Alzheimer disease and related dementias (ADRDs) are a global medical and public health challenge. Early and accurate diagnosis of these conditions can mitigate costs and improve medical care and quality of life. Cognitive assessment is an important part of ADRD diagnosis, and as we move into an age of ADRD therapeutics, accurate syndromic classification and disease monitoring will be a critical component of identifying trial participants and managing these diseases. Since the advent of clinical neuropsychologic assessments, these evaluations have been conducted using paper-and-pencil measures. It is becoming clear, however, that there are myriad benefits to incorporating digital technologies into assessment. Paper-and-pencil tests have typically been administered face-to-face in the clinic. The necessity of validated remote evaluations is apparent, underscored by the recent limitations on in-person visits secondary to the COVID-19 pandemic. Digital technologies are well positioned for a role in remote evaluation.

In this review, we cover several topics we hope will guide clinical neurologists considering adoption of technology in their practice, with a focus on ADRD assessment. We discuss benefits and limitations of adopting digital assessment tools and considerations for deciding which to use. We then review advances in tablet- and smartphone-based digital assessments. Personal computer (PC)-based evaluations have been reviewed elsewhere. Passive digital phenotyping and virtual reality assessments are not covered here. We conclude with practical strategies for bringing digital devices into clinical workflows.

Benefits of Digital Assessments
Integrating digital technologies into clinical practice and research has several possible benefits. Digital assessments may reduce time and cost associated with cognitive testing. If self-administered, there is also potential for reducing costs associated with staffing. This benefit is further increased if testing is performed remotely, rather than in the office or research center. Adaptive tests, such as those using item response theory, could reduce the time associated with determining an individual’s performance level. Many digital assessments provide an automated scoring system, which can reduce clinician time and the likelihood of scoring errors. Some assessments take the additional step of providing interpretative reports and evidence-based care recommendations. Remotely administered (online) technologies can be deployed quickly to a large group of participants and allow for frequent repeated assessments, providing a richer, more accurate classification of cognitive performance and longitudinal change. Naturalistic data collection in a nonclinical uncontrolled environment with everyday distractions may provide a sampling of cognitive abilities that more closely reflects real-world cognitive function. Digital assessments could also increase the reach of our assessment capabilities, potentially addressing health care disparities. For example, self-administered digital tests easily can be administered in different languages, and remote assessments could help reach persons in underserved communities who have barriers to clinic evaluations. Caution must be exerted, however, to ensure that digital assessments do not unintentionally exacerbate health disparities, considering that people in vulnerable groups may have limited digital literacy or access. Digital technologies may improve assessment accuracy and validity. For example, digital assessments have the potential for more accurate quantification of reaction time and motor functioning and analyzing passive data (eg, daily functioning and language) could supplement standard neuropsychologic assessments.

Challenges and Outstanding Questions
Although there is great enthusiasm about augmenting traditional neuropsychologic assessment with innovative technolo-
gories, it is challenging to choose the most appropriate device for a given clinical or research use. Because this is a relatively nascent field, there are several potential pitfalls, and caution should be exerted when considering digital assessment tools.\(^3\,^5\)

Choosing a Device

An initial decision point is whether the examination will be self-administered or if an examiner (eg, medical assistant or trained psychologist) will conduct the evaluation. Although self-administration methods could reduce time and costs, the benefits may be offset by several challenges. For example, it is difficult to control environmental factors and distractions, to ensure that the examinee understands the task instructions, and to evaluate feigned performance.\(^3\,^5\)

For self-administered assessments, a second decision point is use of a “managed” vs “bring-your-own” device (BYOD) model. Using a managed device (ie, a dedicated device for the cognitive assessment rather than the examinee’s device) reduces error variance associated with different hardware and software. Limitations to using a managed device include the resources required for device purchase and reduced flexibility. Individuals will also have differing levels of familiarity with the chosen device, which might affect performance.\(^3\)

A BYOD model, often used in smartphone and browser-based testing, has the potential to reach a wider audience without the added expense of a managed device. Smartphones offer myriad novel data collection methods (see below). In a BYOD model, however, there are unresolved concerns about the effect of different hardware (eg, device size, type, touch screen, or touch-screen responsiveness) and software (eg, iOS vs Android, updates).\(^3\,^5\)\(^\text{3.5}\)\(^\text{3.5}\) Furthermore, browser-based tests may produce inaccurate stimulus-display timing and reaction-time capture when internet connectivity is suboptimal.\(^5\)

When choosing a device, it is important to ensure that the software and hardware of the device are compatible with the minimum requirements specified in the software manual.\(^3\) It is also critical to review the required administrator qualifications (eg, medical assistant vs graduate degree).\(^3\)

Scoring and Interpretation

Automated scoring and report generation are salient benefits of digital assessments. Boilerplate reports or diagnostic levels that are appealing for their simplicity, however, have the potential to mislead by giving the impression that supportive clinical history, imaging, lab studies, and more are unnecessary. Automated reports should be interpreted by qualified clinicians in the context of all other clinical data.\(^3\) When interpreting the scores, it is important to understand how the normative data for that test was derived. It cannot be assumed that digital versions of standard cognitive tests have the same reliability and validity as the original tests, nor should they be compared to the same normative data.\(^3\)

Available Digital Assessments

The list of digital cognitive assessment tools is large and expanding. Recent interest has grown exponentially (Figure). Because this article was written for a special issue on ADRD, we chose to highlight tablet-based measures and batteries validated for ADRD. Specifically, we highlight tools: 1) validated for differential diagnosis or early detection of ADRD, not simply being able to differentiate dementia from controls, 2) shown to be associated with ADRD biomarkers, or 3) evaluated for longitudinal disease monitoring. We loosened these criteria for our review of smartphone assessments, because of the lack of validated measures ready for clinical ADRD evaluations. This brief review is not meant to be exhaustive.

Tablet-Based Assessments

There has been a growing interest in the development and validation of tablet-based cognitive assessment tools. These devices have undergone a major shift in adoption across different demographics and have less variable device characteristics compared with smartphones and PCs.\(^4\) In particular, prior evidence suggests that adults over age 55 have a greater preference for use of touchscreen devices because of the direct and intuitive interaction, lower motor demands, and relative ease of use, even by examinees without prior experience.\(^10\,^11\) Moreover, tablets offer greater mobility than PCs and may be more user-friendly than smartphones for older adults given larger screen sizes and response fields.\(^12\) The expanding market of these appliances warrants attention to variability in features such as response rate logging, screen size and resolution, rendering of visual display elements, and frame rates; all of these factors may contribute to measurement error.\(^4\) Taking these considerations into account, we provide a review of the most widely researched tablet-based cognitive assessments (Table).\(^5\)

\textbf{Brain Health Assessment} (BHA, \url{memory.ucsf.edu/tabcat}). The BHA is a 10-minute 4-task battery developed and vali-
TABLE. TABLET-BASED COGNITIVE ASSESSMENT MEASURES

<table>
<thead>
<tr>
<th>Domains</th>
<th>Peer-reviewed publications</th>
<th>Administration</th>
<th>Automated report</th>
<th>Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Att/EF, Mem, Lang, VS</td>
<td>4</td>
<td>Examiner</td>
<td>10 min, Self</td>
<td>No</td>
</tr>
<tr>
<td>BHA</td>
<td>X</td>
<td>12</td>
<td>8 min, Self</td>
<td>ND</td>
</tr>
<tr>
<td>CAMCI</td>
<td>X</td>
<td>3</td>
<td>100+ Self</td>
<td>6</td>
</tr>
<tr>
<td>CANS-MCI</td>
<td>X</td>
<td>14</td>
<td>100+ Examiner</td>
<td>18</td>
</tr>
<tr>
<td>CANTAB-PAL</td>
<td>X</td>
<td></td>
<td>12-15 min, Self</td>
<td>45</td>
</tr>
<tr>
<td>NIH EXAMINER</td>
<td>X</td>
<td></td>
<td>20-30 min, No</td>
<td>43</td>
</tr>
<tr>
<td>NIHTB-CB</td>
<td>X</td>
<td></td>
<td>31 min, Yes</td>
<td>2</td>
</tr>
</tbody>
</table>

Abbreviations: Att/EF, attention and executive dysfunction; BHA, Brain Health Assessment; CAMCI, Computer Assessment of Mild Cognitive Impairment; CANS-MCI, Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment; CANTAB-PAL, Cambridge Neuropsychological Test Automated Battery Paired Associates Learning; CBB, CogState Brief Battery; Lang, language; Mem, memory; ND, no data; NIH EXAMINER, National Institutes of Health Executive Abilities: Measures and Instruments for Neurobehavioral Evaluation and Research; NIHTB-CB, NIH Toolbox Cognition Battery; VS, visuospatial. For many of these tests, the majority of validation studies have occurred using personal computer (PC) versions as discussed in the text. It is critical that validation studies conducted on a paper-and-pencil or PC version of a test are not mistaken for validation in the tablet. Furthermore, the normative data gathered from paper-and-pencil or PC versions cannot be used for tablet-based adaptations. Scores only.

dated for the detection of mild cognitive impairment (MCI) and mild dementia. The battery is examiner-administered and features automated scoring and interpretation of results in an integrated report. BHA tests measure associative memory, executive functions and processing speed, language, and visuospatial skills and exhibit moderate-to-strong concurrent and neuroanatomic validity. At 85% specificity, the BHA has 100% sensitivity to dementia and 84% sensitivity to MCI in English speakers, and 81% sensitivity to MCI and dementia in Spanish speakers. The BHA has a particular advantage over the Montreal Cognitive Assessment (MoCA) in diagnosing MCI unlikely caused by AD. The BHA has also shown reliability measuring change over time in individuals with no cognitive impairment, MCI, and dementia and is sensitive to longitudinal cognitive decline in individuals who are positive for amyloid β on positron emission tomography (Aβ-PET) but do not have ADRD. A self-administered version is forthcoming.

Computer Assessment of Mild Cognitive Impairment (CAMCI, https://psnet.com/products/camci-research/). The CAMCI is a 20-minute 8-task battery that can be self-administered and features automated scoring and result interpretation. The CAMCI measures attention, episodic memory, executive functions, visuospatial skills, and working memory. Test-retest reliability is good, and the CAMCI has 86% sensitivity and 94% specificity for detecting MCI. Studies support feasibility and practicality of implementing CAMCI in primary care settings.

Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment (CANS-MCI, https://screen-inc.com/). The CANS-MCI is a 30-minute 8-task battery that can be self-administered with automated scoring, result interpretation, and care recommendations. The battery measures episodic memory, executive functions, and language with good test-retest reliability and moderate correlations with standard neuropsychologic measures. In a small study (n=35), the CANS-MCI had 89% sensitivity and 73% specificity to amnestic MCI. Cambridge Neuropsychological Test Automated Battery Paired Associates Learning (CANTAB-PAL, https://www.cambri...). The CANTAB-PAL is a brief standalone portion of the much larger CANTAB battery. Within CANTAB, the CANTAB-PAL has the most validation in ADRD. CANTAB-PAL is an 8-minute nonverbal task of cued recall, widely studied across diagnostically, culturally, and linguistically diverse populations. This measure of visual associative memory has exhibited good reliability over time as well as correlations with conventional neuropsychologic tests and everyday functional measures. CANTAB-PAL showed a sensitivity of 83% and specificity of 82% to MCI. Additionally, CANTAB-PAL has exhibited cross-sectional sensitivity to Aβ-cerebrospinal fluid (CSF) levels in a diagnostically mixed sample.

CogState Brief Battery (CBB, https://www.cogstate.com/clinical-trials/computerized-cognitive-assessment/featured-batteries/). CBB is a 15-minute 4-task battery of attention, episodic memory, speed, and working memory that is self-administered with automated scoring. The PC-version of the CBB has shown test-retest reliability and moderate-to-strong correlations with conventional neuropsychologic tests. CBB measures showed a sensitivity of 41% to 80% and specificity of 85% to 86% to MCI, and sensitivity of 53% to 100% and specificity of 85% to 86% to AD. The battery has also exhibited sensitivity to longitudinal decline in those with Aβ-positive test results with normal cognition and
MCI and mild dementia but was not associated with Aβ- or tau-PET positivity in a cross-sectional study. Notably, CBB is accessed via a web browser, and the majority of validation studies have been conducted on a PC platform.

Computerized Cognitive Composite for Preclinical Alzheimer’s Disease (C3-PAD). Designed as an exploratory endpoint for the A4 trial, the C3-PAD includes CBB measures plus 2 memory paradigms: associative memory and pattern separation. At-home administration of the C3-PAD in 49 adults showed feasibility, alternate-form reliability (Cronbach alpha=0.93), a strong association between in-clinic and at-home versions (r²=0.51), and concurrent validity with standard paper-and-pencil tests. In 3,163 clinically normal adults, in-clinic C3-PAD assessments were moderately correlated with another cognitive composite, and participants who were Aβ-PET had poorer performance statistically, with small effect sizes (unadjusted Cohen d=−0.22, demographic-adjusted d=−0.11).

National Institutes of Health Executive Abilities: Measures and Instruments for Neurobehavioral Evaluation and Research (NIH-EXAMINER, https://memory.ucsf.edu/research-trials/professional/examiner). This 13-task battery, designed and validated on a PC, comprises measures of executive function, social cognition, and behavior; 6 of these executive tasks are considered core and combined into a psychometrically robust composite score. The PC version is sensitive to executive dysfunction in presymptomatic frontotemporal lobar degeneration (FTLD), premanifest Huntington disease, symptomatic FTLD and AD, and associated with brain atrophy. Adaptation for tablet is complete for 4 of the NIH-EXAMINER tasks in TabCAT, and validation studies in ADRD are underway.

NIH Toolbox Cognition Battery (NIHTB-CB, https://www.healthmeasures.net/explore-measurement-systems/niht-toolbox/intro-to-nih-toolbox/cognition). NIHTB-CB is a 31-minute 7-task battery validated across ages and clinical conditions. Designed for PC administration, NIHTB-CB has been adapted for tablets. The battery measures attention, episodic memory, executive functions, processing speed, and language, and the PC version has shown good reliability over time and moderate-to-strong correlations with standard neuropsychologic measures. The PC NIHTB-CB was 85% accurate in discriminating cognitively unimpaired adults mean age 52.5 (SD=12.0) from individuals with MCI and dementia, and poorer performances were associated with decreased hippocampal volumes in adults without cognitive impairment. Most recently, iPad-versions of NIHTB-CB tasks were shown to associate with tau-PET signal. Development and norming of smartphone versions are underway (https://www.mobiletoolbox.org).
Recommendations for Clinicians

The wide range of psychometrically sound computerized cognitive measures allows flexible implementation of these tools into clinical practice depending on the clinical question and needs of the populations served. Most experts agree “ideal” tools for MCI or dementia diagnosis should have affordable cost and reimbursement potential, reliable measurement of multiple cognitive domains across diverse populations, minimization of examiner’s involvement in scoring and interpretation, and sensitivity to longitudinal decline and disease biomarkers.41 Logistically, tools with high potential for scalable implementation should be brief (≤10 minutes), not require a physician for administration (ie, self-administered by the patient or administered by trained clinical staff), and have the functionality to automatically generate an interpretable report with instant electronic medical record (EMR) integration.41 A number of the computerized measures covered in this review meet many of these criteria, and clinicians may consider developing a stepwise workflow to maximize efficient use of digital tools. For example, when the clinical question is whether the patient has a neurodegenerative disorder, the following steps may be used: 1) following the identification of a cognitive concern by the clinician, the patient takes the computerized test either at home (self-administered) or at a quiet place at the clinic (self-administered or administered by trained clinic staff); 2) test results are automatically generated and reported in the EMR, and the clinician reviews the findings; 3) if the results support possible cognitive impairment, the clinician conducts in-depth review of functional and neurologic domains based on the pattern of observed difficulties and completes an evaluation for treatable causes; 4) the clinician integrates all clinical data to determine diagnostic impressions; and 5) the patient retakes the cognitive tests at subsequent visits for monitoring of cognitive functions, particularly if a nonneurodegenerative cause is considered and an intervention is introduced.

Conclusions

Digital assessments may enhance the efficiency of evaluation in neurology and other clinics. Clinicians must choose from a panoply of digital assessments. Selecting an assessment platform requires a careful consideration of the empirical evidence, required device, user qualifications, and the logistics of clinical implementation. Tablet-based assessments are being rapidly validated for use in ADRD populations, and we reviewed several batteries with the greatest empiric support. Smartphone and other BYOD assessments offer several unique benefits, but more validation studies are required before understanding whether they may play a role in clinical evaluations, and until then they should be considered as research tools. We present a potential clinic workflow, acknowledging that greater research on implementation is required.

Adam M. Staffaroni, PhD
Assistant Professor, Department of Neurology
Memory and Aging Center
Weill Institute for Neurosciences
University of California, San Francisco
San Francisco, CA

Elena Tsoy, PhD
Postdoctoral Fellow, Department of Neurology
Memory and Aging Center
Weill Institute for Neurosciences
University of California, San Francisco
San Francisco, CA

Jack Taylor, BS
Research Coordinator, Department of Neurology
Memory and Aging Center
Weill Institute for Neurosciences
University of California, San Francisco
San Francisco, CA

Adam L. Boxer, MD, PhD
Professor, Department of Neurology
Memory and Aging Center
Weill Institute for Neurosciences
University of California, San Francisco
San Francisco, CA

Katherine L. Possin, PhD
Associate Professor in Residence
Department of Neurology, Memory and Aging Center
Weill Institute for Neurosciences
Global Brain Health Institute
University of California, San Francisco
San Francisco, CA

Disclosures
ET, and JT report no disclosures; AMS, ALB, and KLP have disclosures at www.practicalneurology.com