age or disability, the caregiver is considered the assessment user. Researchers must carefully consider the best environment in which to test children using remote assessments. Although children are often more natural and comfortable in the home, this setting presents more confounding factors and a higher chance of protocol deviations than the clinic or school.

Barriers to using remote assessments in special populations include technology hesitance, challenges processing and transforming data into a useful format for clinicians or regulators, and mistrust of researchers and the medical community at large. To establish trust with participants and their families, developers should collaborate with and seek input from the community of interest from the beginning of the research process and continue community engagement between and outside of studies. This approach could include hiring caregivers or other community members as a part of the research team, as well as attending patient conferences and events to talk to and recruit participants.

When working with a small population, collaboration between research groups can be critical, including finding and taking advantage of similar research topics and interests across different rare disease populations. Small population researchers must always ensure that their remote assessment protocols are generating measurements and outcomes that are meaningful and relevant to patients and their families. Achieving this goal may involve including clinical practitioners on research teams and training students and other trainees on how to manage, analyze, and communicate both small and large datasets.

Gaps and Needs in Remote Assessment of Cognition

NIH Data Sharing Policy: How This Impacts You

Tom Cheever, PhD, NINDS

The National Institutes of Health's (NIH) new data sharing policy, instituted on January 25, 2023, requires NIH funding recipients to include and comply with an approved data management and sharing plan (DMSP) and budget. The new data sharing policy applies to most NIH funding mechanisms, except for training (T), fellowship (F), construction (C06), conference

(R13), and resource (G) grants. The minimum required elements in a DMSP are (1) data type; (2) related tools, software, and code; (3) data standards; (4) data repository, access, and timeline; (4) access, distribution, and reuse considerations; and (5) oversight of data management and sharing. Funds to implement the DMSP must be requested at the time of submission; a supplemental requests will unlikely be considered post-award. Additional information can be found on the overall [NIH DMSP website](#) and the [NINDS DMSP webpage](#).

Ethics in Remote Neuropsychological Assessment: Research Applications and Clinical Implications

Gerald Young, PhD, York University

Although there are no official ethical guidelines for TeleNP, the American Psychological Association's (APA) 2017 Ethics Code provides principles and standards that can guide researchers. APA's five general ethical principles are (1) beneficence and nonmaleficence, (2) fidelity and responsibility, (3) integrity, (4) justice, and (5) respect for people's rights and dignity. To adhere to these principles, researchers should inform participants about TeleNP's disadvantages and unknowns, provide populations of interest with equitable access to TeleNP, and ensure the protection and privacy of participants' personal data. To ensure the reliability, validity, and ethics of TeleNP, remote assessments must be compared and found equivalent to face-to-face NP assessments. If implementation of TeleNP requires procedural adjustments, researchers must ensure that the altered test procedures do not compromise test reliability and validity. A study cohort that does not adequately represent disadvantaged populations violates all five APA principles; researchers should consider using a social-justice approach when recruiting participants and collecting data. Institutional Review Boards (IRBs) play an important role in ensuring equitable access to technology for all study participants as well as facilitating appropriate safety and equivalency study conduct.

Budgeting for Remote Neuropsychological Assessments

Ruth Schneider, MD, University of Rochester

When planning a study that uses remote NP assessments, researchers must budget for tests, personnel, and participants. Testing costs may include licensing and materials, data and project management platforms, and technical assistance. In addition to the research team, the study may require the assistance of a research coordinator, data or project manager, or neuropsychologist. Researchers may need to provide some participants with equipment and/or internet access to support real-time audio-video research visits and to complete the remote assessments, as well as financial compensation for their time and effort. Study design can also affect budget considerations.

Remote Assessment: Understanding Implementation and Comparability Evidence

Michelle Campbell, PhD, Center for Drug Evaluation and Research

Remote assessments increase patient clinical trial participation, accommodate greater diversity and inclusion of patients—from a broader range of ages, abilities, and backgrounds—and

enable efficiencies in clinical trial implementation. Researchers should consider how best to prioritize patient safety and privacy, maintain data quality and integrity, and train clinical trial staff and participants. Usability testing should be conducted early in the study design process to ensure that patients can complete assessments remotely and that any accessibility needs have been addressed. Missing data can occur; however, standardizing data collection, patient and personnel training, and testing environment can help mitigate the frequency of missingness. Early discussions of study design and protocols with sponsors, equipment vendors, and trial sites can also reduce missingness.

Summary of Day 1 Breakouts: Special Topics (Gaps/Needs)

Validation of Tools into Different Languages and with Diverse Populations

Elizabeth Peña, PhD, University of California, Irvine

Holly Storkel, PhD, NIDCD

- A limited number of existing remote assessment measures (i.e., LAS-FNAM, BESA, Mayo Stricker Learning Span, PANDABox) are appropriate for use with speakers of other languages or with diverse populations.
- Key issues that should be considered when adapting or translating existing measures include whether a measure can or should be translated; threats to validity when English is centered; financial costs; and accessibility to technology.
- When using measures with populations that may not be well-represented in standardization samples, researchers should account for reduced technological literacy and oversampling.
- Remote assessments should be tailored to a specific purpose.
- Remote assessments should incorporate functional, cultural, and psychometric equivalence.
- Researchers should leverage universal design principles to amplify human efforts when conducting remote assessments.

Ensuring Methodological Fidelity and Integrity with Multi-Site Trials

Svetlana Shore, PhD, Neurocrine Biosciences

Adam Hartman, MD, NINDS

- Develop standardized requirements for administrator qualifications, training, and criteria for retraining.
- Require a certain threshold of agreement with central reader as part of administrator qualification.
- Conduct central overread of a percentage of data to monitor inter-rater reliability.
- Incorporate initiatives to engage administrators, including buy-in options and accountability for data integrity.
- Standardize mobile technologies when possible.
- Standardize and mitigate high-impact variables and capture other relevant data elements of variability for evaluation.

- Ensure data are saved locally in case of WiFi signal loss; consider providing MiFi devices.
- Provide translation, back-translation, and cultural equivalence for prioritized endpoints.
- Collect demographics and common data elements for subgroup analyses.
- Consider context of use for prioritizing data standardization versus naturalistic environment.
- Standardize variables contributing to noise and beware collider bias.

Identify Safety and Reporting Issues

Travis Scott, PhD, Stanford University

Rebecca Hommer, MD, NINDS

- Have a plan for responding to safety concerns prior to starting the study.
- Set participant expectations in advance and set boundaries for when to discontinue an encounter.
- Get pertinent information at the start of the assessment.
- Identify local or accessible resources for the patient.
- Include plans for suicidality risk management.
- Know state reporting requirements regarding abuse and neglect and the associated institutional expectations.
- Expect to encounter challenging situations; temporality may further complicate both reporting considerations and ensuring patient safety.

Identify Equivalence Versus Validity and Reliability (Comparison Between Remote and In-Person Administration)

Elena Plante, PhD, University of Arizona

Tracy King, MD, MPH, NICHD

- Face-to-face tests cannot always translate to a remote assessment and may necessitate developing new measures.
- Face-to-face and remote assessments may be measuring different constructs.
- Gold standards can be flawed; researchers should explore creating a new framework to establish new gold standards.
- When an assessment fails, researchers should explore why it failed to help refine and develop new measures.
- Input from stakeholders with varied backgrounds is important at all steps of the validation process.
- Standardization and reliability are in tension with the inclusion of the breadth of the population.
- Data repositories need to include overlapping measures and internal stability data across rare disorders.

Day 2: Welcome Remarks

Alison Cernich, PhD, NICHD; Josh Gordon, MD, PhD, NIMH; Adam Hartman, MD, NINDS

Dr. Hartman welcomed participants to the second day of the meeting, which focused on the current state and future landscape of using remote assessments in cognition. Invited panelists went into various breakout rooms to discuss best approaches for optimizing the development and delivery of remote assessment tools. Presentations and discussions are summarized below.

Dr. Gordon emphasized the significance of neuropsychological assessments in mental health research. He highlighted the complexity of mental disorders and the limitations of current diagnostic systems, advocating for detailed cognitive and behavioral evaluations. He also emphasized the benefits of remote testing for its accessibility and scalability, particularly in reaching underrepresented populations. In collaboration with the *All of Us* research program, NIH is conducting cognitive tests on a large scale, aiming to understand mental health across diverse backgrounds and to facilitate better clinical predictions in mental health. Furthermore, he stressed the need for more inclusive, reliable neuropsychological tests to enhance understanding and treatment of mental illnesses.

Dr. Cernich discussed the evolution and challenges of remote neuropsychological assessments, which were initially used in military settings for concussion screening. She emphasized the need for these technologies during COVID-19 and highlighted the underrepresentation of diverse populations involved in the research studies, including minoritized groups, non-English speakers, and individuals with limited technology access. Dr. Cernich stressed the importance of including children and people with disabilities throughout the study design process, advocating for universal design to make digital technologies more adaptable. Adapting traditional tests for virtual platforms is complex and may not be as valid for certain populations. However, she expressed optimism about overcoming these inclusivity barriers with available resources and software tools, thanking the planning committee and acknowledging the collaborative efforts of various NIH Institutes and Centers.

Live Data Blast

PANDABox for Remote Monitoring of High-Risk Children

Bridgette Kelleher, PhD, Purdue University

The PANDABox (Parent Administered NeuroDevelopmental Assessment Box) study enables remote monitoring of high-risk youth, particularly children with genetic disorders such as Angelman or Fragile X. PANDABox ships testing kits to families across the United States to conduct lab-like tasks at home by combining behavioral, audio, and heart activity data from both child and parent. More than 200 assessments have been completed, with positive feedback from participating families. Next steps include (1) expanding to include electroencephalography data analysis for children with Angelman syndrome and (2) developing training hubs and software tools for heart rate data analysis. These future discoveries aim to create new metrics for measuring cognition over time.

Associations Between Remote Self-Administered Cognitive Measures and PET Biomarkers of Alzheimer's Disease

Nikki Stricker, PhD, Mayo Clinic

The Mayo Test Drive is a web-based remote self-administered cognitive testing platform. It primarily uses the Stricker Learning Span and Symbols Test functions to assess verbal memory, processing speed, and executive function in about 15 minutes. A study involving 684 older adults, primarily from the Mayo Clinic Study of Aging, investigated the platform's efficacy in identifying cognitive impairments associated with Alzheimer's disease (AD) biomarkers. Results showed significant correlations between test performance and biomarkers, particularly in participants with mild cognitive impairment or dementia. Device differences were considered but had minimal impact on outcomes. The platform's effectiveness in early identification of cognitive impairments and its potential clinical utility were highlighted.

Language and Memory Test: An App-Based Composite Cognitive Measure for Multiple Sclerosis

Vicky Leavitt, PhD, Columbia University

The Language and Memory Test (LMT) is an app-based cognitive tool that identifies early cognitive changes in multiple sclerosis (MS) patients. It assesses lexical access and visual-spatial memory, two cognitive processes most sensitively impacted in MS. Comprising three subtests—finger tapping, rapid naming, and memory accuracy—the LMT is designed for brief (3 to 5 minutes) sessions. LMT is undergoing beta testing but is poised for subsequent validation studies. A notable example of its use is in an ongoing Ethiopian clinical trial on the impact of micronutrient-fortified (i.e., vitamin B12 and folic acid) cereal on adolescent girls. LMT exemplifies the potential of low-cost, adaptable digital cognitive assessment tools.

Experiences of Spanish-Speaking Families with a Remote Developmental Assessment

Jessica Kinard, PhD, University of North Carolina at Chapel Hill

PANDABox has been culturally adapted and translated into Spanish to facilitate remote neurodevelopmental assessments for Spanish-speaking families, particularly for early autism detection. The adaptation addressed challenges such as reduced telehealth usage among Spanish speakers and the inefficiency of in-the-moment translation. After confirming the translations with native Spanish speakers and in-home trials with Spanish-speaking families, PANDABox proved feasible and acceptable. Mothers valued direct communication in Spanish, and technological usability was enhanced by the translations. Participants preferred PANDABox's culturally adapted activities and the natural home environment over clinical settings. The translation process also carefully addressed dialect variations, ensuring clarity and cultural sensitivity.

Remote Assessment of Cognition: Future Directions and Opportunities

Novel Emerging Methodologies and Proposed Framework for Developing Standards

Kathryn Papp, PhD, Harvard Medical School

Remote neuropsychological assessments are conducted outside the clinic without a clinician's supervision, often via smartphones or desktops, and are distinct from TeleNP and passive monitoring. These remote assessments are accessible and scalable, allowing for larger sample collection that is beneficial for genomic and epidemiological studies, particularly in preclinical disease stages. Challenges relate data fidelity, uncontrolled test environments, technical issues, and variability across devices and internet connections. Maintaining test security and addressing digital literacy are key concerns. The implementation of these tools in clinical practice requires establishing robust normative data, developing standardized procedures to guarantee data accuracy, and forming consensus guidelines for feedback and diagnosis. Active involvement of neuropsychologists in tool development and collaboration with big data scientists is crucial for effectively managing and interpreting the extensive digital data generated by remote assessments. These assessments are important for advancing early detection and tracking in conditions such as preclinical AD, enhancing individualized risk metrics, and improving the efficacy of remote assessments in clinical trials.

Accessible Designs

Shifali Singh, PhD, Cambridge Neuropsychology

Traditional neuropsychological tests are sensitive but fail to accurately reflect daily cognitive functioning. Digital technologies should be leveraged to monitor daily cognitive functioning, including ecological momentary assessments (EMA), speech EMA, and context analysis. These tools capture cognitive fluctuations over time and enhance knowledge of individual variability outside clinical settings. The digital toolkit TestMyBrain enables clinicians to remotely administer tests because of the extensive available normative data collected, aligning closely with traditional evaluation domains. Collaborative projects with IBM and MIT are under way to develop prognostic indicators of disease using speech analysis. Future work is focused on creating scalable, rapid testing batteries for various patient groups that can be integrated into clinical practice to enhance precision medicine. One goal of this work is to use digital methods such as wearable devices and speech analysis to improve equity and access in neuropsychology and provide more comprehensive and individualized patient assessment and care.

Discussion

Moderator: Adam Hartman, MD, NINDS

Kathryn Papp, PhD, Harvard Medical School

Shifali Singh, PhD, Cambridge Neuropsychology

Marty Sliwinski, PhD, Penn State University

Panelists discussed the challenges and opportunities of using remote neuropsychological assessments, and emphasized their utility to capture cognitive function that conventional in-

person measurements cannot. One key point was the use of construct validation rather than treating in-person measures as the gold standard. Instead, remote assessments aim for criterion validity, correlating learning curves with biomarkers and in-person measures to ensure they effectively measure cognition. This approach leverages technology to optimize the measurement of cognitive function, enabling the exploration of research questions previously unaskable with conventional methods, such as evaluating cognitive fluctuations and short-term variability.

The presentations highlighted digital tools' ability to gather large amounts of data, offering significant benefits over standard measures. However, a challenge lies in translating these digital tools from research to clinical settings. Translation requires validating that digital tools are not only measuring relevant outcomes but also at least as sensitive and specific as traditional neuropsychological measures. Furthermore, this translation process requires large sample sizes, validation against standard clinical measures, and integration of AI/ML to synthesize data effectively.

Addressing missing EMA data is complex. Adherence can be improved through strategies such as participant compensation, scheduling routine follow-up assessments, and allowing participants to control their assessment schedules. In studies with more demanding assessment schedules, bias in missing data due to cognitive impairment is a concern. Panelists also suggested involving statisticians in study planning and analysis, especially with complex data such as speech samples.

Panelists acknowledged the need for neuropsychologists to actively participate in developing and standardizing digital tools and to collaborate with data scientists to manage the large volumes of collected data. Traditional neuropsychological assessments should not be replaced but enhanced with digital measures. These enhancements will be particularly helpful when triaging patients for common diseases such as AD. Next steps include validating digital tools in larger, more diverse cohorts, understanding how learning curves change over time, and assessing the feasibility of using these digital tools among people with cognitive impairment.

Summary of Day 2 Breakouts: Special Topics in Optimal Development and Delivery of Remote Assessment Tools

Planning Remote Assessment in Natural History Studies

Kevin Krull, PhD, St. Jude Children's Research Hospital

Jacqueline Wright, DrPH, NHLBI

- Cost, feasibility, and time are critical factors in natural history studies, such as childhood cancer survival studies.
- Capturing every data point is critical for studies with smaller sample sizes.
- Study designers should identify and prioritize instruments that provide meaningful data.
- The longitudinal nature of natural history studies is challenging because assessment accuracy over time is key for conducting these types of studies.

- Tension exists among researchers between the use of old and new instruments.
- Cognitive domains are suited for remote assessment administration, including telehealth and computerized testing.
- Conducting assessments with a consistent time of day is more difficult to maintain over long intervals (e.g., 3-year study).
- The field should develop a portfolio of ideas and resources for research design and ways to engage participants.

Planning Remote Assessment in Clinical Trials

Leo Bonilha, MD, University of South Carolina

Richard Benson, MD, NINDS

- Determine the effects of remote assessments on patient anxiety and develop protocols to familiarize participants with these tools.
- Investigate how remote assessments affect anxiety among various demographics to ensure equitable access and tailor tools to diverse needs.
- Ensure that cognitive assessment methodologies in early trial phases align with those in later phases.
- Reduce study biases by assessing participants' access to and comfort with the necessary technology, including quality internet access.
- Implement systems that meet IRB standards across all sites, possibly creating a platform that ensures compliance.
- Research the ethical implications of recording assessments and develop clear protocols for obtaining informed consent from participants.
- Evaluate the consistency and validity of tools when administered remotely versus in-clinic, especially those that are verbally oriented.
- Explore the effectiveness of hybrid trials where local sites handle recruitment and intervention is centralized in order to benefit from participants' trust and familiarity with local site staff while maintaining consistent intervention delivery.
- Study the impact of remote administration on effect sizes, ensuring that such assessments are not disadvantaged compared to traditional methods.

Clinical Unity of Remote Research Tools

Michelle Chen, PhD, Rutgers University

Janna Belser-Ehrlich, PhD, NINDS

- Include relevant stakeholders (e.g., policymakers, consumers) when considering the clinical utility of remote research tools.
- Consider feasibility of tools in the clinical population of interest and their past use in similar populations when selecting measures,
- Provide supplemental funding and resources for researchers to conduct additional expert interpretation of research results.
- Increase validation studies of new measures by comparing them against traditional gold standard measures.

Use and Development of Apps/Computerized Measures

Marty Sliwinski, PhD, Penn State University

Kristina Hardy, PhD, NINDS

- View accessibility as vital, particularly for response input and collection, stimuli presentation, and task instruction.
- Ensure flexibility to adapt measures for local use.
- Balance standardization with innovation to improve accessibility and inclusion.
- Assess time-varying variables that can influence cognitive performance.
- Maintain continuity in software and technology updates, and single insurance provider decisions.
- Reconsider what the research field needs from normative data, including correcting for technical variability, geographic and subgroup factors, and repeated testing variables.

Conclusions and Synthesis

Kristina Hardy, PhD, NINDS

Dr. Hardy thanked participants, speakers, and panelists for their robust discussions throughout both days. Several key themes emerged while exploring the current science, future research landscape, ongoing challenges, and key opportunities in using remote assessments in cognition.

State of the Science

- Remote assessment of cognition is well-established, innovative, and provides essential information about cognitive health in a wide range of populations.
- Remote measures are *new* measures, even when they have been adapted from traditional assessments.
- Remote assessments are impossible to separate from technologies.
- Remote assessment differs in at least three areas that have traditionally been highly standardized, including testing platforms, testing procedures, and testing environments.

Opportunities in Remote Assessment

- Improve analysis of inter- and intra-individual variability.
- Monitor patients' cognitive health and screen them for specific impairments in a cost-effective and expeditious manner.
- Accelerate precision medicine by incorporating cognitive health metrics into remote assessments.
- Increase diversity, equity, inclusion, and accessibility to remote assessments to enhance the overall "user experience" of remote assessments.

Challenges in Remote Assessment

- Integrate clinical and research utility to spur scalability.
- Include underserved populations in research studies and clinical trials.

- Use common data elements to compare new assessments across data platforms and identify ready-for-use methods.
- Strike a balance between standardized and naturalistic methods.
- Plan for and address the unexpected by following research and clinical ethical procedures, including participant consent.

Validation and Standardization

- Validate new methods with gold standards or existing measures, although validation of remote assessments may vary by cognitive skill.
- Consider incorporating more standardization factors, including differences in testing contexts and devices, cultural and native language factors, and groups with health disparities.
- Rethink current approaches to standardization by ensuring a dynamic process and revising data standards as sample collections increase.
- Establish feasibility, reliability, and specificity of assessments for rare disease groups to determine reliable or meaningful outcomes.