OXFORD
UNIVERSITY PRESS

Archives of Clinical Neuropsychology 00 (2022) 1–10

Archives of CLINICAL NEUROPSYCHOLOGY

Preliminary Findings from a Telephone-Based Cognitive Screening of an Adult HIV Research Cohort during the COVID-19 Pandemic

Jairo A. Gonzalez[1,*], Uraina S. Clark[1], Desiree Byrd[1,2], Yusuf Clarke[1], Kaitlyn Greenwood[1], Elizabeth Tell[1], Cira Carrion-Park[1], Maria Pizzirusso[1], Rhonda Burgess[1], Susan Morgello[1,3,4]

[1]*Department of Neurology, The Icahn School of Medicine at Mount Sinai, New York City, NY, USA*
[2]*Department of Psychology, Queens College & The Graduate Center, The City University of New York, New York City, NY, USA*
[3]*Department of Neuroscience, The Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York City, NY, USA*
[4]*Department of Pathology, Icahn School of Medicine at Mount Sinai, New York City, NY, USA*

*Corresponding author at: 1 Gustave Levy Pl, MHBB 1052 10029, New York, NY 10029-6574, USA. Tel: 212-241-1665;
E-mail address: jairo.gonzalez@mssm.edu (G. Jairo)

Accepted 4 June 2022

Abstract

Objectives: Few publications have documented the utility of in-home telephone-based cognitive screeners during COVID-19. This manuscript describes the adaptation of select face-to-face (FTF) neuropsychological tests to telephonic administration in a longitudinal cohort of people with HIV (PWH). Using the cohort's pre-pandemic neuropsychological data, we explore the utility of telephonic administration in this population.
Methods: Of a longitudinal cohort of 170 adult PWH, 59 completed telephonic medical and cognitive screenings with comparable pre-pandemic FTF data. Telephone screeners and FTF evaluations were compared using repeated measures ANCOVAs to examine whether test performance differed between administration types and levels of pre-pandemic cognitive performance. Individuals with pre-pandemic test scores more than a standard deviation below the demographically-corrected mean were categorized as "below average" cognitive performance ($n = 23$), and the remainder as "average" ($n = 36$).
Results: Over 90% of participants gave positive feedback about the telephone encounter. The average cognitive performance group scored higher than the below average group on all measures across both administration types. Telephone and FTF test scores did not differ significantly for measures of category fluency, letter fluency, and verbal learning. However, the below average group scored higher on a verbal memory measure administered via telephone compared with FTF.
Conclusions: Support for telephonic adaptation of select FTF measures in longitudinal research is mixed, with verbal fluency tasks showing the strongest equivalency. When employed carefully with a clear understanding of their limitations, telephone adaptations can provide an opportunity to continue study objectives, promote equity, and monitor participant well-being during times of duress.

Keywords: Remote assessment; HIV; Neuropsychology; COVID-19

Introduction

The Coronavirus disease 2019 (COVID-19) pandemic has been a catastrophic, multi-faceted global crisis that continues to disrupt virtually every aspect of daily life, particularly as we continue to limit in-person contact. This has resulted in an increased interest in and use of video-based telehealth services and overall remote assessments across healthcare and research settings. For research neuropsychologists, the growing practice of teleneuropsychology (TNP) (i.e., use of video teleconferencing [VTC] and other platforms for the delivery of neuropsychological services) is gaining interest given its many benefits, particularly as it allows neuropsychological and other clinical services to continue remotely during the shutdowns due to COVID-19.

https://doi.org/10.1093/arclin/acac045

Several researchers have reported on the benefits of conducting videoconference-based TNP evaluations. There has been wide acceptance by patients and caregivers alike, it has been a critical and feasible way to reach wider patient populations, and it has allowed for degrees of interpersonal connectedness, particularly during the pandemic (Parikh et al., 2013; Barton, Morris, Rothlind, & Yaffe, 2011; Vahia et al., 2015; Brearly et al., 2017, Marra et al., 2020). Recent studies have shown relative validity and reliability of select neuropsychological tests administered via VTC across variable levels of cognitive functioning and with older adults, supporting its use in both major and mild neurocognitive disorders (Cullum, Hynan, Grosch, Parikh, & Weiner, 2014; Brierly et al., 2017; Wadsworth et al., 2018; Marra 2020; Parks, Davis, Spresser, Stroescu, & Ecklund-Johnson, 2021). However, despite this range of benefits and the articulation of practice guidelines, a primary concern continues to be the overall validity and reliability of videoconference administration (Chapman et al., 2020; Hammers, Stolwyk, Harder, & Cullum, 2020; Parks et al., 2021). A recent TNP survey conducted by AACN described this limitation as the most common concern among neuropsychologists in clinical and research practice (American Academy of Clinical Neuropsychology, 2020). Further concerns include lack of normative data, limited familiarity with video-based software platforms by providers and patients, as well as a reduced prospect for building rapport and obtaining behavioral observations (Brearly et al., 2017; Marra et al., 2020). Finally, logistical issues related to feasibility may exist. The "digital divide", whereby portions of the population have limited digital literacy and no access to broadband internet, may preclude VTC. This divide disproportionately impacts ethnic and racial minorities in the US; although 20% of Whites lack home broadband access, this increases to 29% for Black and 35% for Latinx populations (Edmiston & AlZuBi, 2022).

In instances where VTC or FTF assessments are not feasible, telephone-based TNP assessments may provide an alternative. Conducted over mobile or landline devices, telephone-based cognitive assessments (TBCAs) can increase access to a broader population, serve as initial and interim cognitive screeners, as well as decrease no-show rates when compared with FTF neuropsychological evaluations (Carlew et al., 2020; Elliott, Green, Llewellyn, & Quinn, 2020; Caze, Dorsman, Carlew, Diaz, & Bailey, 2020). This is significant because Black and Latinx patient populations, as well as older adults, may have decreased odds of engaging in telehealth services (i.e., video visits, e-consults, among others) as demonstrated during the COVID pandemic (Weber, Miller, Astha, Janevic, & Benn, 2020). TBCA can help bridge the digital divide and directly address systemic barriers faced by technologically and socially marginalized patient populations with limited access to high-speed, stable internet connections, and/or video platforms. Currently, TBCAs are mostly utilized in research settings for initial cognitive screening or longitudinal follow-up to detect changes in gross cognitive function in adults (Manly et al., 2011; Elliott et al., 2020). A recent TBCA meta-analysis identified over 40 cognitive screeners, batteries and single-tests adapted for telephonic administration, with the vast majority developed to screen for cognitive impairments in aging (Carlew et al., 2020). The authors concluded that TBCA may be most helpful in characterizing gross cognitive functioning over time if employed carefully with a clear purpose while increasing access to patient care.

However, there are also important limitations to TCBA, which should be utilized with caution and not serve as a replacement for standard neuropsychological evaluations. Their nature as screening tools limit clinical utility and diagnostic accuracy (Manly et al., 2011; Elliott et al., 2020). Furthermore, given the inherent nature of the telephone modality, only two cognitive domains can be reliably assessed: language and verbal memory. Although TBCA cannot be used to assess visuospatial and psychomotor skills, they may be useful in assessing executive functioning, attention, and processing speed domains with certain limitations (Carlew et al., 2020). Moreover, similar to video-based TNP, there are a lack of norms to guide interpretation of cognitive data and a limited ability to assist with physical manipulation of test materials. Lastly, with the audio-only format, one cannot obtain visual behavioral observations and assess for performance confounds relevant to data interpretation, such as distraction, hearing loss, and the use of covert ancillary aids during assessment (e.g., note writing). Prior to implementation, the researcher must safeguard against these caveats and best determine the utility and appropriateness of TBCA. Nevertheless, when done properly, this modality can address barriers inherent to in-person and VTC-based assessments, and can ensure access to individuals who would otherwise not be represented in neuropsychology research and practice.

In the unique setting of the COVID-19 pandemic, neuropsychology researchers were faced with procedural, practical, and environmental challenges with regard to continuing investigative work. Commentaries, resources, prospective solutions, and ethical considerations to clinical challenges were rapidly published, but without description of the actual processes of adaptation (Arias et al., 2020; Beane et al., 2020; Hantke & Gould, 2020; Padala, Jendro, & Padala, 2020; Stiles-Shields, Plevinsky, Psihogios, & Holmbeck, 2020; Weinberg et al., 2020; Scott, Marton, & Madore, 2021). A few manuscripts have provided evidence supporting video-based TNP conducted during the pandemic (Tailby et al., 2020; Parks et al., 2021; Pulsifer et al., 2021); however, a gap in the current COVID-era TNP literature exists with regard to the use of telephone-based assessments and how to effectively transition from FTF assessments in longitudinal studies. As such, the goal of this manuscript is to provide a description of how our research practices pivoted from in-person to telephone-based testing during early stages of the COVID-19 pandemic in New York City (NYC). We report preliminary equivalency results between FTF neuropsychological assessments and telephonic adaptations of select measures with a longitudinal cohort of mostly ethnic and racial minority research participants

living with HIV. Finally, we discuss participant reception and the benefits and limitations of implementing offsite clinical research. Altogether, it is our hope to also expand and contribute to ongoing telephone-based TNP research by investigating its utility and feasibility within longitudinal neurocognitive HIV research.

Methods

Participants

Participants were enrolled in the Manhattan HIV Brain Bank (MHBB) (U24MH100931), a longitudinal, observational cohort study of the neuropsychological, neuropsychiatric, and neuromedical consequences of HIV infection. Protocols were approved by the Icahn School of Medicine at Mount Sinai (ISMMS) Institutional Review Board. The MHBB longitudinal study includes interval in-person visits as well as phone interviews between onsite visits to query interim medical changes. In April 2020, there were 170 active MHBB participants with a mean age of 61.0 ($SD = 7.7$ years), 48% were female, and 31% Hispanic, 44% African American, 24% non-Hispanic White, and 1% other race/ethnicity, all who had provided written informed consent to participate in MHBB. Of these 170 participants, 69 completed both neuromedical and neuropsychological telephone-based assessments between April 1, 2020 and July 31, 2020, with 15 more participants completing only the neuromedical component, and 1 declining to participate altogether. There was no contact with the remainder of the cohort due to time limitations on staff during COVID re-deployments or due to participant non-responsiveness or ineligibility for remote assessments (e.g., enrolled as organ donation only, domiciled in care facilities, etc.). Of the 69 individuals completing neuromedical and neuropsychological interviews, 10 had incomplete test data as described in sections below, yielding 59 individuals for whom equivalency of phone and FTF neuropsychological test performance could be evaluated. All FTF visits for the 59 participants in this sample were conducted between March 9, 2018 and March 13, 2020. Of these 59 participants, 46% were female, 23% were Hispanic, 52% were African American, and 20% were non-Hispanic White. Mean age was 61.3 ($SD = 6.9$) and average years of education was 12.1 ($SD = 2.8$). These demographic characteristics were representative of the overall MHBB cohort (Pizzirusso et al., 2021).

Procedure

Practical considerations for telephone-based assessments. Telephone contact with our study participants began on April 1, 2020, just under 3 weeks from the last onsite assessment. Pivoting rapidly in this manner allowed the project to preserve momentum, maintain close contact with medically vulnerable participants throughout the quarantine, and prevent significant data collection gaps. All staff who completed the telephone assessments were familiar to research participants. Research staff used Google Voice numbers or Doximity to protect personal contact information when contacting study participants. Telephone cognitive screening was completed using both landlines and cell phones or through voice over IP phone services (i.e., telephone calls over a home internet connection). Participants were offered $20 compensation for the time spent in the study procedures.

Prior to telephone-based testing, MHBB staff and coordinators carefully acknowledged any acute psychosocial stress and overall anguish related to COVID-19, particularly as hundreds of individuals were losing their lives daily in NYC. Staff then non-coercively introduced the option of neurobehavioral phone assessments, explicitly emphasizing the absence of obligation. If the participant declined, the call would end promptly after a brief well-being check conducted by nursing staff. If the participant expressed interest in learning more, staff explained the telephone assessment process and provided "telephone assessment education" prior to the actual phone appointment, explaining the purpose and expectations of remote assessments as well as going over why managing external distractions was important. To address privacy concerns, MHBB staff operationalized steps to confirm consent, participant identity, and participant location. To discourage use of ancillary aids or note taking during the cognitive assessment, participants were given explicit instructions to not engage in any writing or utilize any aids (e.g., "Please do not use any paper, pencils, pens, or calculators to assist you during the evaluation"). Participants also completed a brief hearing screening (i.e., repeat three numbers) to confirm adequate sound quality for testing and were encouraged to minimize external distractions (i.e., turning off TV, moving to a private room in the home, etc.). See Table 1 for a full list of instructions and considerations for implementation of the neurocognitive testing.

Telephone-based assessment measures. Nursing staff conducted a neuromedical assessment (Table 2) in a separate phone call preceding the cognitive screening, which queried COVID-19-related physical symptoms and included short mental health and stress screens. Findings from the neuromedical screens have been previously reported (Pizzirusso et al., 2021). The subsequent cognitive screening was administered by psychometrists and was kept as brief as possible to be sensitive to participants' time and reduce the prospect of cognitively taxing participants during an already stressful time. The battery (Table 2) consisted in part of a

Table 1. Considerations for implementation of TBCAs

Before completing telephone assessment	• Dial to landline telephone if available • If calling to a cellphone, ensure adequate power and turn off notifications
During telephone assessment	• Conduct hearing test (e.g., "please repeat these numbers after me: 2–8 – 3", read at a rate of 2/s, pausing at the end to allow participant to repeat sequence • Confirm participant location. If participant is driving, walking, riding a motor vehicle, ask to call back at a different time. If participant is home, confirm that they are in a private area and continue with assessment • Remind participants to please do not use any paper, pencils, pens, or calculators to assist you during the evaluation • Discuss minimizing of distractions, including turning off TV, refraining from eating, drinking, cooking, reading, texting, socializing, listening to music, smoking, working on computer, and remind participants that the goal is to recreate the type of setting they would normally be in when conducing the assessments • Remind participant to have charger nearby and to silence the phone • Discuss back-up plan if call was to end unexpectedly
After completion of telephone assessment	• Complete a standardized behavioral notes form to document quality of assessment, rapport, auditory and/or comprehension deficits (if any), speech, cooperation, effort, attention, frustration tolerance, behaviors during testing (e.g., perfect serial order, etc.), overall validity of assessment, post-test query comments, and reasons (if any) for non-completion/discontinuation (e.g., logistics, refusal, significant cognitive deficits, hearing problems, fatigue, etc.) • Record duration of phone assessment

Table 2. Telephone-based assessment measures

Neuromedical assessment:
- COVID-19 physical symptom questionnaire
- Patient Health Questionnaire-2 (PHQ-2) (Kroenke, Spitzer, Williams, & Löwe, 2009),
- Generalized Anxiety Disorder-2 (GAD-2) (Kroenke, Spitzer, Williams, Monahan, & Löwe, 2007)
- Perceived Stress Scale-4 (PSS-4) (Cohen, Kamarck, & Mermelstein, 1994)

Neuropsychological assessment:
- HVLT-R (Benedict, Schretlen, Groninger, & Brandt, 1998) Immediate and Delayed Recall trials
- COWAT-FAS (Gladsjo et al., 1999)
- Category Fluency- Animals (Gladsjo et al., 1999)
- Beck Depression Inventory-Fast Screen (BDI-FS) (Beck, Steer, & Brown, 2000)
- A modified 5-item version of the Pittsburgh Sleep Quality Index (PSQI) (Buysse, Reynolds III, Monk, Berman, & Kupfer, 1989)
- A modified functional and cognitive screening measure that combines items from the Activities of Daily Living questionnaire (Lawton & Brody, 1969) and a subset of MHBB standard questions assessing severity and onset of functional cognitive deficits
- Kreek-McHugh-Schluger-Kellogg Scale (KMSK) (Kellogg et al., 2003)
- The Impact of Event Scale-Revised (IES-R) (Christianson & Marren, 2012)
- A COVID-19 Stress Questionnaire (CSQ) authored by the investigators for the purposes of understanding psychosocial impacts of COVID-19

redacted set of tests and questionnaires routinely administered in the MHBB study; three standard neuropsychological tests were adapted for telephone administration as well as psychiatric and functional measures. Lastly, post-assessment examiner notes and participant reception queries were included (i.e., "did you experience any difficulties during the test administration today, such as hearing difficulties, calls coming in, intrusion of others, etc.?" and "do you have any comments about this testing session that you would like to share with us at this time?"). The battery was also developed under the emerging guidance from professional organizations, including resources provided by the American Psychological Association (APA) and the Inter Organization Practice Committee (IOPC) on how to conduct neuropsychological tests remotely during COVID-19 (APA, 2020; IOPC, 2020). Administration time of self-report questionnaires and cognitive tests completed by psychometrists was approximately an hour with a range of 35 min–80 min. Disruptions to the telephone assessments were recorded on a standardized form (e.g., call disconnected, issues with call quality) as means to better contextualize the assessment and identify any likely confounding variables. Lastly, to promote test integrity and security, staff printed study forms at home when necessary and maintained a locked filing cabinet at home to store examiner record forms as well as completed phone assessments until a study clinician personally collected the forms or the team member themselves brought the forms to the medical center.

The specific neuropsychological tests adapted for telephonic administration were the Hopkins Verbal Learning Test-Revised (HVLT-R) immediate and delayed recall trials, the Controlled Oral Word Association Test-FAS (COWAT-FAS), and Category Fluency Animals. The HVLT distractor tasks between immediate recall Trial 3 and delayed recall task included the FAS, Animal Fluency, as well as the cognitive and functional screening questionnaire. The length of delay time between administrations was ∼15 min. In contrast, distractor tasks in the FTF application of the HVLT may include the BVMT-R, WAIS-III Digit Symbol and Symbol Search, FAS, Animal Fluency, and Trail Making Tests A and B. Length of the delay in the FTF assessment is ∼ 20 min. Normative sources for all tests are listed in Table 2. Ten participants were excluded from analysis due to at least one missing test from either their most recent FTF assessment ($n = 3$) or their telephone assessment ($n = 5$), as well as missing depression or stress scores from the telephone assessment ($n = 2$). For each participant, we calculated the average T-score of the aforementioned tests administered during their last FTF visit; those whose average fell below the cut-point of one standard deviation below the standard mean of 40 were classified as below average for the purposes of this study. As such, FTF scores for 23 of 59 individuals (39%) were categorized as below average and for 36 individuals (61%) as average.

Statistical Analyses

Analyses were conducted using IBM SPSS Statistics V24 (IBM Corp, SPSS Statistics V24, Armonk, NY, 2020) with a $p < 0.05$ cutoff as the significance level for primary measures. Independent sample T-tests were conducted to compare cognitive performance groups on psychiatric measures and length of administration time. Repeated measures ANCOVAs were run for each cognitive test, with cognitive performance group as the between-subjects factor (average, below average) and administration type (FTF, phone) as the within-subjects factor. Of note, T-scores were calculated using demographically corrected norms. Given significant negative correlation between levels of stress and cognitive performance during the telephonic administration across the entire sample, stress measure scores (PSS-4) were added as a covariate in each of the ANCOVAs. These models thus examined whether performance differed by administration type (FTF, phone), and whether this difference was consistent across groups. All distributions of variables were examined for normality.

Results

Participant Reception

Our research participants were generally pleased with their telephone assessment participation. In a post-study open-ended survey, 93% of our participants (55 out of 59) gave positive feedback about participating in the telephone-based assessment, either by explicitly stating that they enjoyed the experience or by simply saying they would not make any adjustments to the assessment. Within the average cognitive performance group, seven participants reported hearing difficulties or call quality issues (e.g., poor audio, passing ambulances, knocking on door), and two participants reported having some type of difficulty understanding test instructions. There were no concerns of privacy voiced by any of the participants. Only one participant reported significant fatigue to warrant splitting the neurocognitive portion of the telephone visit. Lastly, three participants reported suicidal ideation as endorsed by the BDI-FS; however, no imminent risk was detected, as level of ideation was consistent with pre-COVID levels among these participants who had documented histories of minimal to mild ideation. No other participant emergency was encountered.

Cognitive Testing in Person and by Telephone

The average time interval between FTF and telephone assessments was 6.36 months ($SD = 5.18$ months). Demographics of the cognitive performance groups, administration times for their neuropsychological phone interviews, and results of their depression and stress telephone screens are presented in Table 3. There were no significant differences in administration times between the two cognitive groups; most assessments were completed in under an hour, with five individuals in each group requiring between 60 and 80 min. The mean times for these prolonged ("upper extreme") assessments were similar for the two cognitive groups. There were no differences in BDI-FS scores across cognitive groups and no significant correlations were found between BDI-FS scores and cognitive measures administered over the telephone. The group with below average cognitive performance reported slightly higher means in the stress screener (PSS-4) over the phone, with mean differences at the trend level, $t\ (57) = 1.79$, $p = 0.078$. Significant negative correlations across the whole sample were found between PSS-4 scores and verbal learning and memory tasks (HVLT-R Immediate and Delayed) administered telephonically, $r\ (57) = -0.345$, $p < 0.01$, $r\ (57) = -0.419$, $p < 0.01$, respectively. Accordingly, stress scores were used as co-variates in subsequent analyses. Mean group T-scores for each of the individual cognitive tests across administration types are demonstrated in Table 4. ANCOVAs for each

Table 3. Demographics of cognitive performance groups, test battery administration times, and scores on stress and depression screening measures

	Below Average ($n = 23$)	Average ($n = 36$)
Age, years	59 (6.4)	62 (7.2)
Education, years	11.2 (2.9)	12.7 (2.7)
Sex, female	13 (56%)	14 (39%)
Race/ethnicity	8 (35%)	18 (50%)
Black	10 (43%)	7 (19%)
Hispanic/Latinx	4 (17%)	8 (22%)
Non-Hispanic White	1 (4%)	3 (8%)
Other	50.7 (11.6)	49.1 (10.1)
Administration time' minutes	68.0 (8.2)	67.4 (5.5)
Upper extreme of administration time"		
PSS-4*	3.3 (2.8)	2.1 (2.1)
BDI-FS	1.3 (1.6)	2 (1.6)

Notes: All results listed as means and standard deviations, with exception of sex and race/ethnicity, which are displayed as numbers and percentages.
PSS-4 = Perceived Stress Scale 4 item version; BDI-FS = Beck Depression Inventory-Fast Screen.
'Administration time for the entire neuropsychological telephone assessment
"Upper extreme refers to prolonged administration times of an hour or more; five individuals from each cognitive performance group were in this upper extreme of testing time.
*$p < 0.10$

Table 4. Results of repeated measures ANCOVAs comparing the test performance (T-scores) of two cognitive performance groups (below average, average) across two administration types (FTF, phone) controlling for scores on Perceived Stress Scale-4

Test	FTF: Mean T-scores (*SD*)		Telephone: Mean T-scores (*SD*)		Below average group-FTF versus phone (*p*-value)	Effect Size	Average group-FTF versus phone (*p*-value)	Effect Size
	Below average ($n = 23$)	Average ($n = 36$)	Below average ($n = 23$)	Average ($n = 36$)				
Animals	42.4 (14.6)	56.8 (7.9)	45.6 (10.3)	54.6 (9.3)	0.234	0.067	0.399	0.021
FAS	46.7 (7.9)	59.4 (10.2)	47.7 (9.1)	57.2 (9.8)	0.782	0.001	0.268	0.036
HVLT-R Total	23.0 (11.5)	39.8 (11.6)	25.2 (11.6)	38.3 (12.2)	0.386	0.036	0.571	0.010
HVLT-R Delayed	16.2 (12.5)	40.4 (14.1)	28.0 (12.4)	38.1 (15.0)	**0.019** *	0.234	0.488	0.014

Note: HVLT-R = Hopkins Verbal Learning Test-Revised; FTF = face-to-face. Covariate: score on Perceived Stress Scale (PSS-4)
*$p < 0.05$. Significant mean T-score differences on HVLT-R Delayed task between administration types found in the 'Below average' group

test revealed a significant main effect of group (Fs [1, 55] >22.2, ps < 0.001), indicating that the group with average cognitive performance did significantly better on each cognitive measure, across both administration types, as expected. In addition, we observed a significant interaction effect between cognitive performance group and type of administration on the verbal memory task (HVLT-R Delayed; F [1] =13.29, $p < 0.01$), revealing that the group with below average cognitive performance did better on the telephone than during the FTF visit (F [1] = 6.4, $p < 0.05$).

Discussion

This manuscript is among the first to explore non-VTC remote assessment modalities with a racially and ethnically diverse research sample during the COVID-19 pandemic. Overall, our preliminary results suggest that test scores did not differ significantly between most recent FTF assessment and telephonic administration for category fluency, letter fluency, and verbal learning tasks among individuals living with HIV. This finding is consistent with research that suggests no significant difference between administration type (in-person, phone) for verbal fluency measures (Bunker et al., 2017; Christodoulou et al., 2016) and for verbal learning, specific to the HVLT-R Immediate Recall task (Bunker et al., 2017; Julian et al., 2012). By contrast, results for verbal memory suggest that the group with below average cognitive performance did better on the phone than during their most recent in-person visit in the HVLT-R Delayed Recall task. This observation is not without precedent. Although most evidence suggests comparable performance between in-person and telephone administration on verbal memory tasks (Julian et al., 2012; Bunker et al., 2017; Barcellos et al., 2018), our findings align with a few others who have found low concurrent validity between modalities. For instance, when compared with FTF assessments, increased performance on telephone verbal memory tasks was observed in younger adults completing the Rivermead Behavioral Memory Test

(Kliegel, Martin, & Jäger, 2007). In contrast, decreased performance in phone administration relative to FTF was observed in an older, female cohort using the Selective Reminding Test (Mitsis et al., 2010). In this latter sample, it was suggested that this may have been related to learning effects, as six learning trials were utilized with in-person visits versus three learning trials over the telephone (Mitsis et al., 2010). Indeed, some of the variability in FTF-telephone concordance may be related to the length and type of the memory test utilized. Overall, shorter tests of verbal memory may be less reliable via phone, particularly for those individuals who demonstrate mild-to-moderate cognitive impairment across multiple cognitive domains. Adding longer list-learning tests or story-based recall measures, while paying close attention to practical issues for telephonic administration (e.g., audio quality, call drop, increased likelihood of fatigue and distractions in longer assessments, etc.), may enhance sensitivity in telephone assessments and may be more effective at assessing episodic memory function across age and diagnostic groups (Carlew et al., 2020). The discrepancy for the recall trial seen in our group with below average cognitive performance may in part have reflected the shorter and potentially more variable nature of the HVLT (with only 3 learning trials). In addition, within this cognitive subgroup, the shorter interval between learning and delayed recall trials, with cognitively lighter distractor tasks, may also have influenced performance. However, it is unclear if these differences are relevant to performance in our cohort.

There are other potential explanations for our observation of better performance in telephone-based assessment of verbal memory in cognitively impaired individuals. It may be related to the fluctuation in neurocognitive performance seen in individuals aging with HIV, which is particularly seen in individuals with milder forms of cognitive impairment (e.g., Mild Neurocognitive Disorder) (Antinori et al., 2007). At different examinations, these bidirectional changes (normal, abnormal) in neurocognitive symptomatology have been observed in three separate HIV+ cohorts over time (Antinori et al., 2007). Although changes in neurocognitive symptomatology rely on the assessment of multiple cognitive domains, it is important to note that observed improvement in verbal memory may be a small part of the overall fluctuating nature of neurocognitive functioning over time in this patient population. It is also possible that the test taking environment impacts test performance; individuals with mild cognitive impairments may be more comfortable in their home surroundings, which may serve to enhance their capabilities. Lastly, there is always a risk of participants writing down items during list-learning tasks administered over the telephone. However, this was actively discouraged in our telephone assessments with explicit instructions before administration (e.g., “Please do not use any paper, pencils, pens, or calculators to assist you during the evaluation”).

Although not asked directly, a handful of participants inquired about VTC administration or said they “wished this was their normal visit,” which may be indications of a positive attitude towards remote assessments in general. Some elements that may have contributed to a positive reception from participants included a streamlined process of remote assessment handoffs among nursing team and psychometrists, clear explanations of the purpose and expectations of remote assessments, as well as an empathic approach to data collection, whether done FTF or over the telephone. Altogether, these preliminary results align well with patient acceptability of VTC-based neuropsychological assessment, as was evidenced in a recent study where the majority of surveyed older adults reported being satisfied with TNP (Parikh et al., 2013). Overall, participants were generally appreciative of the phone call, as some expressed gratitude towards the research team for also “checking in” during the pandemic. Although the telephone neuropsychology assessment lasted from 35 to 80 min, these results show that remote telephone assessments were generally well accepted and tolerated by our research participants. Altogether, our results, although preliminary, offer a specific perspective on the utility and feasibility on the telephonic adaptation of select FTF cognitive tests with research participants who may not be able to complete FTF or VTC cognitive assessments and expand TNP research specific to ethnic minority populations utilizing non-VTC assessment modalities.

Limitations

Limitations to the current findings include a small sample size and the use of an HIV cohort, which may limit generalizability of the results, the limited number of cognitive tests administered to assure brevity of telephone assessment, and the lack of familiarity with telephone-based cognitive testing in our cohort. Limitations specific to the verbal memory findings include shorter delay periods than typically employed in FTF and verbal fluency tests used as distractor tasks. Another limitation is related to how cognitive performance was defined in this study. Although MHBB participants are assigned HIV-informed neurocognitive diagnoses based on FTF assessment results, for this study, average T-scores were utilized as the cognitive performance classification system to remain consistent between administration modalities (not enough cognitive domains were assessed over the phone for diagnostic purposes). Lastly, there may be limited applicability of existing norms for tests administered over the telephone. Although our study has demonstrated the ability to conduct TNP during the stressful, early stages of the pandemic, future research should focus on the utility of TNP to elaborate the overall impact of the COVID-19 pandemic on neurocognitive functioning. This may include both short- and long-term neuropsychological impacts of COVID-19 infection, as well as exploring neurocognitive functioning in the context of psychological, medical, and environmental variables related to the pandemic (e.g., medical comorbidity, socioeconomic changes due to COVID-19 pandemic, stress, sleep

quality, substance use, etc.). This will be particularly useful in cohort studies with pre-pandemic neurocognitive, medical, and psychological data.

Conclusion

This study provides limited support for telephonic adaptation of select FTF neuropsychological measures in longitudinal research. Verbal fluency tasks showed the strongest equivalency between administration types and across levels of cognitive functioning. However, more work needs to be done to understand the suitability of telephonic adaptation of verbal memory measures with individuals of variable cognitive ability. Telephonic adaptation of screeners, batteries and single cognitive tests, although not meant to replace standard FTF neuropsychological assessments, can increase equity by directly addressing technological and environmental barriers that may limit individuals from completing FTF or VTC-based research assessments, particularly during the ongoing COVID-19 pandemic. Furthermore, it can help ensure that research groups remain connected to participants whose need of institutional support may be more critical, specifically for ethnic and racial minorities. In addition, these adaptations may reduce operational costs for research programs by allowing more flexibility and, along with VCT-based assessments, can be feasible and practical ways to complement FTF research visits. Remote evaluations, whether through the telephone or through VTC platforms, bring very promising prospects for researchers; however, they remain in early stages with regard to standardization of administration guidelines and validation of normative scores. The emergence of the COVID-19 pandemic provides an opportunity to reconsider and expand the optimal medium for delivery of research cognitive assessments and longitudinal data collection, in the continual quest to find accurate and validated means that are balanced against the exigencies of cost and environmental challenges.

Acknowledgements

The authors would like to thank the patients and staff of the MHBB and would like to dedicate this manuscript to the three MHBB participants whose lives were lost during the early stages of the COVID-19 crisis.

Funding

This work was supported by NIMH under grant U24MH100931.

References

American Academy of Clinical Neuropsychology. (2020). *AACN COVID-19 survey responses*. Retrieved August 14, 2021, from https://theaacn.org/wp-content/uploads/2020/04/AACN-COVID-19-Survey-Responses.pdf.

American Psychological Association. (2020). *Office and technology checklist for telepsychological services*. Retrieved April, 27 2020. https://www.apa.org/practice/programs/dmhi/research-information/telepsychology-services-checklist.pdf

Antinori, A., Arendt, G., Becker, J. T., Brew, B. J., Byrd, D. A., Cherner, M., et al. (2007). Updated research nosology for HIV-associated neurocognitive disorders. *Neurology*, *69*(*18*), 1789–1799. https://doi.org/10.1212/01.WNL.0000287431.88658.8b.

Arias, F., Safi, D. E., Miranda, M., Carrión, C. I., Diaz Santos, A. L., Armendariz, V., et al. (2020). Teleneuropsychology for monolingual and bilingual Spanish-speaking adults in the time of COVID-19: Rationale, professional considerations, and resources. *Archives of Clinical Neuropsychology*, *35*(*8*), 1249–1265. https://doi.org/10.1093/arclin/acaa100.

Barcellos, L. F., Bellesis, K. H., Shen, L., Shao, X., Chinn, T., Frndak, S., et al. (2018). Remote assessment of verbal memory in MS patients using the California verbal learning test. *Multiple Sclerosis Journal*, *24*(*3*), 354–357. https://doi.org/10.1177/1352458517694087.

Barton, C., Morris, R., Rothlind, J., & Yaffe, K. (2011). Video-telemedicine in a memory disorders clinic: Evaluation and management of rural elders with cognitive impairment. *Telemedicine Journal and E-Health*, *17*(*10*), 789–793. https://doi.org/10.1089/tmj.2011.0083.

Beane, J. D., Dedhia, P. H., Ejaz, A., Contreras, C. M., Cloyd, J. M., Tsung, A., et al. (2020). Conducting clinical trials in the time of a pandemic. *Annals of Surgery*, *272*(*3*), e219–e221. https://doi.org/10.1097/SLA.0000000000004114.

Beck, A. T., Steer, R. A., & Brown, G. K. (2000). *Beck depression inventory fast screen for medical patients*. San Antonio, TX: Psychological Corporation.

Benedict, R. H., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins verbal learning test–revised: Normative data and analysis of inter-form and test-retest reliability. *The Clinical Neuropsychologist*, *12*(*1*), 43–55. https://doi.org/10.1076/clin.12.1.43.1726.

Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., et al. (2017). Neuropsychological test administration by videoconference: A systematic review and meta-analysis. *Neuropsychology Review*, *27*(*2*), 174–186. https://doi.org/10.1007/s11065-017-9349-1.

Bunker, L., Hshieh, T. T., Wong, B., Schmitt, E. M., Travison, T., Yee, J., et al. (2017). The SAGES telephone neuropsychological battery: Correlation with in-person measures. *International Journal of Geriatric Psychiatry*, *32*(*9*), 991–999. https://doi.org/10.1002/gps.4558.

Buysse, D. J., Reynolds, C. F., III, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(*2*), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4.

Carlew, A. R., Fatima, H., Livingstone, J. R., Reese, C., Lacritz, L., Pendergrass, C., et al. (2020). Cognitive assessment via telephone: A scoping review of instruments. *Archives of Clinical Neuropsychology*, *35*(*8*), 1215–1233. https://doi.org/10.1093/arclin/acaa096.

Caze, T., Dorsman, K. A., Carlew, A. R., Diaz, A., & Bailey, K. C. (2020). Can you hear me now? Telephone-based teleneuropsychology improves utilization rates in underserved populations. *Archives of Clinical Neuropsychology*, *35*(*8*), 1234–1239. https://doi.org/10.1093/arclin/acaa098.

Chapman, J. E., Ponsford, J., Bagot, K. L., Cadilhac, D. A., Gardner, B., & Stolwyk, R. J. (2020). The use of videoconferencing in clinical neuropsychology practice: A mixed methods evaluation of neuropsychologists' experiences and views. *Australian Psychologist*, *55*(*6*), 618–633. https://doi.org/10.1111/ap.12471.

Christianson, S., & Marren, J. (2012). The impact of event scale–revised (IES-R). *Medsurg Nursing*, *21*(*5*), 321–323.

Christodoulou, G., Gennings, C., Hupf, J., Factor-Litvak, P., Murphy, J., Goetz, R. R., et al. (2016). Telephone based cognitive-behavioral screening for frontotemporal changes in patients with amyotrophic lateral sclerosis (ALS). *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *17*(*7–8*), 482–488. https://doi.org/10.3109/21678421.2016.1173703.

Cohen, S., Kamarck, T., & Mermelstein, R. (1994). Perceived stress scale. In *Measuring stress: A guide for health and social scientists*, Oxford University Press on Demand (Vol. *10*, pp. 1–2).

Cullum, C., Hynan, L., Grosch, M., Parikh, M., & Weiner, M. (2014). Teleneuropsychology: Evidence for video teleconference-based neuropsychological assessment. *Journal of the International Neuropsychological Society*, *20*(*10*), 1028–1033.

Edmiston, K. D., & AlZuBi, J. (2022). *Trends in telehealth and its implications for health disparities*. https://www.researchgate.net/profile/Kelly-Edmiston/publication/360851588_Trends_in_Telehealth_and_Its_Implications_for_Health_Disparities/links/628e8a0255273755ebb51cad/Trends-in-Telehealth-and-Its-Implications-for-Health-Disparities.pdf

Elliott, E., Green, C., Llewellyn, D. J., & Quinn, T. J. (2020). Accuracy of telephone-based cognitive screening tests: Systematic review and meta-analysis. *Current Alzheimer Research*, *17*(*5*), 460–471. https://doi.org/10.2174/1567205017999200626201121.

Gladsjo, J. A., Schuman, C. C., Evans, J. D., Peavy, G. M., Miller, S. W., & Heaton, R. K. (1999). Norms for letter and category fluency: Demographic corrections for age, education, and ethnicity. *Assessment*, *6*(*2*), 147–178. https://doi.org/10.1177/107319119900600204.

Hammers, D. B., Stolwyk, R., Harder, L., & Cullum, C. M. (2020). A survey of international clinical teleneuropsychology service provision prior to and in the context of COVID-19. *The Clinical Neuropsychologist*, *34*(*7–8*), 1267–1283. https://doi.org/10.1080/13854046.2020.1810323.

Hantke, N. C., & Gould, C. (2020). Examining older adult cognitive status in the time of COVID-19. *Journal of the American Geriatrics Society*, *68*(*7*), 1387–1389. https://doi.org/10.1111/jgs.16514.

Inter Organizational Practice Committee. (2020). *Tele-neuropsychology during the COVID-19 pandemic*. Available at https://iopc.squarespace.com/teleneuropsychology. Accessed on April 2020.

Julian, L. J., Yazdany, J., Trupin, L., Criswell, L. A., Yelin, E., & Katz, P. P. (2012). Validity of brief screening tools for cognitive impairment in rheumatoid arthritis and systemic lupus erythematosus. *Arthritis Care & Research*, *64*(*3*), 448–454. https://doi.org/10.1002/acr.21566.

Kellogg, S. H., McHugh, P. F., Bell, K., Schluger, J. H., Schluger, R. P., LaForge, K. S., et al. (2003). The Kreek–McHugh–Schluger–Kellogg scale: A new, rapid method for quantifying substance abuse and its possible applications. *Drug and Alcohol Dependence*, *69*(*2*), 137–150. https://doi.org/10.1016/S0376-8716(02)00308-3.

Kliegel, M., Martin, M., & Jäger, T. (2007). Development and validation of the cognitive telephone screening instrument (COGTEL) for the assessment of cognitive function across adulthood. *The Journal of Psychology*, *141*(*2*), 147–170. https://doi.org/10.3200/JRLP.141.2.147-172.

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ–4. *Psychosomatics*, *50*(*6*), 613–621. https://doi.org/10.1176/appi.psy.50.6.613.

Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, *146*(*5*), 317–325. https://doi.org/10.7326/0003-4819-146-5-200703060-00004.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist*, *9*(*3 Part 1*), 179–186. https://doi.org/10.1093/geront/9.3_Part_1.179.

Manly, J. J., Schupf, N., Stern, Y., Brickman, A. M., Tang, M. X., & Mayeux, R. (2011). Telephone-based identification of mild cognitive impairment and dementia in a multicultural cohort. *Archives of Neurology*, *68*(*5*), 607–614. https://doi.org/10.1001/archneurol.2011.88.

Marra, D. E., Hamlet, K. M., Bauer, R. M., & Bowers, D. (2020). Validity of teleneuropsychology for older adults in response to COVID-19: A systematic and critical review. *The Clinical Neuropsychologist*, *34*(*7–8*), 1411–1452. https://doi.org/10.1080/13854046.2020.1769192.

Mitsis, E. M., Jacobs, D., Luo, X., Andrews, H., Andrews, K., & Sano, M. (2010). Evaluating cognition in an elderly cohort via telephone assessment. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences*, *25*(*5*), 531–539. https://doi.org/10.1002/gps.2373.

Padala, P. R., Jendro, A. M., & Padala, K. P. (2020). Conducting clinical research during the COVID-19 pandemic: Investigator and participant perspectives. *JMIR Public Health and Surveillance*, *6*(*2*), e18887. https://doi.org/10.2196/18887.

Parikh, M., Grosch, M. C., Graham, L. L., Hynan, L. S., Weiner, M., Shore, J. H., et al. (2013). Consumer acceptability of brief videoconference-based neuropsychological assessment in older individuals with and without cognitive impairment. *The Clinical Neuropsychologist*, *27*(*5*), 808–817. https://doi.org/10.1080/13854046.2013.791723.

Parks, A. C., Davis, J., Spresser, C. D., Stroescu, I., & Ecklund-Johnson, E. (2021). Validity of in-home teleneuropsychological testing in the wake of COVID-19. *Archives of Clinical Neuropsychology*, *36*(*6*), 887–896. https://doi.org/10.1093/arclin/acab002.

Pizzirusso, M., Carrion-Park, C., Clark, U. S., Gonzalez, J., Byrd, D., & Morgello, S. (2021). Physical and mental health screening in a New York City HIV cohort during the COVID-19 pandemic: A preliminary report. *Journal of Acquired Immune Deficiency Syndromes*, *86*(*3*), e54–e60. https://doi.org/10.1097/QAI.0000000000002564.

Pulsifer, M. B., Grieco, J. A., Burstein, S. M., Parsons, M. W., Gardner, M. M., & Sherman, J. C. (2021). The development and implementation of teleneuropsychology in an academic lifespan neuropsychology center: Lessons learned from the COVID-19 pandemic. *Journal of Clinical and Experimental Neuropsychology*, *43*(*8*), 774–785. https://doi.org/10.1080/13803395.2021.1963683.

Scott, T. M., Marton, K. M., & Madore, M. R. (2021). A detailed analysis of ethical considerations for three specific models of teleneuropsychology during and beyond the COVID-19 pandemic. *The Clinical Neuropsychologist*, *36*(*1*), 24–44. https://doi.org/10.1080/13854046.2021.1889678.

Stiles-Shields, C., Plevinsky, J. M., Psihogios, A. M., & Holmbeck, G. N. (2020). Considerations and future directions for conducting clinical research with pediatric populations during the COVID-19 pandemic. *Journal of Pediatric Psychology*, *45*(*7*), 720–724. https://doi.org/10.1093/jpepsy/jsaa055.

Tailby, C., Collins, A. J., Vaughan, D. N., Abbott, D. F., O'Shea, M., Helmstaedter, C., et al. (2020). Teleneuropsychology in the time of COVID-19: The experience of the Australian epilepsy project. *Seizure*, *83*, 89–97. https://doi.org/10.1016/j.seizure.2020.10.005.

Vahia, I. V., Ng, B., Camacho, A., Cardenas, V., Cherner, M., Depp, C. A., et al. (2015). Telepsychiatry for neurocognitive testing in older rural Latino adults. *The American Journal of Geriatric Psychiatry*, *23*(*7*), 666–670. https://doi.org/10.1016/j.jagp.2014.08.006.

Wadsworth, H. E., Dhima, K., Womack, K. B., Hart, J., Jr., Weiner, M. F., Hynan, L. S., et al. (2018). Validity of teleneuropsychological assessment in older patients with cognitive disorders. *Archives of Clinical Neuropsychology*, *33*(*8*), 1040–1045. https://doi.org/10.1093/arclin/acx140.

Weber, E., Miller, S. J., Astha, V., Janevic, T., & Benn, E. (2020). Characteristics of telehealth users in NYC for COVID-related care during the coronavirus pandemic. *Journal of the American Medical Informatics Association*, *27*(*12*), 1949–1954. https://doi.org/10.1093/jamia/ocaa216.

Weinberg, M. S., Patrick, R. E., Schwab, N. A., Owoyemi, P., May, R., McManus, A. J., et al. (2020). Clinical trials and tribulations in the COVID-19 era. *The American Journal of Geriatric Psychiatry*, *28*(*9*), 913–920. https://doi.org/10.1016/j.jagp.2020.05.016.