

# Computerised cognitive assessment in patients with traumatic brain injury: an observational study of feasibility and sensitivity relative to established clinical scales



Martina Del Giovane,<sup>a,b,\*</sup> William R. Trender,<sup>b</sup> Maria Bălăeț,<sup>b</sup> Emma-Jane Mallas,<sup>a,b</sup> Amy E. Jolly,<sup>a,b,c</sup> Niall J. Bourke,<sup>a,b,d</sup> Karl Zimmermann,<sup>a,b</sup> Neil S. N. Graham,<sup>a,b</sup> Helen Lai,<sup>a,b</sup> Ethan J. F. Losty,<sup>b</sup> Garazi Araña Oiarbide,<sup>b</sup> Peter J. Hellyer,<sup>e</sup> Irene Faiman,<sup>f</sup> Sarah J. C. Daniels,<sup>a,b</sup> Philippa Batey,<sup>a,g</sup> Matthew Harrison,<sup>a,g</sup> Valentina Giunchiglia,<sup>b</sup> Magdalena A. Kolanko,<sup>a,b</sup> Michael C. B. David,<sup>a,b</sup> Lucia M. Li,<sup>a,b</sup> Célia Demarchi,<sup>a,b</sup> Daniel Friedland,<sup>a,b</sup> David J. Sharp,<sup>a,b</sup> and Adam Hampshire<sup>b</sup>



<sup>a</sup>UK Dementia Research Institute, Care Research & Technology Centre, Imperial College and the University of Surrey, 9th Floor, Sir Michael Uren Hub, 86 Wood Ln, W12 0BZ, London, United Kingdom

<sup>b</sup>Department of Brain Sciences, Imperial College London, London, United Kingdom. Burlington Danes, The Hammersmith Hospital, Du Cane Road, W12 0NN, London, United Kingdom

<sup>c</sup>Department of Neuroinflammation, UCL Queen Square Institute of Neurology, Faculty of Brain Sciences, Queen Square, WC1N 3BG, London, United Kingdom

<sup>d</sup>Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, 16 De Crespigny Park, SE5 8AB, London, United Kingdom

<sup>e</sup>Institute of Psychiatry, Psychology and Neuroscience, King's College London, 16 De Crespigny Park, SE5 8AF, London, United Kingdom

<sup>f</sup>Clinical Neuropsychology Service, St George's University Hospital NHS Foundation Trust, Blackshaw Road, SW17 0QT, London, United Kingdom

<sup>g</sup>The Helix Centre, Imperial College London, and the Royal College of Arts, St. Mary's Hospital, 3rd Floor Paterson Building, 20 South Wharf Road, Paddington, W2 1PE, London, United Kingdom

## Summary

**Background** Online technology could potentially revolutionise how patients are cognitively assessed and monitored. However, it remains unclear whether assessments conducted remotely can match established pen-and-paper neuropsychological tests in terms of sensitivity and specificity.

**Methods** This observational study aimed to optimise an online cognitive assessment for use in traumatic brain injury (TBI) clinics. The tertiary referral clinic in which this tool has been clinically implemented typically sees patients a minimum of 6 months post-injury in the chronic phase. Between March and August 2019, we conducted a cross-group, cross-device and factor analyses at the St. Mary's Hospital TBI clinic and major trauma wards at Imperial College NHS trust and St. George's Hospital in London (UK), to identify a battery of tasks that assess aspects of cognition affected by TBI. Between September 2019 and February 2020, we evaluated the online battery against standard face-to-face neuropsychological tests at the Imperial College London research centre. Canonical Correlation Analysis (CCA) determined the shared variance between the online battery and standard neuropsychological tests. Finally, between October 2020 and December 2021, the tests were integrated into a framework that automatically generates a results report where patients' performance is compared to a large normative dataset. We piloted this as a practical tool to be used under supervised and unsupervised conditions at the St. Mary's Hospital TBI clinic in London (UK).

**Findings** The online assessment discriminated processing-speed, visual-attention, working-memory, and executive-function deficits in TBI. CCA identified two significant modes indicating shared variance with standard neuropsychological tests ( $r = 0.86$ ,  $p < 0.001$  and  $r = 0.81$ ,  $p = 0.02$ ). Sensitivity to cognitive deficits after TBI was evident in the TBI clinic setting under supervised and unsupervised conditions ( $F(15,555) = 3.99$ ;  $p < 0.001$ ).

**Interpretation** Online cognitive assessment of TBI patients is feasible, sensitive, and efficient. When combined with normative sociodemographic models and autogenerated reports, it has the potential to transform cognitive assessment in the healthcare setting.

eClinicalMedicine  
2023;59: 101980

Published Online xxx  
<https://doi.org/10.1016/j.eclinm.2023.101980>

\*Corresponding author. UK Dementia Research Institute (DRI), Care Research & Technology Centre, Imperial College London and the University of Surrey, Imperial College London, 9th floor, Sir Michael Uren Hub, 86 Wood Lane, W12 0BZ, London, United Kingdom.

E-mail address: [m.del-giovane@imperial.ac.uk](mailto:m.del-giovane@imperial.ac.uk) (M. Del Giovane).

**Funding** This work was funded by a National Institute for Health Research (NIHR) Invention for Innovation (i4i) grant awarded to DJS and AH (II-LB-0715-20006).

**Copyright** © 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** Cognition; Deficits; Traumatic brain injury; Online assessment; Attention; Memory; Executive functions

### Research in context

#### Evidence before this study

A PubMed search for 'traumatic brain injury', 'cognition', 'computerised testing', and 'remote testing' conducted on the 15th of January 2023 highlighted growing interest in remote and computerised assessments over the past decade. The heterogeneity of TBI-related cognitive deficits and the debilitating effect these have on functional recovery highlights the benefit that computerised cognitive testing would have in terms of supporting more detailed assessments, early detection and longitudinal tracking of the impairments. However, these assessments should be designed to specifically target the cognitive domains that are relevant to TBI and benchmarked against standard established cognitive tests.

#### Added value of this study

We designed a brief assessment with low sensitivity to the type of device used that could discriminate different TBI-related cognitive impairments. Modelling massive normative data enabled correction of patient scores for multiple demographic variables and enabled a culture fair assessment. The online assessment performed favourably relative to standard neuropsychological measures.

#### Implications of all the available evidence

Our findings suggest that computerised cognitive testing in TBI is feasible and sensitive to deficits across different cognitive domains. Our platform offers a rapid way of screening patients to identify those in greatest need of further clinical investigations and support. Future studies should focus on administering remote computerised testing longitudinally to track recovery trajectories.

### Introduction

Over the past decade, interest has grown in the use of online cognitive assessments in population and clinical research,<sup>1-4</sup> and some neuropsychologists have called for this technology to be incorporated into clinical practice.<sup>5</sup> This is motivated by the potential benefits that online technologies offer as an adjunct to face-to-face assessments, including time savings through automatic deployment, test scoring and data export, analyses based on more detailed behavioural recording at the individual response level, and enhanced engagement through a gamified format (e.g., real-life stimuli, points scoring, rules of playing, progress bar), which allows to simulate the experience of game-playing.<sup>3,6,7</sup>

Surveys of neurological patients, their carers and stakeholders also highlight higher frequency of longitudinal monitoring of cognitive problems as a top research priority.<sup>8</sup> The traditional approach to neuropsychological testing relies on supervised administration of pen-and-paper tests which require costly clinician time in services facing increasingly limited resources and is therefore expensive and burdensome on patients expected to attend face-to-face appointments. Automated online assessment is ideal for this purpose because it is deployable in peoples' homes, potentially under unsupervised conditions and via computer, tablet and smartphone devices that they already own. The

automated generation of unique stimuli can also reduce practice effects, enabling repeat assessments.<sup>2,7</sup> Importantly, computerised and remote cognitive tests can offer a cost-effective way of screening patients for those most needing of the more extensive, formal specialist neuropsychological assessment.

However, the potential of online cognitive assessment in clinical practice is still in its infancy. Partly this reflects generic challenges of translating research technology into digital healthcare solutions, such as ensuring robustness and data validity across operating systems and devices. Regarding cognitive assessment, such challenges are substantial because differences in screen sizes and response interface latencies can confound results which primarily rely on performance speed and accuracy. There are also more idiosyncratic challenges; most notably, human cognition is complex, spanning multiple domains. Patients have widely differing levels and types of impairment. Tests sensitive to subtle executive problems may be unfeasible for more pronounced memory problems, whereas tests for the latter may be insensitive to the former. Neuropsychologists therefore call upon a repertoire of carefully designed measures to fit different assessment purposes. Some patients may not tolerate lengthy assessments or understand task instructions. Critically, the performance of cognitive tasks covaries with a variety of

variables in the normal population, e.g., age, education level and first language.<sup>7</sup> Therefore, for utility in healthcare, an online assessment battery must be optimised to discriminate the cognitive domains that are relevant to the target population whilst being accessible in terms of instructions and duration. It should be ‘culture fair’—accounting for normal performance variation with population factors. Such technology also must result in sufficient engagement and compliance and be evaluated against the current available tests for sensitivity and precision in discriminating affected domains.

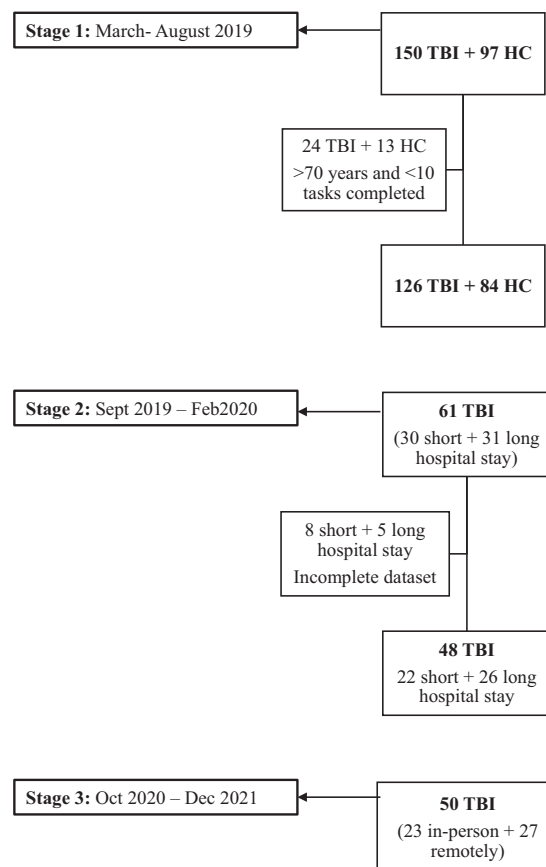
The management of traumatic brain injury (TBI) exemplifies the challenges and potential of online assessments. TBI is a leading cause of death and disability worldwide, affecting ~10 million individuals annually.<sup>9</sup> The management of chronic cognitive symptoms after TBI is particularly difficult, partly because of limited resources, but also the intrinsic heterogeneity of this population.<sup>10</sup> Along with physical and psychiatric sequelae,<sup>11,12</sup> approximately 65% of moderate-severe TBI patients show long-term cognitive deficits.<sup>13</sup> These deficits are heterogeneous at baseline. Furthermore, different patterns of damage to brain networks correlate with deficits in different cognitive domains and predict differential responsiveness to pharmacotherapy and neurostimulation.<sup>14–16</sup> Impairments in memory, executive functions, and attention can persist for many years and have significant functional consequences.<sup>13,17,18</sup> Deficits also follow diverse trajectories.<sup>19</sup> There can be substantial recovery from the acute to chronic phase, but patients are likely to experience accelerated age-related decline,<sup>20</sup> including increased dementia incidence.<sup>21</sup> It is therefore important to identify cognitive problems early to plan rapid interventions, i.e., before emerging problems interfere with life trajectories. Online assessment could provide detailed diagnostic information for this purpose cheaply, including longitudinal tracking of deficits.

Here, we sought to optimise and validate an online assessment battery for patients with moderate-severe TBI that could be used not only for research, but also in practical healthcare. Specifically, to help clinicians and rehabilitation therapists (i) prioritise and allocate resources by identifying patients most in need of formal neuropsychological assessments, and (ii) identify those who need follow-up based on remote monitoring of their progression over time.

## Methods

### Study summary

This is an observational study structured in three stages (Fig. 1). First, we evaluated a superset of available computerised tasks that were based on classic neuropsychological paradigms; these were modified to be brief, engaging, and suitable for online unsupervised



**Fig. 1:** Recruitment flowchart. Legend: Flowchart illustrating the number of participants recruited and included in the analysis at each stage of the study.

deployment. We collected data from 126 patients and 87 controls to identify an optimal subset of tasks that maximised sensitivity to the TBI group, minimised sensitivity to the device used, and measured different cognitive domains. We then undertook a second stage to investigate how this optimal battery correlated with two standard face-to-face neuropsychological tests, the Montreal Cognitive Assessment (MoCA) and Repeatable Battery for the Assessment of Neuropsychological Status (RBANS).<sup>22,23</sup> We measured the relative sensitivity of the online and face-to-face assessment to TBI severity, defined by length of hospital stay. Finally, the online assessment was embedded into a novel clinical tool that calculated each individual’s scores via using a very large normative dataset and correcting for demographic characteristics. This tool automatically generates a clinician-facing results report that is immediately available. It was piloted for use in a regional referral TBI outpatient clinic, and sensitivity to deficits was evaluated under both unsupervised remote and supervised clinic conditions. The tertiary referral clinic we have been using this tool in, tends to see patients a minimum of 6

months post-injury. The mean time since injury of the patients in the 3 stages of this study was ~3 years. Based on this study, we propose that the tool is applicable at this stage of the injury – after the acute cognitive effects such as post-traumatic amnesia have worn off, and clinicians are able to assess the residual sequelae that need rehabilitating in the chronic phase (i.e., months to years). We also envision the tool to be used by clinicians to then monitor patients over time and identify those who need rehabilitation and follow-up appointments.

However, we did also include a small number of patients who were <1 month post-injury at the time of testing. This was important to validate the fact that the tool was accessible for patients in this stage of their recovery. Further work would be required to validate the utility of this tool in aiding clinical decision making in the more acute phase of recovery following TBI.

#### Online cognitive assessment tasks

A large library of tasks is available on the Cognitron platform. These tasks are designed to measure diverse cognitive domains, including with substantial redundancy (e.g., Tower of London and Blocks both measure the same aspects of spatial planning but in different ways). From this library, 23 tasks were selected that (a) covered multiple cognitive domains, (b) included aspects of cognition reported to be affected by moderate-severe TBI, and (c) that were considered to be accessible, brief, understandable and engaging by a panel of cognitive neuroscientists, clinical neuropsychologists, neurologists, other stakeholders, and patient consultants who had suffered a TBI.

The cognitive platform and task designs have been reported extensively elsewhere and in [Supplementary Table S1](#).<sup>4,7,24,25</sup> Briefly, each task was based on a classic paradigm used for measuring different aspects of cognition, including working memory, attention, reasoning, planning, processing speed, and executive functions. These paradigms, typically pen-and-paper and delivered under controlled conditions, had been reworked into HTML5 with JavaScript tasks that were (a) brief, (b) minimally gamified to maximise engagement, (c) testing non-language-dependent abilities, (d) using stimuli automatically generated with balancing of difficulty dimensions, minimising learning effect, and (e) deployable via practically any modern PC, tablet, or smartphone web browser.

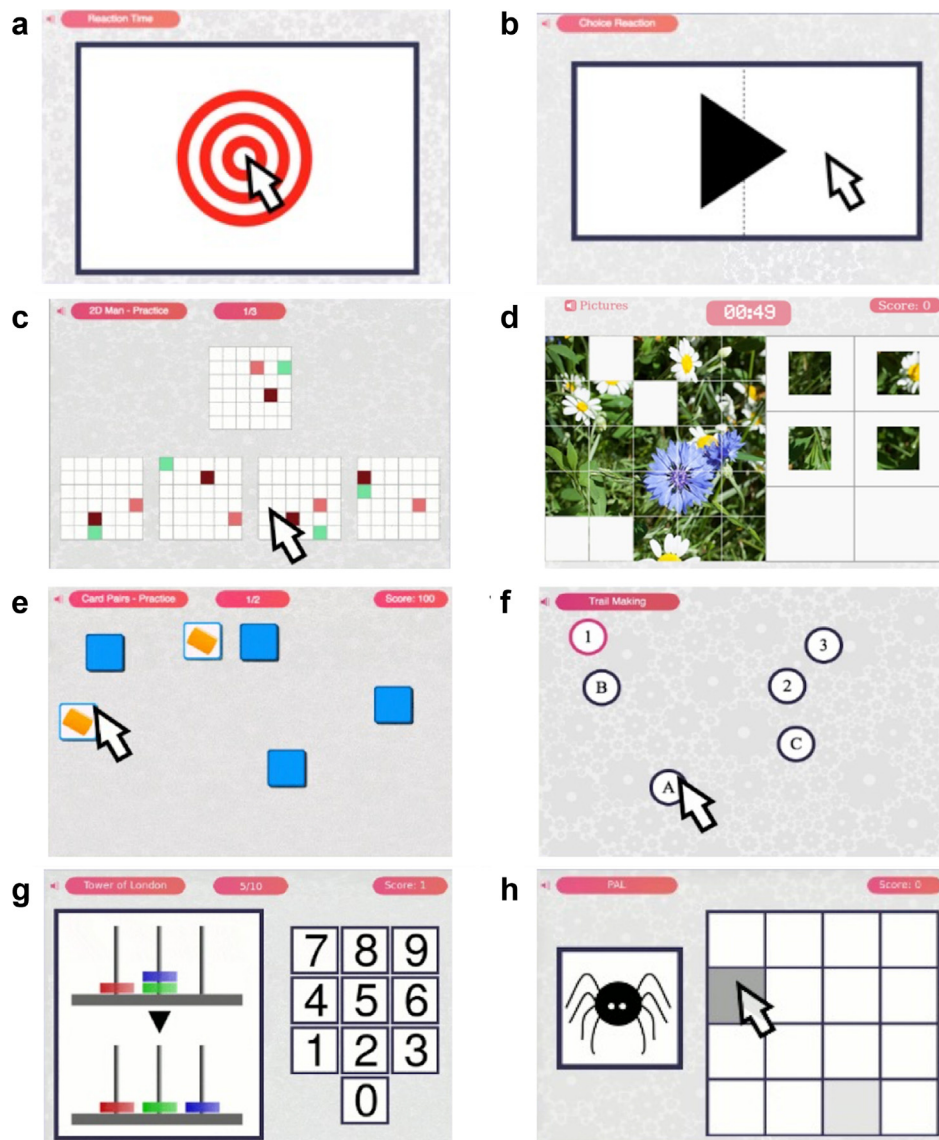
#### Data collection

Recruitment took place across inpatient major trauma wards and at an outpatient TBI clinic at St Mary's Hospital in London, which is a tertiary referral clinic and receives referrals from across the whole country. Control participants were recruited through friends and relatives accompanying patients, and via the Join Dementia Research website, funded by the Department of Health and delivered by the National Institute for

Health and Care Research in partnership with Alzheimer Scotland, Alzheimer's Research UK and Alzheimer's Society ([www.joindementiaresearch.nihr.ac.uk](http://www.joindementiaresearch.nihr.ac.uk)). Exclusion criteria for all participants included a prior history of neurological or psychiatric diseases and the presence of significant substance abuse. People with significant pre-existing visual and motor impairments which would influence their ability to interact with an electronic device were also excluded. This study received ethical approval by the NHS Health Research Authority, London - Camberwell St Giles Research Ethics Committee (REC reference: 17/LO/2066, IRAS: 230221). All participants gave written and/or electronic informed consent.

This study had three stages ([Fig. 1](#)). Stage 1 took place between March and August 2019 at the St. Mary's Hospital TBI clinic, and in major trauma wards at Imperial College Healthcare NHS trust and St. Georges Hospital in London (UK). 247 individuals aged 18–80 years were recruited, 97 healthy controls, and 150 participants with moderate-severe TBI based on the Mayo Classification system,<sup>26</sup> which we consider an appropriate measure of TBI long-term outcomes and prognosis among the currently available tools.<sup>26,27</sup> Individuals 70 years or younger, with 10 or more completed cognitive tasks and complete demographic information were included in the analysis. Participants older than 70 were excluded due to lack of sufficient numbers to consider them representative of the general population, especially in light of the increased risk of cognitive impairments due to neurodegenerative processes in this age group.<sup>28</sup> 23 cognitive tasks are too many for patients to tolerate in a single session. Therefore, during Stage 1, participants were asked to complete one of four different batteries ([Supplementary Table S1](#)) on a tablet device at the research centre. To minimise order effects, the batteries were pseudo-randomised, cycling through A–D for each successive participant. During this stage, participants were also instructed to download an app on their home devices and asked to complete the batteries remotely. Participants were free to choose the time between batteries. To minimise practice effects, only data from the first-time subjects completed each task were analysed. Instances of participants failing to complete the other batteries at home as requested, would therefore have led to a random selection of tasks not being completed.

Stage 2 took place between September 2019 and February 2020. A further sample of 61 moderate-severe TBI patients recruited from the St Mary's Hospital TBI clinic was assessed at the Imperial College research centre using a single battery of 8 tasks ([Supplementary Table S1](#) and [Fig. 2](#)) derived from optimisation analyses of data from Stage 1. An index of injury severity used duration of stay in hospital, splitting the cohort into two, 30 participants being hospitalised for ≤7 days and 31 participants being hospitalised for >7 days.



**Fig. 2:** Schematic of the 8 tasks included in the refined battery of Stage 2. Legend: (A) simple reaction time, (B) choice reaction time, (C) 2D manipulations, (D) picture completion, (E) card pairs, (F) trail making, (G) Tower of London, (H) paired associates learning.

Participants also completed standard neuropsychological scales—the MoCA and RBANS. Only participants with a complete dataset were retained for analysis.

During Stage 3, between October 2020 and December 2021, the 8-task battery was administered with a third group of 50 moderate-severe TBI patients who attended the outpatient TBI clinic at St. Mary's Hospital in London. Participants completed the computerised tasks either in-person or remotely, which divided the cohort into two groups of 23 and 27 patients, respectively.

Before administering the tasks in the clinic, designers in Imperial College London's Helix Centre improved the accessibility and usability of the webpages

of the tasks and result reports. The work was structured as a 'design sprint', first mapping the user experience, and then interviewing representative users and observing them using the webpages to guide improvements. Designers also worked with people affected by dementia and clinicians to inform this process. After interviews and observations, they created static mock-ups to specify improvements to the pages' design.

### Statistical analysis

#### *Stage 1—optimisation of tasks for cognitive battery*

Unpaired t-tests and Chi-squared tests were performed to characterise differences in age, gender, handedness,



first language, education, and the device used in the patient and control groups. Multiple measures can be derived from the same task; therefore, our analysed dataset comprised 39 scores accounting for the 23 tasks (Supplementary Table S1). Rank inverse normalisation was applied to each of the vectors of scores to handle outliers, in order to ensure that they were normally distributed and to place them on a normalised scale.

To investigate which task scores were sensitive to TBI and insensitive to device, general linear models were conducted for each task score using all available data as the dependent variables, and group (TBI or control) and testing device as independent variables. 74 participants undertook the tasks with a PC, 55 with a Apple mobile device, 46 with a Samsung mobile device and 35 with a different Android mobile device, forming a factor with four conditions. Age, gender and education were included as covariates in the models. Tasks with low sensitivity to group as quantified by low effect size ( $|t| < 1.9$ ) were excluded. Task with scores showing high sensitivity to device ( $p < 0.05$ ) were also removed.

Principal component analysis (PCA) with orthogonal varimax rotation was then applied to the remaining data to reduce the number of available variables and group the computerised task scores into cognitive components. The number of components was defined based on the Kaiser convention (eigenvalue  $> 1$ ). Two tasks having high and discrete loading for each component (i.e., having a high influence on one specific component and low influence on the others) were retained, while the remaining were excluded. This way, tasks that capture the greatest variance for each cognitive component remained, with the cognitive battery assessing multiple aspects of cognition. The resulting brief battery, comprising 8 tasks, was further evaluated in Stages 2–3.

#### *Stage 2—validation of the refined battery to standard neuropsychological tests and assessment of sensitivity to injury severity*

Scores derived from the online tasks and standard neuropsychological tests were winsorised at 3 standard deviations (STDs) to account for extreme outliers. Four composite measures were calculated from the online tasks to reflect the four cognitive components identified using PCA in Stage 1. This was achieved by taking the first unrotated factor for the pair of tasks associated with each cognitive component. A global composite score for each patient was also calculated as the first unrotated factor across all eight vectors of task scores.

Patients were categorised as having either short ( $\leq 7$  days) or long hospital stays ( $> 7$  days) as a proxy for injury severity. General linear models (GLMs) were conducted to determine whether individual tasks, as well as the composite and total scores for both the computerised battery and the standard neuropsychological scales, were sensitive to injury severity, age, first language, and education.

To assess the relationship between the computerised battery and standard neuropsychological scales, a CCA was performed, whereby data from the computerised tasks formed the first set of variables (U) and data from the standard neuropsychological battery (RBANS & MoCA) formed the second (V). In CCA, the shared variance between the two sets of variables is defined as a correlational mode. To determine which tasks from each dataset contributed to each mode, correlation analyses were performed between each variable and the statistically significant modes that were identified.

#### *Stage 3—evaluating tasks sensitivity using large scale normative data*

Data were winsorised at three STDs to remove outliers. Accuracy and reaction time (RT) scores of the 8-task battery were converted into ‘deviation from expected’ (DfE) scores, which describe the extent that patients deviate from the expected score of a cognitively healthy person with the same demographic characteristics. Specifically, linear models were trained on a large online-collected normative dataset to predict each task score based on a detailed combination of population factors including age, age<sup>2</sup>, gender, handedness, education, occupation status, ethnicity, residency, and first language. The trained models were then applied to the participants’ demographics to derive their expected scores. For each task, the difference between the participant’s observed and expected accuracy and RT scores was then computed and divided by the control population STD, thereby deriving the DfE score.

Normative data were collected via the Great British Intelligence Test, a collaborative project with the BBC2 Horizon started in December 2019.<sup>4</sup> The number of normative datapoints available varied across tasks (mean = 131,736.82, min = 2283, max = 392,855). For RT, thresholds indicative of lack of compliance (e.g., for 2D Manipulations, spending more than 88,059.5 ms or less than 169 ms to provide an answer) were applied to the normative and participants datasets, but no non-compliant participants were identified. One-way ANOVA was used to check for the sensitivity of different online tasks to TBI-related cognitive deficits. Then, t-tests against zero defining whether patient performance was significantly lower than the normative group were executed on the DfE scores to test the sensitivity of each online task. Finally, to determine whether the administration of the tasks in-person vs. remotely substantially affected performance, two-way mixed ANOVA was performed with the 17 online task measures as the within-subject factor and the modality of administration as the between-subject factor.

#### **Role of the funding source**

The funder of the study had no role in the design of the study, data collection, data analysis, interpretation, or writing of the report. All authors had full access to all

data and accept responsibility for the decision to submit for publication.

**Results**

**Stage 1—stepwise selection of optimal task sub-set**

Participants undertook different combinations of tasks, producing a sparse dataset. Therefore, each task was analysed individually for sensitivity to device and TBI using all available data. In total, 126 TBI patients (91 males, mean age = 43.16, STD = 12.99, mean time since the injury = 140 weeks, median = 28, IQR = 157) and 84 healthy controls (32 males, mean age = 34.60, STD = 13.45) were included in the analyses. Details regarding the number of participants per task and participants’ demographics are displayed in [Supplementary Tables S2 and S3](#), respectively.

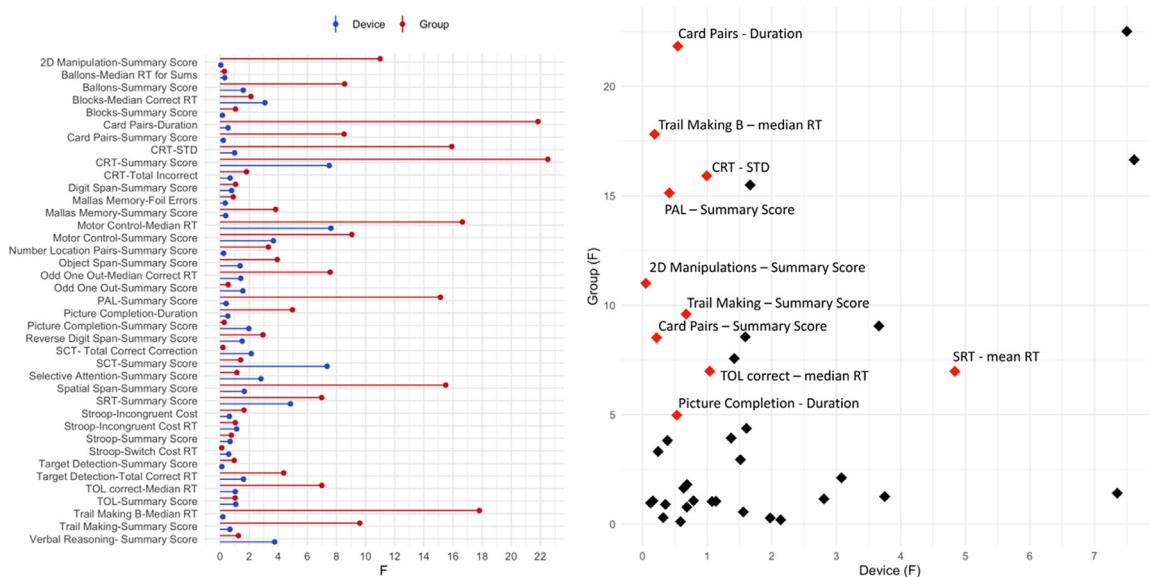
After accounting for socio-demographics factors including age, sex, education level and device, 15 tasks were selected as they showed substantial sensitivity to the TBI vs. control contrast ( $F > 4.0$ , all  $p < 0.05$ ) ([Supplementary Table S4](#) and [Fig. 3](#)). Notably, the  $F$  values for device and TBI sensitivity were not significantly correlated across tasks, indicating that the observed variability did not have a common basis ( $r = 0.24$ ,  $p > 0.05$ ).

PCA of the most sensitive scores from each of the 15 TBI-sensitive tasks identified four components with eigenvalues  $>1$ , with the first explaining  $\sim 37\%$  of the variance, and the remaining three collectively explaining  $\sim 62\%$ . The identified components captured different

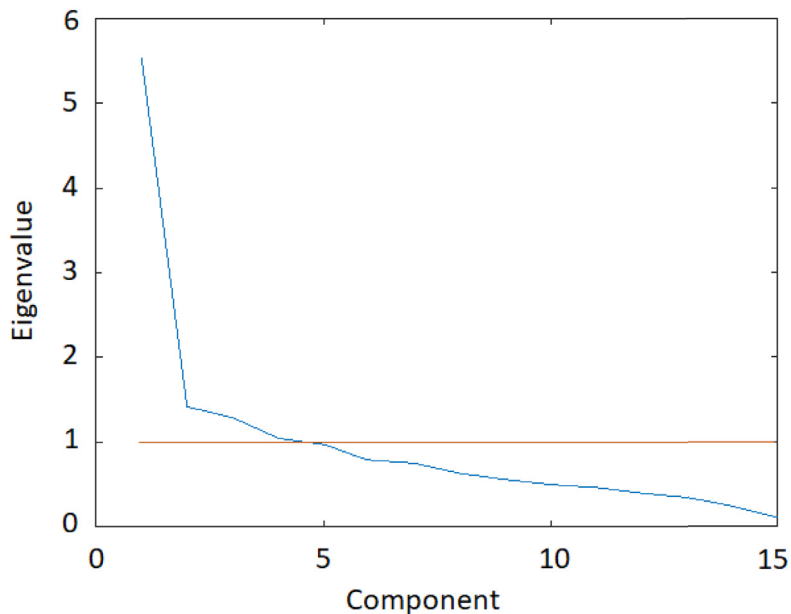
cognitive domains and were processing speed, working memory, visual attention, and executive functions ([Fig. 4](#)). Two tasks that had high loading onto each of the four components, had low loadings onto other components and low sensitivity to device ( $F < 1.4$ ), were selected to derive 4 composite scores. These tasks were Card Pairs, Trail Making, Paired Associate Learning, 2D Manipulations, Tower of London (TOL), Picture Completion, Choice Reaction Time (CRT), and the Simple Reaction Task (SRT) ([Fig. 2](#)). Notably, mean/median RTs for the latter two tasks were sensitive to device, but they were retained for two reasons. First, they provide rapid response time measures that are prominent in the literature on TBI. Second, the standard deviation measure provided a correlated proxy for median response time having low device sensitivity. This 8-task assessment battery was used in Stage 2 and 3 because it enabled to measure the impact of TBI on cognition at the level of (a) individual tasks, (b) pairs of tasks that measure different broad cognitive factors, and (c) overall composite performance.

**Stage 2—benchmarking sensitivity of online tasks to TBI severity against established neuropsychological scales**

Exclusion of participants without a complete dataset due to extensive cognitive impairments or lack of tolerance of the testing session led to a further sample of 48 moderate-severe TBI patients (39 males, mean age = 45.23, STD = 14.12, mean time since the injury = 224 weeks, median = 64, IQR = 209), 22



**Fig. 3:** Sensitivity of the computerised tasks to TBI-related cognitive deficits and the device used. Legend: **Left.** Tasks sensitivity to group (red) and device (blue), as indicated by the  $F$  values derived from the general linear models. **Right.** Computerised tasks on a scatter plot of TBI (Y axis) vs. device (X axis) sensitivity. In red are the tasks that were selected based on high sensitivity to the TBI group ( $F > 4.0$ ) and low sensitivity to device ( $F < 1.4$ ).



Processing Speed	Visual Attention	Working Memory	Executive Functions	
0.5	0.07	0.05	0.01	<b>Choice Reaction Time</b>
0.47	0.08	0.11	0.14	<b>Simple Reaction Time</b>
0.46	0	0.02	0.03	Motor Control
0.36	0	0.18	0.09	Balloons
0.26	0.61	0.02	0.05	Odd One Out
0.04	0.49	0.04	0.09	<b>Picture Completion</b>
0.19	0.44	0.16	0.2	<b>2D manipulations</b>
0.06	0.1	0.53	0.2	Object Span
0.09	0.09	0.45	0.11	Spatial Span
0.02	0.05	0.43	0.11	<b>Trail Making</b>
0.06	0.11	0.4	0.16	<b>Paired Associate Learning</b>
0.08	0.16	0.01	0.54	Target Detection
0.05	0.31	0.08	0.51	<b>Tower of London</b>
0.03	0.11	0.25	0.44	<b>Card Pairs</b>
0.24	0.12	0.14	0.31	Mallas Memory

**Fig. 4:** Factor analysis conducted to refine the 8-task computerised battery and generate component scores. Legend: **Upper.** Eigenvalues derived from principal component analysis of cognitive tasks performed by TBI patients. **Lower.** Loadings of the computerised tasks onto the different cognitive components derived from principal component analysis. In bold are the tasks that were retained for each component.



participants being hospitalised for  $\leq 7$  days and 26 participants being hospitalised for  $> 7$  days. The Glasgow Outcome Scale -Extended (GOSE) mean score was 5.71, STD = 1.20. Participants' demographics are illustrated in [Supplementary Table S5](#). Component scores were extracted by applying PCA to each of the four pairs of tasks, producing composite scores for each patient as follows: 1) Processing speed (CRT-STD and SRT-mean RT), 2) Working Memory (Card Pairs-Summary Score and PAL-Summary Score), 3) Visual attention (2D Manipulations-mean RT and Picture Completion-Duration), and 4) Executive functions (TOL correct-mean RT, Trail Making B-median RT, and Trail Making-Summary Score). An overall composite score was also extracted across all tasks.

As expected, the GLMs indicated sensitivity to age for some of the final battery task scores: CRT-STD ( $F(1,44) = 4.88$ ;  $p = 0.03$ ), PAL-summary score ( $F(1,44) = 4.88$ ;  $p = 0.03$ ), Card Pairs-summary score ( $F(1,44) = 18.48$ ;  $p < 0.001$ ), 2D Manipulation-summary score ( $F(1,44) = 8.74$ ;  $p = 0.01$ ), Picture Completion-Duration ( $F(1,44) = 14.34$ ;  $p < 0.001$ ), Trail Making B-mean RT ( $F(1,44) = 8.02$ ;  $p = 0.01$ ), and Trail Making-Summary Score ( $F(1,44) = 10.34$ ;  $p = 0.02$ ). Sensitivity to age was also shown by the majority of the composite scores, including Working Memory ( $F(1,44) = 18.50$ ;  $p < 0.001$ ), Visual Attention ( $F(1,44) = 14.34$ ;  $p < 0.001$ ), Executive Functions ( $F(1,44) = 8.56$ ;  $p = 0.01$ ), and the total composite score ( $F(1,44) = 13.91$ ;  $p < 0.001$ ). Regarding the classic neuropsychological scales, the MoCA delayed recall subtest was sensitive to age ( $F(1,44) = 8.33$ ;  $p = 0.01$ ) and age<sup>2</sup> ( $F(1,44) = 5.68$ ;  $p = 0.02$ ). Sensitivity to age<sup>2</sup> was also shown by the RBANS delayed memory subtest ( $F(1,44) = 5.32$ ;  $p = 0.03$ ). Only the MoCA naming subtest was sensitive to first language ( $F(1,44) = 7.34$ ;  $p = 0.01$ ).

Sensitivity to TBI severity ([Fig. 5](#)) was shown by SRT-STD ( $F(1,46) = 5.83$ ;  $p = 0.02$ ), Picture Completion-Duration ( $F(1,46) = 8.10$ ;  $p = 0.01$ ), Trail Making B-median RT ( $F(1,46) = 8.18$ ;  $p = 0.01$ ), Trail Making-summary score ( $F(1,46) = 7.18$ ;  $p = 0.01$ ), Processing Speed ( $F(1,46) = 5.63$ ;  $p = 0.02$ ), Visual Attention ( $F(1,46) = 8.10$ ;  $p = 0.01$ ), Executive Functions ( $F(1,46) = 8.51$ ;  $p = 0.01$ ), and the total composite score ( $F(1,46) = 9.39$ ;  $p < 0.01$ ). Among the standard neuropsychological scales and sub-scales, only the RBANS visuospatial subscale was able to discriminate between the lengths of hospital stay ( $F(1,46) = 16.62$ ;  $p < 0.001$ ).

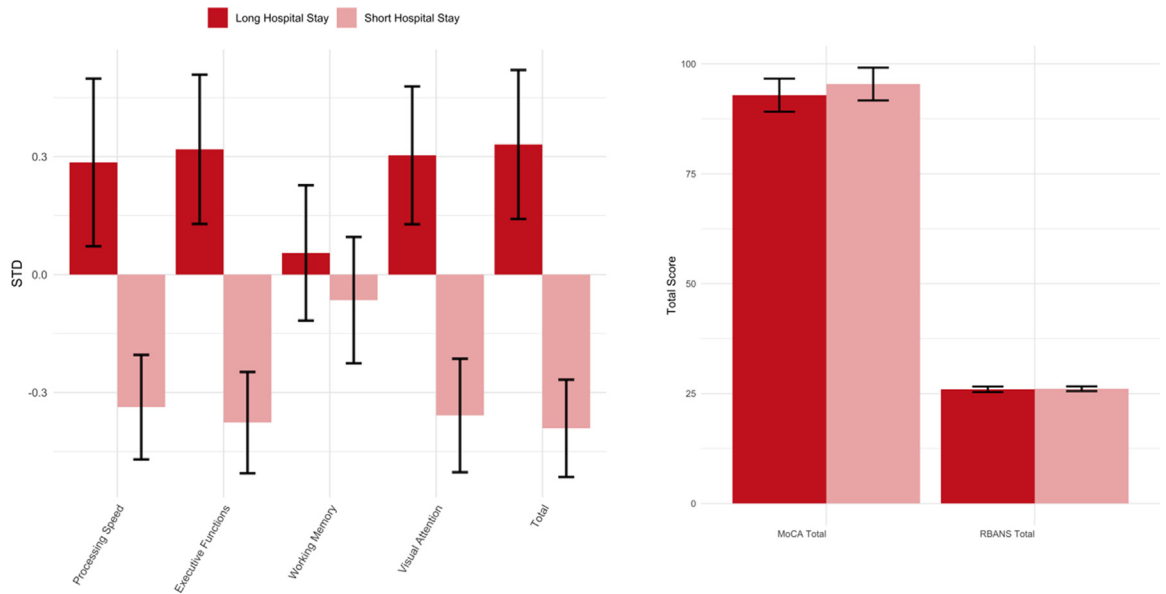
Sensitivity to educational level was found for the CRT-STD ( $F(2,45) = 4.62$ ;  $p = 0.02$ ), SRT-STD ( $F(2,45) = 4.00$ ;  $p = 0.03$ ), TOL correct-median RT ( $F(2,45) = 5.64$ ;  $p = 0.001$ ), Picture Completion-Duration ( $F(2,45) = 7.04$ ;  $p < 0.01$ ), Trail Making A-mean RT ( $F(2,45) = 3.15$ ;  $p = 0.05$ ), and Trail making B-mean RT ( $F(2,45) = 10.33$ ;  $p < 0.001$ ), Processing speed ( $F(2,45) = 4.18$ ;  $p = 0.02$ ), Visual Attention ( $F(2,45) = 7.04$ ;  $p < 0.01$ ), Executive Functions

( $F(2,45) = 10.01$ ;  $p < 0.001$ ), and the total composite score ( $F(2,45) = 7.80$ ;  $p = 0.001$ ). Among the standard neuropsychological scales, those sensitive to education were the RBANS immediate memory ( $F(2,45) = 4.13$ ;  $p = 0.02$ ), visuospatial ( $F(2,45) = 3.19$ ;  $p = 0.05$ ), language ( $F(2,45) = 4.53$ ;  $p = 0.02$ ), and attention ( $F(2,45) = 6.42$ ;  $p < 0.01$ ) subscales, the RBANS total scale ( $F(2,45) = 6.28$ ;  $p < 0.01$ ), the MoCA visuospatial/executive ( $F(2,45) = 8.87$ ;  $p < 0.001$ ), naming ( $F(2,45) = 14.18$ ;  $p < 0.001$ ), and abstraction ( $F(2,45) = 9.41$ ;  $p < 0.001$ ) subscales, and the MoCA Total score ( $F(2,45) = 8.94$ ;  $p < 0.001$ ).

CCA identified two significant modes, M1 ( $r = 0.86$ ,  $p < 0.001$ ) and M2 ( $r = 0.81$ ,  $p = 0.02$ ), indicating shared variance between the final battery and the standard neuropsychological scales ([Fig. 6](#)). Back projected  $r$  values indicating loading for each of the computerised tests (U) and each of the standard neuropsychological scores (V) onto the two significant modes are shown in [Tables 1 and 2](#), respectively.

### Stage 3—modelling a large normative dataset to correct scores for demographic variables and embedding these in a novel framework that autogenerates clinician-facing reports

Data for 17 task scores across the 8 tasks of the final battery were analysed for 50 moderate-severe TBI patients (39 males, mean age = 42.6, STD = 16.41, mean time since the injury = 103 weeks, median = 54, IQR = 86). Participants' demographics are illustrated in [Supplementary Table S6](#). One-way ANOVA indicated a significant effect of the online tasks on participants' performance ( $F(15,570) = 4.03$ ;  $p < 0.001$ ) ([Fig. 7](#)). T-tests of the DfE scores calculated for all the 17 task scores confirmed that participants scored significantly lower than what is expected in the control population on the Card Pairs-summary score (estimate =  $-0.69$ ,  $t = -3.88$ ,  $p < 0.001$ ) and duration (estimate =  $-1.37$ ,  $t = -5.85$ ,  $p < 0.001$ ), CRT-summary score (estimate =  $-0.87$ ,  $t = -5.65$ ,  $p < 0.001$ ) and STD (estimate =  $-0.61$ ,  $t = -3.49$ ,  $p = 0.001$ ), manipulations-summary score (estimate =  $-0.75$ ,  $t = -5.93$ ,  $p < 0.001$ ) and median correct RT (estimate =  $-0.67$ ,  $t = -3.30$ ,  $p < 0.01$ ), PAL-summary score (estimate =  $-0.52$ ,  $t = -3.52$ ,  $p = 0.001$ ) and median RT (estimate =  $-0.94$ ,  $t = -4.80$ ,  $p < 0.001$ ), Picture Completion-summary score (estimate =  $-0.58$ ,  $t = -3.18$ ,  $p < 0.01$ ) and duration (estimate =  $-1.45$ ,  $t = -5.36$ ,  $p < 0.001$ ), SRT-summary score (estimate =  $-1.57$ ,  $t = -6.96$ ,  $p < 0.001$ ) and STD (estimate =  $-0.81$ ,  $t = -5.02$ ,  $p < 0.001$ ), TOL-summary score (estimate =  $-0.44$ ,  $t = -2.87$ ,  $p = 0.01$ ) and median correct RT (estimate =  $-0.69$ ,  $t = -2.45$ ,  $p < 0.05$ ), Trail Making-summary score (estimate =  $-1.32$ ,  $t = -4.85$ ,  $p < 0.001$ ) and part B-median RT (estimate =  $-1.64$ ,  $t = -7.32$ ,  $p < 0.001$ ). Results from the two-way mixed ANOVA showed a significant effect of the tasks ( $F(15,555) = 3.99$ ;  $p < 0.001$ ), suggesting



**Fig. 5:** Sensitivity to length of hospital stay for both the computerised tasks and standard neuropsychological scales. Legend: **Left.** Bar plot showing sensitivity to length of hospital stay for the composite scores derived from the computerised tasks. Scores are indicated in standard deviation units (STD). **Right.** Bar plot showing sensitivity to length of hospital stay for the standard neuropsychological scales. Red = long hospital stay, pink = short hospital stay. Error bars represent the standard errors.

these were sensitive to TBI-related deficits. The modality of administration of the test (in-person vs. remotely) showed no significant effect on participants' performance and no interaction with the tasks ( $p > 0.05$ ).

Through the work conducted by the Helix design centre in conjunction with the software developers, the online tasks were embedded in a user-friendly interface designed to make the testing experience more approachable for patients. Patients' performance was then summarised in an optimally accessible results page for clinicians (Fig. 8). As the tasks were designed also to be performed unsupervised, three categories of 'effort tests' were included in the clinician result page to indicate which tasks might have not been performed to the best of the patients' abilities. These consisted of (1) a self-perceived effort test based on whether the participant felt distracted and on what kind of environment they were in, (2) a loss of focus test, indicated by the number of times the patient clicked onto another browser tab during the tests, (3) a specific effort test for each task derived from thresholds applied on tasks-specific output measures (e.g., for the SRT, for not responding on >50% of trials).

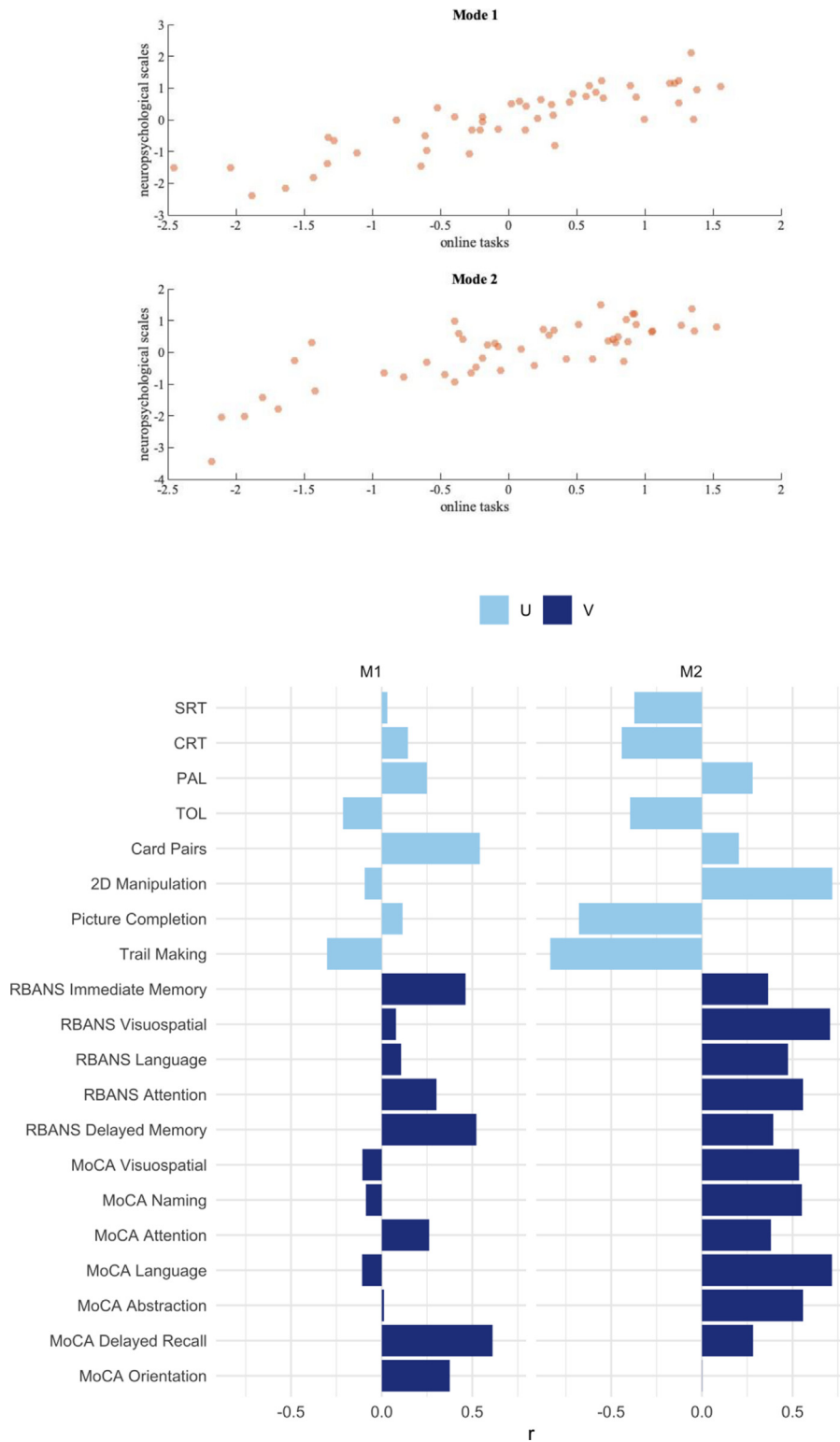
### Case studies

Two case studies illustrate the use of the online cognitive tasks at the multi-disciplinary TBI clinic at St. Mary's Hospital. The online cognitive tasks were used as

an initial screening of patients' cognitive abilities to define the need for referral for formal neuropsychological assessment.

### Case 1

A 22-year-old right-handed male who sustained a moderate-severe TBI when hit by a car whilst cycling. He underwent a left craniotomy for a traumatic extradural haematoma involving the left parietal and temporal lobes. Following admissions to intensive care, he was stepped down to the major trauma unit and then to the neurological rehabilitation unit before discharge. He was then assessed in the TBI clinic via telephone appointment. Our online cognitive assessment was performed remotely, and his performance in all tests was in the average to high-average range relative to the large normative dataset. The exception was the Paired Associates Learning task, which was in the low-average range (24th percentile). At formal neuropsychological assessment, he reported mild, subjective cognitive difficulties, in particular relating to attention. Memory, processing speed, and executive functions were tested, and were all found to be reflective of a largely intact cognitive functioning. This was consistent with the on-line cognitive assessment, but our tool was able to detect an area of cognitive impairment missed by the standard neuropsychological assessment. In terms of treatment, the assessing neuropsychologist advised cognitive and fatigue management strategies.



**Fig. 6:** Results of the Canonical Correlation Analysis. Legend: **Upper.** Correlation of the computerised tests with the standard neuropsychological scales for the two significant modes (M1 and M2). **Lower.** Back projected r values indicating contribution of the computerised tests (U - light blue) and the MoCA and RBANS subtests (V - dark blue) to the two significant modes derived from the canonical correlation analysis.

Domain U	M1	M2
CRT	0.145	-0.442
PAL	0.250	0.280
SRT	0.031	-0.372
TOL	-0.213	-0.396
Card Pairs	0.542	0.204
2D manipulation	-0.094	0.719
Picture completion	0.115	-0.678
Trail making	-0.301	-0.836

CRT = choice reaction time, PAL = paired associate learning, SRT = simple reaction time, TOL = Tower of London.

**Table 1:** Back projected r values indicating loading for each of the computerised tests (Domain U) onto the two significant modes (M1 and M2) identified by the canonical correlation analysis and showing shared variance between the computerised tasks and the standard neuropsychological scales.

Case 2

A 36-year-old right-handed male who sustained a moderate-severe TBI following a fall. The CT brain scan showed a right extra-axial haemorrhage, frontal and left temporal haemorrhagic contusions, fracture through the left middle ear extending to the skull base, and a right hairline skull base fracture. He attended the TBI clinic in-person, and his performance on the online cognitive tasks pointed to impairments in Motor Control (1st percentile), Picture Completion (5th percentile), SRT (17th percentile), and CRT (19th percentile). During formal neuropsychological assessment, he reported difficulties with memory, processing speed, and executive functions. The assessment indicated executive functions performance was within expected levels, although there was collateral information pointing to

Domain V	M1	M2
RBANS immediate memory	0.462	0.366
RBANS visuospatial	0.079	0.707
RBANS language	0.107	0.475
RBANS attention	0.302	0.558
RBANS delayed memory	0.522	0.394
MoCA visuospatial	-0.107	0.536
MoCA naming	-0.088	0.551
MoCA attention	0.262	0.381
MoCA language	-0.109	0.717
MoCA abstraction	0.012	0.558
MoCA delayed recall	0.611	0.282
MoCA orientation	0.375	0.002

RBANS = repeatable battery for the assessment of neuropsychological status, MoCA = montreal cognitive assessment.

**Table 2:** Back projected r values indicating loading for each of the neuropsychological subscales (Domain V) onto the two significant modes (M1 and M2) identified by the canonical correlation analysis and showing shared variance between the computerised tasks and the standard neuropsychological scales.

significant everyday executive difficulties. The online cognitive testing highlighted processing speed difficulties that were not picked up in the formal neuropsychological assessment. The clinical recommendation was for community-based neurorehabilitation including neuropsychology and neuro-occupational therapy.

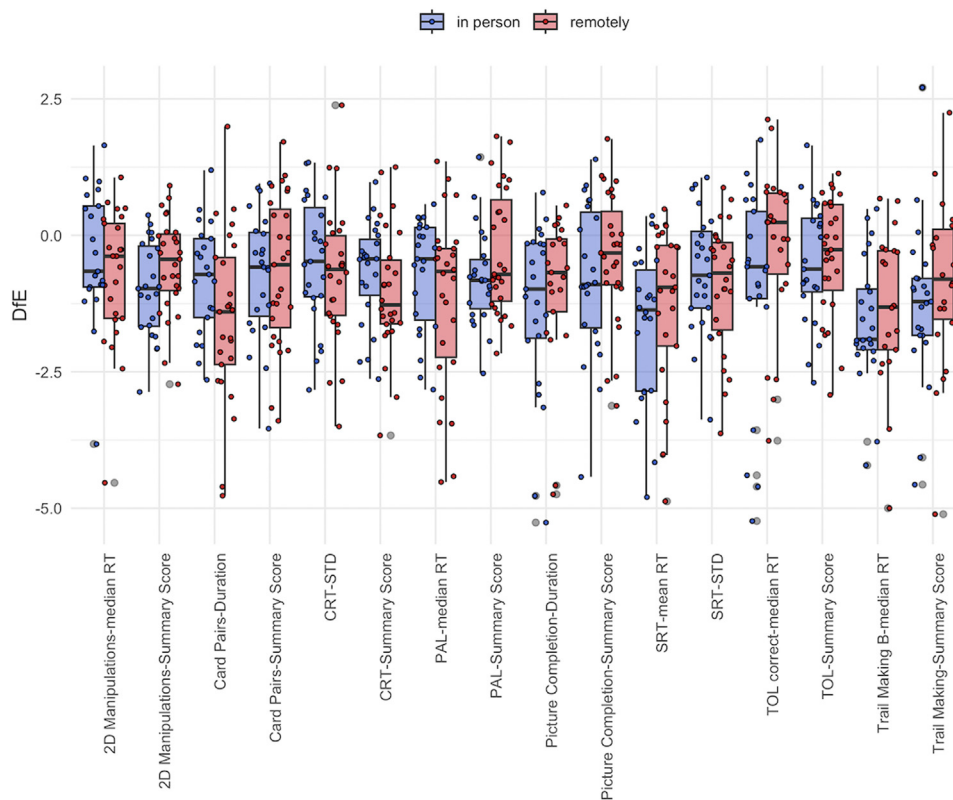
Discussion

The results of our study suggest that online cognitive assessment in patients with moderate-severe TBI is feasible and comparable to standard neuropsychological tests. It was successfully trialled in the authors' TBI clinic helping to guide further referral for formal neuropsychological assessment and neurorehabilitation for patients. We identified a brief battery of online cognitive tasks that can detect specific TBI-related cognitive impairments across different devices. The use of online tools for cognitive assessments reduces the need for lengthy assessments and scoring sessions by trained professionals and can be utilised by different healthcare workers including neuropsychology and occupational therapists, as well as neurologists. It also provides a more granular cognitive assessment while recording multiple measures simultaneously – something unfeasible in traditional assessments performed by a single clinician. Factor analysis on the selected tasks allowed us to define a brief battery that optimises sensitivity to deficits while avoiding redundancy. This, together with the gamified format, optimises engagement, and performance. By selecting tasks with a low sensitivity to device, we have justified the possibility of having patients complete these tasks at home, saving time spent to travel to the hospital and optimising resources.

Stage 2 found that our online cognitive tasks and the derived composite scores share variance with classic neuropsychological tests, suggesting that they can measure some of the same constructs, and that they can have superior sensitivity to injury severity.

Alternative online cognitive batteries, such as the CNS-Vital Signs, Immediate Post-Concussion Assessment and Cognitive Testing (impACT), Automated Neuropsychological Assessment Metric (ANAM) and Axon/CogState/CogSport (CogState), previously proved to be sensitive to cognitive effects of concussion and showed significant correlations to standard neuropsychological scales.<sup>29,30</sup> However, their diagnostic utility seems unclear in the post-acute phase, and correlations to traditional neuropsychological tests range from weak to moderate. Moreover, there has been limited research on the use of these tests in moderate-severe TBI.

In Stage 3 we showed the sensitivity of a tool that autogenerates reports with immediate readout, supporting the potential of our tool in remotely tracking the trajectory of patients' cognitive impairments. As it is known that cognitive performance varies greatly based on demographic factors such as age, gender, or



**Fig. 7:** Comparison of patients' performance on the battery of computerised tasks completed in person and remotely. Legend: Boxplot displaying comparison of DfE scores between patients who completed the computerised task in-person (blue) and remotely (red).

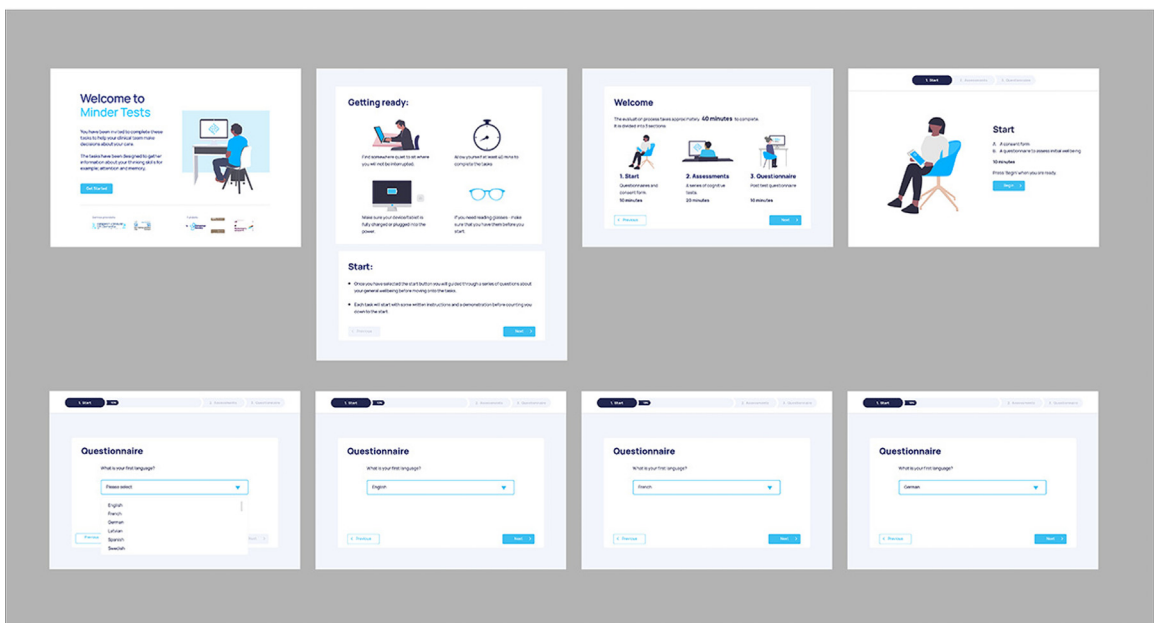
education level, and that these can even act as markers of injury severity,<sup>31</sup> it is important to consider these when evaluating cognitive performance. Using our novel approach, we have achieved this systematically and precisely by leveraging a uniquely large normative dataset to allow a more 'culture fair' scoring procedure.

Despite the benefits of such a tool, we acknowledge some limitations. The usability of the tool is reduced in people without access to an online device, although the test may be applied in clinic prior to the appointment with the clinician. For people without the requisite computer literacy, the tool is also unlikely to be effective. This has been pre-emptively mitigated through very simple task designs, with animated instructions, and is also expected to diminish in future generations. Moreover, for data collected remotely, we cannot be certain of the exact conditions under which the tasks were completed, and the degree of supervision provided.

Our findings should also be considered in light of certain methodological limitations. First, patients completed different tasks in Stage 1, resulting in different datasets for each task. However, the minimum number of patients completing each task was equal to 27, which is comparable to previous similar studies.<sup>32</sup> Second, some gender and age differences were found

in Stage 1, although this is reflective of the epidemiology of TBI, which occurs more in males.<sup>33</sup> Third, we did not seek to retain patients' participation longitudinally across all three stages of the study. In Stage 3, we do not have data from the same patients performing the tasks both in-person and remotely, for comparison. This arose partly due to the COVID-19 pandemic, which prevented many patients from attending hospital appointments in-person. However, this provided us with the opportunity to demonstrate a core benefit of online testing, as we were able to acquire cognitive data for the purpose of clinical assessment in patients who remained at home. Future studies could employ a crossover design to investigate the potential effect of remote vs. in-person testing further. Also, we did not include tests that measure aspects of memory at longer temporal scales, impulsivity, or emotion processing/control. Such tasks are available, as are many others, within the same online task framework, and they will be evaluated for their potential value as an extension to this TBI-optimised battery. The variability of the time since the injury across stages also represents a limitation, with average time since the injury higher at later stages of the study, which may complicate the results due to recovery and decline trajectories.<sup>34</sup> Critically though, the tests were





**Fig. 8:** Result page generated to summarise patients’ performance with a clinician-facing report. Legend: **Top.** Example of the result report automatically generated following completion of the online cognitive tasks. Just 3 of the 8 tasks are shown, with others on subsequent pages. **Bottom.** Design sketches of an enhanced user experience, that emphasises preparing the user for the exercises to increase the likelihood of successful completion.

consistently sensitive across the three stages of the study, meaning there is likely reliability across a range of post-injury timepoints. This may also limit the generalisability of our findings to the TBI population. However, the fact that St Mary’s Hospital clinic is a tertiary referral clinic which receives referrals from across the country, and that patients were also recruited from major trauma wards that were not limited to the Imperial College NHS trust, contribute to the representativeness of our sample of the national TBI population.

It is also important to note that the computerised tasks were presented in a set sequence, with no option to skip them. Therefore, missing data were more likely to relate to tasks towards the end of the battery. This may lead to underrepresenting participants with higher level of impairments and the associated difficulty in completing the tasks. However, this is an unavoidable phenomenon of all cognitive assessment techniques, and failure to complete the assessment can be considered a useful outcome measure in itself. Finally, although the length of Hospital stay represents

a surrogate marker of injury severity,<sup>35</sup> we recognise that it can be extended by the presence of polytrauma or other causes of delayed discharge, and consequently, participants' TBI severity may have been overestimated here. However, as our data were collected retrospectively, other indicators of injury severity, such as length of post-traumatic amnesia, would have been unreliable.

Despite the limitations mentioned, this study shows the feasibility of using online cognitive testing in TBI. Cognitive impairments represent a major complaint and cause of disability in TBI.<sup>31</sup> However, the nature and magnitude vary greatly, dictating the need for more precise and targeted assessments that are validated in this specific population and that pick up multiple aspects of cognition. Our tool offers the potential for improved cost-efficacy ratio compared to traditional assessments. It allows rapid and relatively cheap assessment, so it can provide a standardised way of screening patients to determine those needing referral for further investigations and neurorehabilitation. This is done with increased precision in measuring multiple behavioural measures, reduced risk of learning effect and a culture-fair approach to the scoring process. The possibility of administering the tool remotely, without the need for a clinician present, and on any device, confers clear financial benefit. This is potentially applicable to other clinical groups and has been previously used and reported by our research group in the context of people with multiple sclerosis,<sup>36</sup> dementia,<sup>25</sup> and people who contracted COVID-19 and consequently received extensive care.<sup>4,25</sup>

Notably, our assessment battery measures cognitive domains that overlap with those included in the neuropsychological tests recommended by the National Institute of Neurological Disorders and Stroke (NINDS) common data elements to assess cognition in adults with moderate-severe TBI, and the established Cambridge Neuropsychological Test Automated Battery (CANTAB). However, our software is specifically designed for repeat remote deployments under unsupervised conditions and can be normed against a comprehensive set of population variables collected from hundreds of thousands of members of the public. Additionally, the tool produces an associated, clinician-facing report created by software developers and design experts, providing it with unique clinical utility.

Rehabilitation therapists and neuropsychologists can leverage this tool to conduct longitudinal assessments with added depth, without requiring multiple clinical appointments. Limited resources can then be diverted to those most in need. Moreover, moderate-severe TBI has long-term sequelae, and there are now questions as to whether associated cognitive deficits are progressive, underlying the importance of close monitoring over time.<sup>31</sup> Future studies should focus on administering these tasks longitudinally to characterise impairment.

Hopefully these results can direct bespoke, personalised rehabilitation plans for more effective recovery, in a way that is not possible with current assessment methods.

Parties interested in using this technology should contact the senior author.

#### Contributors

All authors had full access to the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

AH and MDG had accessed and verified the data.

AH, DS, DF, CD, MDG, NB, and NG contributed to the conceptualisation of the study.

AH and DS were responsible for the investigation, administration, and supervision of the study.

AH and DS obtained the funding which supported the study.

AH was responsible for the methodology used in the study.

WT, AH, PH, and MB contributed to the development of the software used in this study.

MDG, IF, NG, KZ, LL, EM, AJ, MK, GO, and HL contributed to the data collection.

MDG, AH, VG, WT, MB, EL, and HL contributed to the data curation.

AH, MDG, VG, AJ, MB, and NB contributed to the analysis and interpretation of the data.

MDG and AH contributed to the data visualisation.

AH, MDG, MD, and DF contributed to the drafting of the manuscript.

MH and PB contributed to the graphic design of the tests.

All authors provided critical revision of the manuscript for important intellectual content.

#### Data sharing statement

Requests for data should be directed to the corresponding authors. Data will be available upon reasonable request.

#### Declaration of interests

Prof. Hampshire reports grants from NIHR, and is owner/director of H2 Cognitive Designs Ltd and Future Cognition Ltd, which produce online assessment technology and provide online survey data collection for third parties. Dr. Hellyer is co-founder and co-owner of H2 Cognitive Designs Ltd and H2 Cognitive Designs Ltd. Prof. Sharp reports grants from NIHR Invention for innovation (i4i); NIHR Imperial Biomedical Research Centre; Advance Brain Health Clinic & Aligned Research programme (Football Association (FA); Advance Brain Health Clinic & Aligned Research programme; Rugby Football Union & Premiership Rugby; ADVANCE Charity; Armed Services Trauma Rehabilitation Outcome Study (ADVANCE); Action Medical Research; UK Dementia Research Institute Care & Technology Centre; Era-Net-Neuron European collaborative award; St Mary's Development Charity; Royal British Legion Centre for Blast Studies. Prof. Sharp also took part in the Early Detection of Neurodegenerative diseases (EDoN) project and the Rugby Football Union concussion advisory board. Mr. Trender reports payments from Imperial NHS trust, Imperial College London, the Engineering and Physical Sciences Research Council, and H2 Cognitive Designs Ltd. Dr. David reports grants from the Medical Research Council (Clinical Research Fellowship). Miss Bălăeț reports payments from the Medical Research Council Doctoral Training Programme. Dr. Graham reports grants from Alzheimer's Research UK (Clinical Research Fellowship). Mr. Friedland reports payments for expert testimony outside this study. Dr. Li reports grants from LML: UK DRI Pilot Grant 2021 and is trustee of Headway West London. Miss Helen Lai reports payments from Imperial College London (President's PhD Scholarship) and the UK Dementia Research Institute Care & Technology Centre. Dr. Bourke reports payments for medicolegal imaging analysis outside this work and payments as part of the BRAN travel grant. All the other authors have nothing to disclose.

**Acknowledgements**

This work was funded by a National Institute for Health Research (NIHR) Invention for Innovation (i4i) grant awarded to DJS and AH (II-LB-0715-20006).

**Appendix A. Supplementary data**

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.eclinm.2023.101980>.

**References**

- Bauer RM, Iverson GL, Cernich AN, Binder LM, Ruff RM, Naugle RI. Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Arch Clin Neuropsychol*. 2012;27(3):362–373.
- Brooker H, Williams G, Hampshire A, et al. FLAME: a computerized neuropsychological composite for trials in early dementia. *Alzheimers Dement*. 2020;12(1):e12098.
- Hampshire A, Ballard C, Williams G. Computerized neuropsychological tests undertaken on digital platforms are cost effective, achieve high engagement, distinguish and are highly sensitive to longitudinal change: data from the PROTECT and GBIT studies. *Alzheimers Dement*. 2020;16(S10):e041122.
- Hampshire A, Trender W, Chamberlain SR, et al. Cognitive deficits in people who have recovered from COVID-19. *EclinicalMedicine*. 2021;39:101044.
- Parsons TD, McMahan T, Kane R. Practice parameters facilitating adoption of advanced technologies for enhancing neuropsychological assessment paradigms. *Clin Neuropsychol*. 2018;32(1):16–41.
- Feenstra HEM, Vermeulen IE, Murre JMJ, Schagen SB. Online cognition: factors facilitating reliable online neuropsychological test results. *Clin Neuropsychol*. 2017;31(1):59–84.
- Hampshire A, Highfield RR, Parkin BL, Owen AM. Fractionating human intelligence. *Neuron*. 2012;76(6):1225–1237.
- Deane KHO, Flaherty H, Daley DJ, et al. Priority setting partnership to identify the top 10 research priorities for the management of Parkinson's disease. *BMJ Open*. 2014;4(12):e006434.
- Dewan MC, Rattani A, Gupta S, et al. Estimating the global incidence of traumatic brain injury. *J Neurosurg*. 2019;130(4):1080–1097.
- Covington NV, Duff MC. Heterogeneity is a hallmark of traumatic brain injury, not a limitation: a new perspective on study design in rehabilitation research. *Am J Speech Lang Pathol*. 2021;30(2S):974–985.
- Reeves RR, Panguluri RL. Neuropsychiatric complications of traumatic brain injury. *J Psychosoc Nurs Ment Health Serv*. 2011;49(3):42–50.
- Ponsford J, Sloan S, Snow P. *Traumatic brain injury: rehabilitation for everyday adaptive living*. 2nd ed. Psychology Press; 2017.
- Draper K, Ponsford J. Cognitive functioning ten years following traumatic brain injury and rehabilitation. *Neuropsychology*. 2008;22(5):618–625.
- Fregni F, Li S, Zaninotto A, Santana Neville I, Paiva W, Nunn D. Clinical utility of brain stimulation modalities following traumatic brain injury: current evidence. *Neuropsychiatric Dis Treat*. 2015;11:1573–1586.
- Jolly AE, Scott GT, Sharp DJ, Hampshire AH. Distinct patterns of structural damage underlie working memory and reasoning deficits after traumatic brain injury. *Brain*. 2020;143(3):1158–1176.
- Jenkins PO, de Simoni S, Bourke NJ, et al. Stratifying drug treatment of cognitive impairments after traumatic brain injury using neuroimaging. *Brain*. 2019;142(8):2367–2379.
- Gorgoraptis N, Zaw-Linn J, Feeney C, et al. Cognitive impairment and health-related quality of life following traumatic brain injury. *NeuroRehabilitation*. 2019;44(3):321–331.
- Rassovsky Y, Satz P, Alfano MS, et al. Functional outcome in TBI I: neuropsychological, emotional, and behavioral mediators. *J Clin Exp Neuropsychol*. 2006;28(4):567–580.
- Millis SR, Rosenthal M, Novack TA, et al. Long-term neuropsychological outcome after traumatic brain injury. *J Head Trauma Rehabil*. 2001;16(4):343–355.
- Cole JH, Leech R, Sharp DJ. Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Ann Neurol*. 2015;77(4):571–581.
- Fann JR, Ribe AR, Pedersen HS, et al. Long-term risk of dementia among people with traumatic brain injury in Denmark: a population-based observational cohort study. *Lancet Psychiatr*. 2018;5(5):424–431.
- Nasreddine ZS, Phillips NA, Bédirian V, et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695–699.
- Randolph C, Tierney MC, Mohr E, Chase TN. The repeatable battery for the assessment of neuropsychological status (RBANS): preliminary clinical validity. *J Clin Exp Neuropsychol*. 1998;20(3):310–319.
- Soreq E, Violante IR, Daws RE, Hampshire A. Neuroimaging evidence for a network sampling theory of individual differences in human intelligence test performance. *Nat Commun*. 2021;12(1):2072.
- Hampshire A, Chatfield DA, MPhil AM, et al. Multivariate profile and acute-phase correlates of cognitive deficits in a COVID-19 hospitalised cohort. *EclinicalMedicine*. 2022;47:101417.
- Malec JF, Brown AW, Leibson CL, et al. The Mayo classification system for traumatic brain injury severity. *J Neurotrauma*. 2007;24(9):1417–1424.
- Friedland DP. Improving the classification of traumatic brain injury: the Mayo classification system for traumatic brain injury severity. *J Spine*. 2013;S4:005. <https://doi.org/10.4172/2165-7939.S4-005>.
- Prince M, Knapp M, Guerchet M, et al. Dementia UK update. Available from: [https://www.alzheimers.org.uk/sites/default/files/migrate/downloads/dementia\\_uk\\_update.pdf](https://www.alzheimers.org.uk/sites/default/files/migrate/downloads/dementia_uk_update.pdf) [cited 2023 Mar 10].
- Arriex JP, Cole WR, Ahrens AP. A review of the validity of computerized neurocognitive assessment tools in mild traumatic brain injury assessment. *Concussion*. 2017;2(1):CNC31.
- Gualtieri CT, Johnson LG. A computerized test battery sensitive to mild and severe brain injury. *Medscape J Med*. 2008;10(4):90.
- Ruet A, Bayen E, Jourdan C, et al. A detailed overview of long-term outcomes in severe traumatic brain injury eight years post-injury. *Front Neurol*. 2019;10:120.
- Richter S, Stevenson S, Newman T, et al. Study design features associated with patient attrition in studies of traumatic brain injury: a systematic review. *J Neurotrauma*. 2020;37(17):1845–1853.
- Mollayeva T, Mollayeva S, Colantonio A. Traumatic brain injury: sex, gender and intersecting vulnerabilities. *Nat Rev Neurol*. 2018;14(12):711–722.
- Lennon MJ, Brooker H, Creese B, et al. Lifetime traumatic brain injury and cognitive domain deficits in late life: the PROTECT-TBI cohort study. *J Neurotrauma*. 2023. <https://doi.org/10.1089/neu.2022.0360>.
- Yue JK, Krishnan N, Chyall L, et al. Predictors of extreme hospital length of stay after traumatic brain injury. *World Neurosurg*. 2022;167:e998–e1005.
- Lerede A, Middleton R, Rodgers J, et al. 056 Online assessment and monitoring of cognitive decline in neurological conditions. *J Neurol Neurosurg Psychiatry*. 2022;93(9):e2.3.