Check for updates

Original Research Article

Health Informatics Journal



Health Informatics Journal
1–19
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14604582221087890
journals.sagepub.com/home/jhi

Evaluation of a clinical decision support tool for matching cancer patients to clinical trials using simulation-based research

Clarissa Gardner
Institute of Global Health Innovation, Imperial College London, London, UK

Jack Halligan
Institute of Global Health Innovation, Imperial College London, London, UK

Gianluca Fontana
Institute of Global Health Innovation, Imperial College London, London, UK

Roberto Fernandez Crespo
Institute of Global Health Innovation, Imperial College London, London, UK

Matthew Prime
Roche Information Solutions, F. Hoffmann-La Roche AG, Basel, Switzerland

Chaohui Guo
Roche Information Solutions, F. Hoffmann-La Roche AG, Basel, Switzerland

Okan Ekinci
Roche Information Solutions, F. Hoffmann-La Roche AG, Basel, Switzerland

Saira Ghafur
Institute of Global Health Innovation, Imperial College London, London, UK

Ara Darzi
Institute of Global Health Innovation, Imperial College London, London, UK

Corresponding author:
Clarissa Gardner, Institute of Global Health Innovation, Imperial College London, Level 10, QEQM Building, St Marys Hospital, Paddington Basin, London W2 1NY, UK.
Email: clarissa.gardner@imperial.ac.uk



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

Abstract

There is a growing need for alternative methodologies to evaluate digital health solutions in a short timeframe and at relatively low cost. Simulation-based research (SBR) methods have been proposed as an alternative methodology for evaluating digital health solutions; however, few studies have described the applicability of SBR methods to evaluate such solutions. This study used SBR to evaluate the feasibility and user experience of a clinical decision support (CDS) tool used for matching cancer patients to clinical trials. Twenty-five clinicians and research staff were recruited to match 10 synthetic patient cases to clinical trials using both the CDS tool and publicly available online trial databases. Participants were significantly more likely to report having sufficient time ($p = 0.020$) and to require less mental effort ($p = 0.001$) to complete trial matching with the CDS tool. Participants required less time for trial matching using the CDS tool, but the difference was not significant ($p = 0.093$). Most participants reported that they had sufficient guidance to participate in the simulations (96%). This study demonstrates the use of SBR methods is a feasible approach to evaluate digital health solutions and to collect valuable user feedback without the need for implementation in clinical practice. Further research is required to demonstrate the feasibility of using SBR to conduct remote evaluations of digital health solutions.

Keywords

Digital health, simulation-based research, evidence generation, clinical decision support tools

Introduction

Clinical trials in oncology are essential for the scientific advancement of treatments available to existing and future cancer patients.[1] They require the enrollment of suitable study participants to assess the clinical efficacy and safety of novel treatments. They also offer an opportunity for patients with cancer to try alternative new treatments provided they meet the strict, and sometimes lengthy, eligibility criteria outlined in trial protocols.[2] Anecdotal evidence from practitioners in the field suggests that manually screening patients for eligibility criteria is a time-consuming activity responsible to medical and non-medical practitioners, such as consultants, oncologists, and research nurses. It requires a review of patients' records, including any previous medical history such as pre-existing conditions, medications, test results, as well as holistic information such as patient mentality to determine a patient's eligibility. In a busy clinical environment, this could lead to some degree of bias or a lack of equal opportunity for patients.[2]

NAVIFY Clinical Trial Match (NAVIFY CTM) is a clinical decision support (CDS) tool within NAVIFY Tumor Board (NAVIFY TB) (Roche Information Solutions Inc., Santa Clara, CA), a cloud-based workflow solution that is used to prepare for and conduct cancer multidisciplinary team meetings. CDS tools aim to improve healthcare delivery by enhancing medical decisions with targeted clinical knowledge, patient information, and other health information.[3] CDS tools are primarily used at the point-of-care for the clinician to combine their knowledge with information or suggestions provided by the tool. NAVIFY CTM has been developed to streamline the process of searching for and identifying clinical trials by using patient-specific data such as cancer stage, genomic alterations, and the treating institution's postcode to perform an automated search for suitable clinical trials across 21 international trial databases. The tool has been developed to address the challenges of under-recruitment in cancer clinical trials, as well as the decision-making

challenges faced by clinicians seeking to identify eligible patients from electronic health record data.[4–6]

Evaluating digital health solutions

Digital health solutions (DHSs) have been shown to address healthcare challenges in various disease areas and healthcare contexts.[7,8] However, to justify their implementation in clinical settings and ensure public trust and patient safety, DHSs must be evaluated for usability, acceptability, clinical effectiveness, and cost-effectiveness.[9–11] Randomized controlled trials (RCTs) have traditionally been recognized as the gold standard to demonstrate the clinical effectiveness of novel interventions and therapeutics.[12] However, few innovators have sought to evaluate DHSs using RCTs, possibly due to the fast-paced nature of product development and the arduous process of setting up an RCT in a clinical setting.[12,13] The relatively high costs and resources associated with conducting an RCT reduces the return on investment and is therefore not a sustainable option for small and medium-sized enterprises (SMEs) or products in the early stage of development.[13] Furthermore, recruiting participants to RCTs and pilot studies to evaluate DHSs in real-life clinical settings comes with challenges such as patient data privacy and security concerns, the commitment required to participate in longitudinal studies for self-management interventions, and a lack of endorsement for trial participation from treating clinicians.[13–16]

The regulation and publication of guidance to evaluate DHSs is outpaced by the rapidly evolving digital health landscape and the increasingly diverse types of solutions entering the market.[17,18] In addition, there is a lack of clarity regarding what evidence generation methodologies are appropriate to evaluate solutions in the early stages of product development. Evidence generation methodologies are required that can generate high-quality insights in a short timeframe and at relatively low cost.[13,19] A perspective piece published in npj Digital Medicine on the challenges of evaluating DHSs proposed alternative methodologies that facilitate safe and responsible growth across the sector, such as the use of simulation-based research (SBR) methods.[13]

Simulation-based research methods

Simulation is described as the replication of real-life events in an interactive manner, and has been applied in research and the trial of new healthcare interventions.[20] Simulation has also been widely applied to healthcare in the context of education and clinical rehearsal, and allows healthcare providers the opportunity to learn new protocols and procedures in a safe and controlled environment.[21–26] SBR in the context of healthcare involves replicating clinical scenarios, with the aim of eliciting reactions and behaviors from healthcare professionals that are as close as possible to how they would react in a real situation.[27] The use of SBR methods could support faster evidence generation of the usability, acceptability, and effectiveness of DHSs without involving real patients and/or real patient data. Such simulations could take place outside of a clinical setting, providing opportunities to obtain behavioral and cognitive measures that would otherwise be impractical to gather in routine practice.[13] SBR methods have been used to evaluate non-digital healthcare interventions[28,29]; however, few studies have reported the use of SBR methods to evaluate DHSs and the results are seldom reported.[13,30–32]

To assess the feasibility of using SBR methods to evaluate DHSs, we sought to apply these methods to evaluate NAVIFY CTM, which primarily supports existing clinical services yet has no direct impact on patient outcomes. According to guidelines published by the National Institute for Health and Care Excellence (NICE), such solutions would be functionally classified as “Tier 1” and

require the minimum standards of evidence for effectiveness.[19] The intended outcome would be to demonstrate that SBR methods could be used in the first instance to generate the minimum level of evidence required to support the adoption or larger-scale evaluation of any DHS.

The aims of this study is to assess the feasibility of using SBR methods to evaluate DHSs. The objective of this study is to use SBR methods to compare the efficiency, quality and cognitive burden of using NAVIFY CTM to match patients to cancer clinical trials with the use of publicly available online trial databases.

Methods

Study design and participants

We performed an experimental within-subject, non-blinded study to evaluate NAVIFY CTM using simulation-based research (SBR) methods. Participants were approached as previous participants of digital health studies conducted at the university or via the research team's professional networks. Snowball sampling was used in this study to reach additional participants, which is a commonly used recruitment technique when engaging individuals in a specific population.[33] As per the inclusion/exclusion criteria for the study, participants were required to have experience matching patients to cancer clinical trials in the UK and must not have had prior experience using NAVIFY CTM. Roles eligible to participate in the study included consultant oncologists (i.e., board certified attendings), specialist registrars in oncology (i.e., senior residents), research nurses, and clinical trial practitioners.

Participation in this study was pseudonymous and ethics approval for this study was sought from and approved by the Research Governance & Integrity Team at the university (reference 19IC5457). Informed consent was obtained from participants at the beginning of each simulation session. The study was funded by F. Hoffman-La Roche Ltd, and participants were reimbursed for their time and travel expenses.

Synthetic patient case development

The research team identified 10 actively recruiting Investigational Medicinal Product (IMP) trials that represented a mix of trials for small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). A total of 11 synthetic patient cases were created to match each of these trials (two of the cases were eligible for the same trial). The cases were developed by two consultant thoracic oncologists, one consultant histopathologist, and one consultant interventional radiologist, and included clinical details (e.g., past medical history and social history), pathology reports, and radiology reports with accompanying images. The patient cases were added to NAVIFY TB to facilitate trial matching using the NAVIFY CTM application. Online trial databases that are known to be in use in clinical practice were used as the study comparator. These databases are publicly accessible and are designed to enable members of the public, clinicians, and research staff to identify clinical trials. To facilitate trial matching using online trial databases, clinical details and reports were provided on a laptop computer in PDF format with images provided in JPG format.

Data collection

Participants were instructed to assume the role of a clinician or research staff member at St Mary's Hospital in London who is preparing to see 10 patients at an outpatient clinic. Participants were advised to find at least one suitable clinical trial for each patient, with the intention of discussing trial enrollment, and to treat the session as they would a real-life clinical scenario.

Table 1. The order of procedures during the simulation sessions.

Activity	Duration (mins)
Informed consent and introductions	5
Review of simulation session guide	10
Practice stroop color and word test (SCWT)	<1
SCWT 1[a]	<5
Trial matching exercise 1[b]	60
SCWT 2[a]	<5
BREAK	10
SCWT 3[a]	<5
Trial matching exercise 2[b]	60
SCWT 4[a]	<5
BREAK	10
Post-simulation survey	10
Total	*~180*

[a]Participants completed the stroop color and word test (SCWT) before and after each trial matching exercise. This was used to infer participants' mental agility before and after each trial matching exercise and estimate the cognitive burden associated with matching the synthetic patient cases using NAVIFY CTM or online trial databases.
[b]Participants were randomly assigned to complete the trial matching exercise with NAVIFY CTM or online trial databases first.

Table 1 shows the order of procedures during the simulation sessions. Each participant was asked to match 10 synthetic patient cases to lung cancer clinical trials—five cases using NAVIFY CTM and five using other online trial databases—and participants were free to match a case to more than one trial. The order of the trial matching exercises was randomly assigned so that half of the participants started with NAVIFY CTM and half started with online trial databases. The synthetic patient cases were also randomly assigned. All trial matches were recorded by participants in Microsoft Excel.

Before the NAVIFY CTM trial matching exercise, participants were given a brief demonstration of how to navigate NAVIFY TB and how to access NAVIFY CTM. Participants were also provided with a printed copy of the NAVIFY CTM User Guide and shown a 5-min video on how NAVIFY CTM is used. For the comparator, participants were provided with a list of links to online trial databases that are used in clinical practice. Participants were free to use one or multiple of the databases to find a suitable trial. The online trial databases used in the study included ClinicalTrials.gov,[34] EU Clinical Trials Register,[35] National Institute for Health Research (NIHR) CRN Portfolio Search website,[36] NIHR "Be Part of Research" website,[37] International Standard Randomized Controlled Trial Number (ISRCTN) registry,[38] and North West London Clinical Research Network (NWL CRN) Lung Cancer Portfolio.[39]

A simulation session guide was developed by the research team to describe the simulated clinical scenario. Participants were instructed to read this document at the beginning of each session. A member of the research team was present throughout each session to resolve technical issues and to ensure data collection procedures were adhered to.

To reflect the real-life time pressure of a clinician tasked with this responsibility, participants were given 60 min for each of the two trial matching exercises (i.e., an average of 12 min per patient). All trial matches were recorded by participants in Microsoft Excel. The simulation session concluded with an online Qualtrics survey. Questions in the survey were informed by a review of the literature related

to clinical trial matching in the UK and by interviews conducted with oncologists and research staff. When all study procedures were accounted for, each simulation session lasted approximately 3 hours.

Measured outcomes

Using simulation-based research to evaluate digital health solutions. Participants completed a survey after the simulation session and were invited to provide feedback on the usability and potential applications of NAVIFY CTM in clinical practice, as well as their experience of participating in the simulation sessions and how realistic the synthetic patient cases were.

Trial matching efficiency. The start time of each trial matching exercise was recorded by a member of the research team. Participants were required to record the time they finished each case in Microsoft Excel. Screen recording software Snagit was used to record the computer screen during the simulation sessions. A web browser plug-in called Webtime Tracker was used to record the amount of time participants spent on each website during the sessions. In addition, participants were asked in the post-simulation survey whether they felt they had sufficient time to complete the tasks. This question was structured in a Likert scale with the following possible answers: Completely agree, somewhat agree, neither agree nor disagree, somewhat disagree, and completely disagree.

Quality of trial matches. The trial matches identified by participants were scored by a consultant thoracic oncologist. Trials were scored from 0 to 5 based on the following criteria: 0—No eligibility criteria matched; 1—Age, tumor type and location matched; 2—Age, tumor type, location and TNM staging matched; 3—Age, tumor type, location, TNM staging, performance status, RECIST criteria matched; 4—Age, tumor type, location, TNM staging, performance status, RECIST criteria, previous lines of treatment matched; and 5—Age, tumor type, location, TNM staging, performance status, RECIST criteria, previous lines of treatment, biomarkers, and all other eligibility criteria met. In addition, participants were asked in the post-simulation survey whether they felt trial suggestions provided by each database were relevant for the patient case. This question was structured in a Likert scale with the following possible answers: Completely agree, somewhat agree, neither agree nor disagree, somewhat disagree, and completely disagree.

Cognitive burden. Participants were required to complete an online Stroop Color and Word Test (SCWT) before and after each trial matching exercise (programmed using Psytoolkit.org). The SCWT measures participants' reaction times and the accuracy of their response to text color and wording pairs (a "Stroop task") each of which are either congruent between text and color (e.g., "GREEN") or incongruent (e.g., "RED"). The SCWT has been used pre- and post-task to measure the mental agility of study participants who have completed mentally demanding activities.[40] The SCWT has been used to measure conflict-controlling selective attention/response inhibition and working memory.[41,42] Each SCWT contained 40 individual Stroop tasks and the maximum response time of each task was limited to 1.5 s. If no response was detected within the response time, this task was considered an error. The post-simulation survey also included a subjective measure of mental effort via the Paas scale[43] which is a popular measure of cognitive load and has been used to validate other measures of cognitive load.[44]

Data analysis

Free-text responses to the post-simulation survey underwent thematic analysis which was performed by one researcher using an inductive approach whereby the onus to derive themes was placed on them.[45] The free-text responses were reviewed line-by-line to identify recurring concepts or words, from which a list of codes was derived. The codes were then grouped into distinct themes, which described a positive or negative reflection of participants' experience, until all codes were categorized under a theme. The free-text responses from each participant were reviewed once again and a numerical score was assigned to each theme based on the number of participants that had referenced a concept or word related to the theme.

All Likert scale questions were converted into a one (Completely disagree) to five (Completely agree) scale for analysis. The scores for the trial matches and the data from the SCWTs, Webtime Tracker, and survey were compiled using Microsoft Excel.

Data was imported into and analyzed on R version 4.0.[46] All outcome variables used in this analysis were non-normally distributed, as confirmed by Shapiro–Wilk test results. For assessment of trial matching efficiency, outliers were first identified using Ronser's test (EnvStats package, version 2.4).[47] This method was selected as it allows simultaneous identification of multiple outliers within sufficiently large datasets. Trial matches over 30 min both belonging to matches found using the ClinicalTrials.gov database) were identified as outliers and subsequently removed from the analysis.

Quality of the trials found by the participants of the study was scored by a thoracic oncologist who was part of the team that developed the synthetic patient cases and who was blinded to participants. All matches were rated on a one (worst possible match) to five (best possible match) scale. Cognitive burden was measured in a Paas scale, and converted to a one (very, very low mental effort) to nine (very, very high mental effort) scale. All pairwise comparisons were performed using a Wilcoxon Signed-Rank test, and p values were adjusted using the false discovery rate (FDR) method due to the small number of comparisons being performed at each step. The p value significance threshold used for this study was 0.05.

Pilot of the methodology

Prior to this study, we piloted our methodology by conducting simulation sessions with 10 clinicians, five with experience matching patients to cancer clinical trials in the UK (experts) and five without experience (non-experts). The aim of the pilot phase of the study was to rapidly test and refine the SBR methods (including data collection procedures) with a smaller number of participants. The results of the pilot phase are not reported in this article; however, a comparison of the study methods for the pilot and study phase is shown in Table 2.

Results

A total of 25 clinicians and research staff with experience screening patients for clinical trials in oncology participated in the study. All participants had experience screening patients in public teaching hospitals operated by the National Health Service (NHS) in the UK. Roles of participants included clinical trial practitioners (n = 8, 32%), research nurses (n = 8, 32%), specialist trainees in oncology (n = 7, 28%) and consultant oncologists (n = 2, 8%). Approximately half of participants (n = 13, 52%) had less than 5 years' experience screening patients for clinical trials in oncology.

Table 2. Comparison of study design and data collection procedures between the pilot and study phases.

	Pilot phase	Study phase*
Primary objective(s)	Rapidly test methodology and data collection	Evaluate efficiency, quality, and cognitive burden of trial matching
Sample size	*n* = 10 (5 experts, 5 non-experts)	*n* = 25 (all experts)
NAVIFY CTM software version	v1.3a	v1.4
Comparator		
Presentation of patient cases	Printed materials	Digital (PDF and JPG files)
Online trial databases	ClinicalTrials.gov	ClinicalTrials.gov
	EU clinical trials register	EU clinical trials register
	NIHR Be Part of research	NIHR Be Part of research
	NIHR CRN portfolio search	NIHR CRN portfolio search
	ISRCTN	ISRCTN
		NWL CRN lung cancer portfolio
Data collection		
Efficiency	Time to complete each trial match	Time to complete each trial match
Quality	Eligibility score assigned by thoracic oncologist	Eligibility score assigned by thoracic oncologist
Cognitive burden	SCWT	SCWT
	No practice test	Practice test = 10 tasks
	Time limit per task = 2 s	Time limit per task = 1.5 s
	Paas scale	Paas scale
Use of SBR	Post-simulation survey	Post-simulation survey

* reported in this article. NAVIFY CTM was updated between phases (from v1.3 to v1.4). Changes were also made to the SCWT between the pilot and study phases. Based on the review of studies where the time limit is not calibrated,[57,58] a time limit of 2.0 s was set for all participants in the pilot. The accuracy observed across all participants was above 97%, suggesting a ceiling effect (i.e., performance was not significantly impacted by mental agility). The time limit was therefore shortened to 1.5 s for the study phase; however, accuracy rates remained above 95%, suggesting that the ceiling effect remained.

Using simulation-based research to evaluate digital health solutions

Feedback on simulation sessions and synthetic patient cases. On the post-simulation survey, the majority of participants (*n* = 24, 96%) stated that they were provided with sufficient guidance to participate in the simulations and complete the trial matching exercises. Most participants (*n* = 18, 72%) also reported that they were provided enough clinical information in the synthetic patient cases. A small number of participants said that it would have been beneficial to have more histology information (*n* = 2, 8%), more information on previous treatments (*n* = 2, 8%), and the inclusion of clinical information via a referral letter (*n* = 1, 4%).

Feedback on the use of simulation-based research methods. Table 3 shows the themes derived from the free-text responses in the post-simulation survey on the benefits and challenges of using SBR to evaluate digital health solutions. Benefits identified by participants included the opportunity for clinicians to preview solutions before they are implemented in clinical practice (*n* = 9, 36%) and to provide feedback on early iterations of digital health solutions (*n* = 7, 28%). Challenges reported by participants included a lack of familiarity with the novel solution (*n* = 3, 12%) and the time constraints to complete the relevant exercises (*n* = 2, 8%). A small number of

Table 3. Analysis of free-text responses to the post-simulation survey regarding the use of simulation-based research methods to evaluate digital health solutions.

	n, (%)
Simulation Benefits	
Preview of app in real-time/experience before launch	9, 36%
Provide amendments/feedback before app launch	7, 28%
Broad staff perspective during testing	3, 12%
Training/familiarity before use in clinical practice	3, 12%
Simulation challenges	
Additional time/training required to learn how to use search tools	5, 20%
Lack of time for a challenging exercise	4, 16%
More clinical information available in a real scenario	2, 8%
Participant selection bias/personal limitations	2, 8%
Small sample size of participants/cases	2, 8%
Taking time from clinical duties to participate	2, 8%
WiFi connectivity during simulation	2, 8%
Feeling rushed to complete trial matching exercises	2, 8%
Small range of cases/tumor types	2, 8%
Length of simulation session should be shorter	1, 4%
Unable to use real hospital data	1, 4%
Participants' lack of familiarity with lung cancer	1, 4%
Inability to get feedback from clinical team on eligibility	1, 4%
Difficult to simulate institutional logistics	1, 4%

participants also stated that simulation sessions could be challenging to run if the patient cases lack sufficient data (n = 2, 8%) or if the simulations took too much time away from clinical duties (n = 2, 8%).

Evaluating NAVIFY clinical trial match

Overview of trial matches. A total of 342 trial matches were made across the 10 synthetic patient cases. Of these, 172 trial matches were made using NAVIFY CTM and 170 using online trial databases. Participants failed to match a trial for 34 cases (14%), with 14 of these 34 failures occurring while using NAVIFY CTM and 20 while using online trial databases. Failing to match a case to a trial was not linked to a specific patient case, with six case match failures being the highest for any patient case. When using online trial databases, ClinicalTrials.gov was used for the majority of trial matches (71%) followed by the North West London CRN Lung Cancer Portfolio (12%).

Trial matching efficiency. Measured objectively, participants matched trials faster on average when using NAVIFY CTM (10 min) compared to the average across all online trial databases (11.7 min), although the difference was not significant (p = 0.063). Results were similar when outliers (n = 2) were removed from the dataset (p = 0.092) (Figure 1).

The time taken to match the synthetic patient cases to clinical trials using NAVIFY CTM and each of the online databases used by participants in this study is summarized in Table 4.

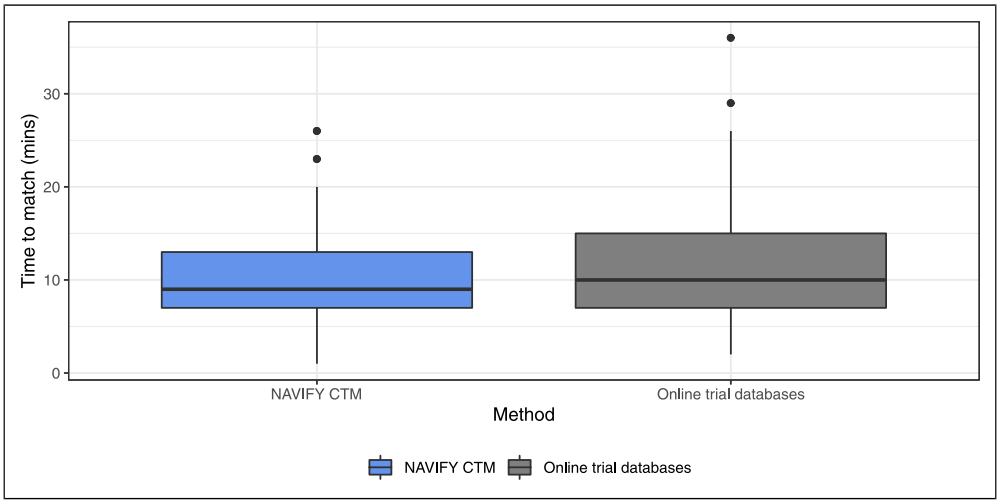


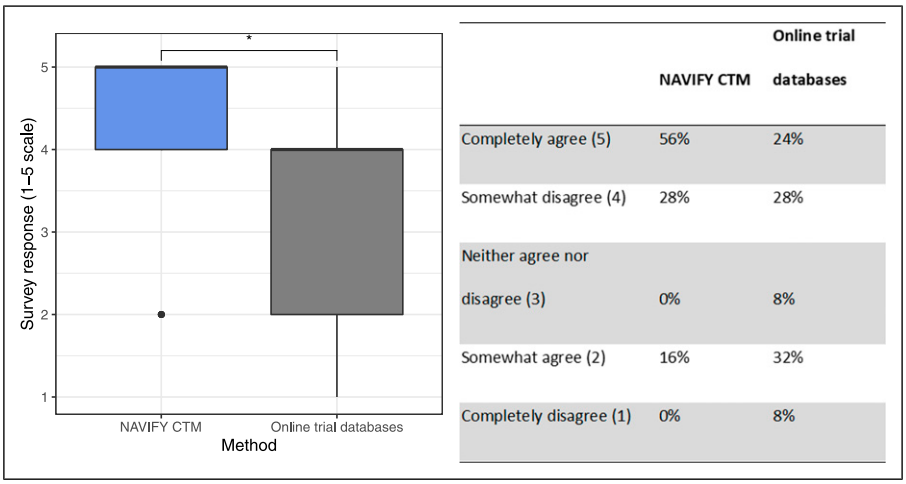
Figure 1. Time to match patient cases to clinical trials, NAVIFY CTM versus all online trial databases combined. (Outliers illustrated as black dots and removed in the statistical tests). Participants matched trials faster using NAVIFY CTM, although the difference was not statistically significant (p = 0.092).

Table 4. Summary of time to match patient cases to clinical trials (NAVIFY CTM vs. online trial databases).

	Interquartile						
	Median (mins)	range (mins)	Mean (mins)	Std. Dev.	Lower CI (95%)	Upper CI (95%)	Trial matches (#)
NAVIFY CTM	9	6	10.0	5.0	9.2	10.7	172
All online trial databases	10	8	11.7	6.6	10.6	12.7	162

	Interquartile						
	Median (mins)	range (mins)	Mean (mins)	Std. Dev.	Lower CI (95%)	Upper CI (95%)	Trial matches (#)
NAVIFY CTM	9	6	10.0	5.0	9.2	10.7	171
ClinicalTrials.gov	10	8	11.6	6.4	10.4	12.7	121
NWL CRN lung cancer portfolio	12	6	11.0	4.3	9.1	13.0	21
EU clinical trials register	15	14	15.4	9.0	7.1	23.7	7
NIHR Be Part of research	7.5	5.5	11.6	11.1	2.4	20.9	8
NIHR CRN portfolio search	9.5	12.8	11.2	7.4	3.4	18.9	6

On subjective measurement in the post-simulation survey, we identified significant differences when comparing the perception of having sufficient time to match patients to clinical trials (Figure 2). Participants reported a significantly higher degree of agreement with having sufficient time to complete the trial matching exercise using NAVIFY CTM (p = 0.020). There was a larger



	NAVIFY CTM	Online trial databases
Completely agree (5)	56%	24%
Somewhat disagree (4)	28%	28%
Neither agree nor disagree (3)	0%	8%
Somewhat agree (2)	16%	32%
Completely disagree (1)	0%	8%

Figure 2. Participants' perception of having enough time to complete the trial matching exercise (by method). Participants reported a significantly higher degree of agreement with having sufficient time to complete the trial matching exercise using NAVIFY CTM ($p = 0.020$). Asterisk (*) indicates results of statistical significance.

proportion of individuals indicating they "completely agree" that they had enough time to complete the task using NAVIFY CTM ($n = 14$, 56%) compared to online trial databases ($n = 6$, 24%).

Quality of trial matches. There was no statistically significant difference in eligibility score for trials matched with NAVIFY CTM compared with all other online trial databases combined ($p = 0.790$). There was a statistically significant difference when comparing NAVIFY CTM to NWL CRN Lung Cancer Portfolio ($p = 0.026$) and NIHR Be Part of Research ($p = 0.040$), although there were few trial matches using these databases ($n = 21$ and $n = 8$, respectively). There were no statistically significant differences compared with ClinicalTrials.gov ($p = 0.590$), the EU Clinical Trials Register ($p = 0.840$), or NIHR CRN Portfolio Search ($p = 0.690$) (Figure 3).

On the post-simulation survey, participants were more likely to agree that trial suggestions via NAVIFY CTM and ClinicalTrials.gov were relevant compared with other online trial databases. No statistically significant differences were identified between NAVIFY CTM and the next best scoring databases, ClinicalTrials.gov ($p = 0.780$), the EU Clinical Trials Register ($p = 0.150$), and NIHR Be Part of Research ($p = 0.150$). Both NAVIFY CTM and ClinicalTrials.gov scored significantly higher compared to NIHR CRN Portfolio Research ($p = 0.049$ and $p = 0.049$, respectively) and ISRCTN ($p = 0.049$ and $p = 0.008$, respectively) (Figure 4).

Cognitive burden

The Stroop effect—defined as the difference in average reaction times (RTs) for incongruent and congruent Stoop tasks—after using NAVIFY CTM was equal to the Stroop effect after using online trial databases. The difference in average RT before and after using NAVIFY CTM was smaller (28 msec) than the same difference for online trial databases (36 msecs), although the difference was not statistically significant ($p = 0.662$ and $p = 0.776$, respectively). Stroop task errors were higher before trial matching using online trial databases compared to after (3.1% vs. 2.4%). For NAVIFY

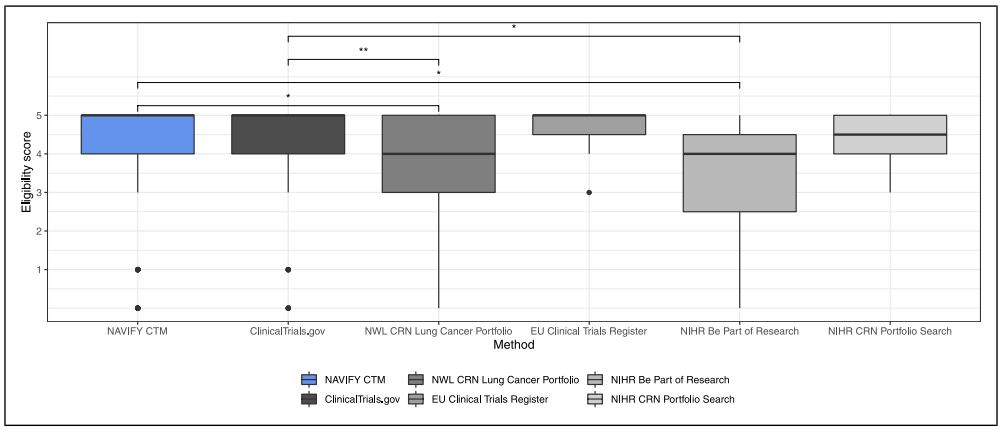


Figure 3. Comparison of eligibility scores for trial matches using NAVIFY CTM and using online trial databases. Single asterisk (*) indicates results of statistical significance, double asterisk (**) indicates non-significant results.

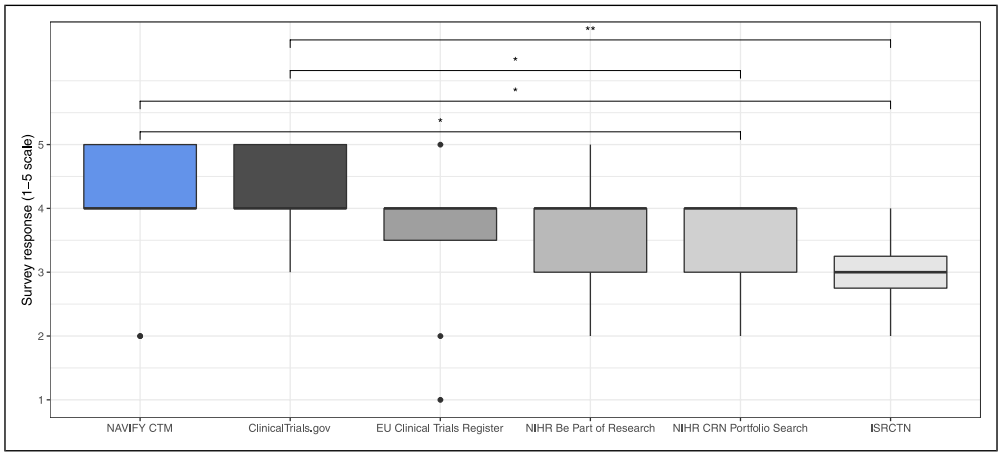


Figure 4. Perceived relevance of top trial suggestions as per post-simulation survey (NAVIFY CTM vs. online trial databases). Participants were significantly more likely to agree that trial suggestions via NAVIFY CTM and ClinicalTrials.gov were relevant compared with the other online trial databases. Single asterisk (*) indicates results of statistical significance, double asterisk (**) indicates non-significant results.

CTM, errors were lower before trial matching compared to after (2.3% vs 2.6%). This difference in errors before and after each trial matching exercise was not statistically significant ($p = 0.993$ and $p = 0.092$, respectively).

When cognitive burden was measured subjectively on the post-simulation survey, participants were significantly more likely to agree that less mental effort was required to complete trial matching using NAVIFY CTM compared to all other online trial databases combined ($p = 0.001$). The difference was significant when compared individually with each of the online trial

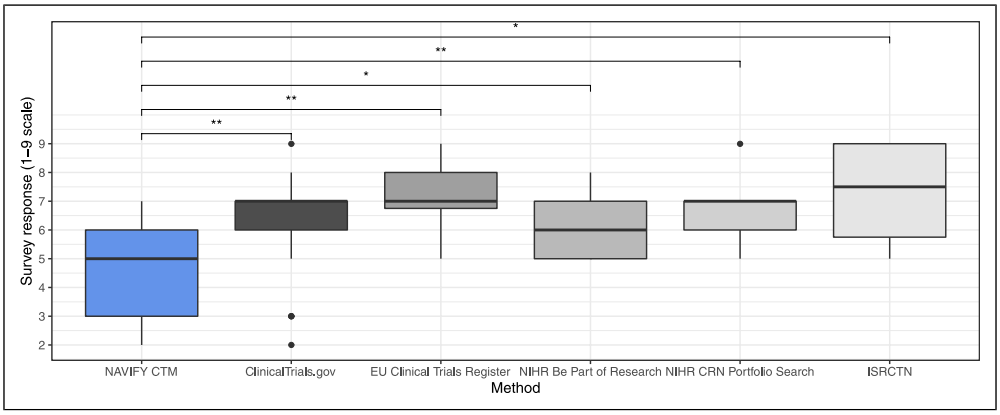


Figure 5. Measure of mental effort required to use NAVIFY CTM compared to online trial databases. Participants reported that less mental effort was required to use NAVIFY CTM and this difference was statistically significant compared to all online trial databases ($p = 0.001$). Single asterisk (*) indicates results of statistical significance, double asterisk (**) indicates non-significant results.

databases, including ClinicalTrials.gov ($p = 0.002$), the EU Clinical Trials Register ($p = 0.001$), NIHR Be Part of Research ($p = 0.047$), NIHR CRN Portfolio Search ($p = 0.005$), and the ISRCTN ($p = 0.039$). Participants were more likely to require "very low," "low," or "rather low" mental effort to use NAVIFY CTM ($n = 11$, 44%) compared to all online trial databases taken together ($n = 2$, 8%). This difference remained when compared against online trial databases individually (Figure 5) (Table 5).

Discussion

Using simulation-based research to evaluate digital health solutions

This study demonstrates the ability to evaluate a novel clinical decision support tool in a simulated clinical scenario using synthetic patient cases. We were able to collect quantitative insights on the performance of NAVIFY CTM compared to online trial databases and to collect valuable user feedback without the need for implementation in clinical practice. The ability to reproduce the same simulated clinical environment before and after, and for all participants, also increases the confidence that any measurable differences between NAVIFY CTM and online trial databases can be attributed to the tools themselves as opposed to external factors.[20]

This study demonstrated the value of using synthetic patient data to support faster evaluations of digital health solutions. Using real patient data in research introduces additional ethical and information governance considerations, such as data privacy and confidentiality. Ethics approval for this study was obtained via a university ethics committee and was completed within two to 3 weeks. Approval from the Health Research Authority in England—with a typical timeline of two to 3 months—was not required as the study did not involve real patients or real patient data, did not take place at an NHS organization, and participants were not recruited via the NHS. Using real patient data also typically involves de-identifying or pseudonymizing data, adding "noise" to data or grouping variables together. However, these approaches are subject to risk as well as delays obtaining data.[48,49]

Table 5. Measure of mental effort required to use NAVIFY CTM compared to online trial databases. Summary table of survey responses via the Paas Scale.

	NAVIFY CTM, %	ClinicalTrials.gov, %	EU clinical trials register, %	NIHR Be part of research, %	NIHR CRN portfolio search, %	ISRCTN, %
9 = very, very high mental effort	0	4	13	0	13	50
8 = very high mental effort	0	8	25	14	0	0
7 = rather high mental effort	20	56	38	29	50	0
6 = high mental effort	12	8	13	14	25	25
5 = neither low nor high mental effort	24	8	13	43	13	25
4 = low mental effort	12	0	0	0	0	0
3 = rather low mental effort	28	12	0	0	0	0
2 = very low mental effort	4	4	0	0	0	0
1 = very, very low mental effort	0	0	0	0	0	0

Another benefit of SBR is the ability to collect measurements of participants' responses during the study and simulated scenarios in ways that would be otherwise impossible or impractical in a clinical setting. We were able to obtain extensive measurements related to cognitive burden with the Stroop Color Word Test (SCWT), record screen activity and gather additional feedback from participants immediately after they conducted the trial matching exercises. It would not be practical or in the best interest of patient care to gather such measurements and feedback during a real-world evaluation conducted in a clinical setting.

One future area of research could be the use of SBR to conduct remote evaluations of digital health solutions. This has important implications in light of the COVID-19 pandemic, which has resulted in shifts to remote working. Guidance published by Public Health England has supported the use of remote approaches to evaluate digital health solutions, as opposed to face-to-face approaches.[50] For certain digital health solutions, such as clinical workflow solutions, teleconsultation services or any other software-based solutions, the design and methodology used in this study could be translated to remote settings and enable evaluations in different geographical locations simultaneously. This approach could generate evidence within a short timeframe, further enhancing the potential for return on investment for innovators.

Evaluating NAVIFY clinical trial match

The hypothesis of the study was that participants would identify trial matches quicker using NAVIFY CTM than online trial databases. While the difference in time-to-match between the tool approaches was not significant, participants were significantly more likely to feel like they had enough time to complete the exercise when using NAVIFY CTM. This effect could be explained by the lower mental effort required to use NAVIFY CTM or as a result of bias towards the solution being evaluated due to the absence of blinding. Risk of bias could be reduced using a single- or double-blind study design, however blinding of any kind can be difficult when evaluating a novel digital health solution. Alternatively, a between-subject study design could be used to mitigate bias amongst participants.

Participants subjectively reported that trial match suggestions from NAVIFY CTM and ClinicalTrials.gov were more relevant compared to the other online trial databases. This finding could be due to factors affecting user satisfaction which caused participants to favor NAVIFY CTM, such as information format, content, consistency and ease of navigation.[51] However, this effect could also be due to information biases which effect participants' judgment, that is, participants may have been influenced or biased towards providing positive feedback about trial matches from NAVIFY CTM because they knew the tool was developed to provide better matches than other online trial databases.[52] Further studies could be conducted to explore the factors influencing the perceived quality of trial match suggestions.

The challenges associated with decision-making in healthcare contexts are well understood and associated with high levels of cognitive load.[6] Therefore, there is a need for solutions that reduce the cognitive load associated with decision-making processes, including clinical trial matching. However, there is limited research on sensitive and objective measurement instruments for the cognitive burden of a task placed on healthcare professionals. Behavior paradigms to measure executive ability or mental agility have been well established in the Experimental Psychology and Neuroscience.[53–55] Future work could explore the use of other paradigms to measure the brain's response to a cognitive event, such as the event-related potential (ERP) paradigm.[56]

Limitations

This study was conducted during the COVID-19 pandemic and lockdown measures were in place which impacted travel,[59,60] as well as the availability of clinicians to participate in research. This may have created a bias towards participants who were located closer to St Mary's Hospital, London or who had more flexible clinical commitments. Furthermore, participants were recruited to complete trial matching exercises individually, while the literature on cancer clinical trial recruitment in the UK shows evidence of multidisciplinary team involvement in decision-making.[4,61] Due to the clinical responsibilities of study participants, it was not reasonable to calibrate the reaction time limit for each participant and to control for sleep and caffeine intake (considered best practice when using the SCWT in experimental psychology).[40]

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: C.H.G., M.P. and O.E. are currently employees of Roche Diagnostics. C.H.G., M.P., and O.E. hold Roche shares. J.H., G.F., and S.G. are Directors at Prova Health, which in turn holds consulting contracts with Roche unrelated to the work described in this manuscript.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study was funded by F. Hoffman-La Roche Ltd.

ORCID iDs

Clarissa Gardner https://orcid.org/0000-0001-9990-7961
Jack Halligan https://orcid.org/0000-0002-8094-5420

References

1. *Feasibility of Cancer Clinical Trial Enrollment Goals Based on Cancer Incidence | JCO Clinical Cancer Informatics [Internet]*. Available at: https://ascopubs.org/doi/full/10.1200/CCI.19.00088 (accessed 7 December 2020).
2. *The Role of Clinical Trial Participation in Cancer Research: Barriers, Evidence, and Strategies | American Society of Clinical Oncology Educational Book [Internet]*. Available at: https://ascopubs.org/doi/full/10.1200/EDBK_156686 (accessed 7 December 2020).
3. Osheroff JA, Teich JM, Levick D, et al. *Improving Outcomes with Clinical Decision Support: An Implementer's Guide*. 2nd ed. Boca Raton: HIMSS Publishing, 2012, p. 348.
4. Cox K, Avis M, Wilson E, et al. An evaluation of the introduction of clinical trial officer roles into the cancer clinical trial setting in the UK. *Eur J Cancer Care (Engl)* 2005; 14(5): 448–456.
5. Patel VL, Kaufman DR and Arocha JF. Emerging paradigms of cognition in medical decision-making. *J Biomed Inform* 2002; 35(1): 52–75.
6. Burgess DJ. Are providers more likely to contribute to healthcare disparities under high levels of cognitive load? How features of the healthcare setting may lead to biases in medical decision making. *Med Decis Making* 2010; 30(2): 246–257.
7. WHO. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. WHO. 2016:1–44.
8. Snowdon DA. *Digital Health: A Framework for Healthcare Transformation*. Chicago, IL: HIMSS, p. 70.
9. Mathews SC, McShea MJ, Hanley CL, et al. Digital health: A path to validation. *Npj Digit Med* 2019; 2(1): 1–9.
10. Hollis C, Falconer CJ, Martin JL, et al. Annual research review: Digital health interventions for children and young people with mental health problems – A systematic and meta-review. *J Child Psychol Psychiatry* 2017; 58(4): 474–503.
11. Singh K, Drouin K, Newmark LP, et al. Many Mobile health apps target high-need, high-cost populations, but gaps remain. *Health Aff Proj Hope* 2016; 35(12): 2310–2318.
12. Murray E, Hekler EB, Andersson G, et al. Evaluating digital health interventions: Key questions and approaches. *Am J Prev Med* 2016; 51(5): 843–851.
13. Guo C, Ashrafian H, Ghafur S, et al. Challenges for the evaluation of digital health solutions—A call for innovative evidence generation approaches. *Npj Digit Med* 2020; 3(1): 1–14.
14. O'Connor S, Hanlon P, O'Donnell CA, et al. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: A systematic review of qualitative studies. *BMC Med Inform Decis Mak* 2016; 16(1): 120.
15. O'Connor S, Hanlon P, O'Donnell CA, et al. Barriers and facilitators to patient and public engagement and recruitment to digital health interventions: Protocol of a systematic review of qualitative studies. *BMJ Open* 2016; 6(9): e010895.
16. Pratap A, Neto EC, Snyder P, et al. Indicators of retention in remote digital health studies: A cross-study evaluation of 100,000 participants. *Npj Digit Med* 2020; 3(1): 1–10.

17. Ong J, Parchment V and Zheng X. Effective regulation of digital health technologies. *J R Soc Med* 2018; 111(12): 439–443.
18. Greaves F, Joshi I, Campbell M, et al. What is an appropriate level of evidence for a digital health intervention? *The Lancet* 2018; 392(10165): 2665–2667.
19. *Evidence Standards Framework for Digital Health Technologies | Our Programmes | What we do | About [Internet]*. NICE. Available at: https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies (accessed 30 October 2020).
20. Lamé G and Dixon-Woods M. Using clinical simulation to study how to improve quality and safety in healthcare. *BMJ Simul Technol Enhanc Learn* 2020; 6(2). Available at: https://stel.bmj.com/content/6/2/87 (accessed 03 February 2021).
21. Aebersold M. Simulation-based learning: No longer a novelty in undergraduate education. *Online J Issues Nurs* 2018; 23: 1–1.
22. Kim J, Park J-H and Shin S. Effectiveness of simulation-based nursing education depending on fidelity: A meta-analysis. *BMC Med Educ* 2016; 16. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4877810/ (accessed 30 October 2020).
23. Higham H and Baxendale B. To err is human: Use of simulation to enhance training and patient safety in anaesthesia. *BJA Br J Anaesth* 2017; 119(suppl_1): i106–1114.
24. Pottle J. Virtual reality and the transformation of medical education. *Future Heal J* 2019; 6(3): 181–185.
25. Cook DA, Brydges R, Zendejas B, et al. Mastery learning for health professionals using technology-enhanced simulation: A systematic review and meta-analysis. *Acad Med* 2013; 88(8): 1178–1186.
26. Gaba DM. The future vision of simulation in healthcare. *Simul Healthc* 2007; 2(2): 126–135.
27. Marshall SD and Flanagan B. Simulation-based education for building clinical teams. *J Emerg Trauma Shock* 2010; 3(4): 360–368.
28. Katsaliaki K and Mustafee N. Applications of simulation within the healthcare context. *J Oper Res Soc* 2011; 62(8): 1431–1451.
29. Cooper K, Brailsford SC and Davies R. Choice of modelling technique for evaluating health care interventions. *J Oper Res Soc* 2007; 58(2): 168–176.
30. Low D, Clark N, Soar J, et al. A randomised control trial to determine if use of the iResus© application on a smart phone improves the performance of an advanced life support provider in a simulated medical emergency. *Anaesthesia* 2011; 66(4): 255–262.
31. Jensen S, Kushniruk AW and Nøhr C. Clinical simulation: A method for development and evaluation of clinical information systems. *J Biomed Inform* 2015; 54: 65–76.
32. Ammenwerth E, Hackl WO, Binzer K, et al. Simulation studies for the evaluation of health information technologies: Experiences and results. *Health Inf Manag J* 2012; 41(2): 14–21.
33. Goodman LA. Snowball sampling. *Ann Math Stat* 1961; 32(1): 148–170.
34. *ClinicalTrials.gov [Internet]*. Available at: https://clinicaltrials.gov/ (accessed 04 May 2021).
35. *EU Clinical Trials Register [Internet]*. Available at: https://www.clinicaltrialsregister.eu/ (accessed 04 May 2021).
36. *NIHR CRN Portfolio Search [Internet]*. Available at: https://public-odp.nihr.ac.uk/QvAJAXZfc/opendoc.htm?document=crncc_users%5Cfind%20a%20clinical%20research%20study.qvw&lang=en-US&host=QVS%40crn-prod-odp-pu&anonymous=true (accessed 04 May 2021).
37. *Be Part of Research [Internet]*. Available at: https://bepartofresearch.nihr.ac.uk/ (accessed 16 June 2020).
38. *ISRCTN Registry [Internet]*. Available at: https://www.isrctn.com/ (accessed 04 May 2021).
39. *CRN North West London [Internet]*. Available at: https://local.nihr.ac.uk/lcrn/north-west-london/ (accessed 04 May 2021).
40. Tanaka M, Ishii A, Shigihara Y, et al. Impaired selective attention caused by mental fatigue. *J Neurol Sci* 2012; 29: 542–553.

41. Stoet G. PsyToolkit: A software package for programming psychological experiments using Linux. *Behav Res Methods* 2010; 42(4): 1096–1104.
42. Stoet G. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teach Psychol* 2017; 44(1): 24–31.
43. Paas FGWC and Van Merriënboer JJG. The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Hum Factors* 1993; 35(4): 737–743.
44. Sweller J. Measuring cognitive load. *Perspect Med Educ* 2018; 7(1): 1–2.
45. Braun V and Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; 3(2): 77–101.
46. R Core Team. R. *A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020. Available at: https://www.R-project.org/
47. Millard SP. EnvStats, an R package for environmental statistics. In: *Wiley StatsRef: Statistics Reference Online [Internet]*. American Cancer Society, 2014. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat07181 (accessed 18 March 2021).
48. Goncalves A, Ray P, Soper B, et al. Generation and evaluation of synthetic patient data. *BMC Med Res Methodol* 2020; 20(1): 108.
49. Tucker A, Wang Z, Rotalinti Y, et al. Generating high-fidelity synthetic patient data for assessing machine learning healthcare software. *Npj Digit Med* 2020; 3(1): 1–13.
50. *Rapid Evaluation of Digital Health Products during the COVID-19 Pandemic [Internet]*. GOV.UK. Available a: https://www.gov.uk/guidance/rapid-evaluation-of-digital-health-products-during-the-covid-19-pandemic (accessed 08 March 2021).
51. Sindhuja PN and Dastidar SG. Impact of the factors influencing website usability on user satisfaction. *IUP J Manag Res* 2009; 8(12): 54–66.
52. Hartmann J, De Angeli A and Sutcliffe A. Framing the user experience: Information biases on website quality judgement. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems [Internet], Florence, Italy, 5 – 10 April, 2008. New York, NY, USA: Association for Computing Machinery; 2008, p. 855–864.
53. Barch DM, Carter CS, Arnsten A, et al. Selecting paradigms from cognitive neuroscience for translation into use in clinical trials: Proceedings of the third CNTRICS meeting. *Schizophr Bull* 2009; 35(1): 109–114.
54. Kessels RPC. Improving precision in neuropsychological assessment: Bridging the gap between classic paper-and-pencil tests and paradigms from cognitive neuroscience. *Clin Neuropsychol* 2019; 33(2): 357–368.
55. Luck SJ and Gold JM. The translation of cognitive paradigms for patient research. *Schizophr Bull* 2008; 34(4): 629–644.
56. Ghani U, Signal N, Niazi IK, et al. A novel approach to validate the efficacy of single task ERP paradigms to measure cognitive workload. *Int J Psychophysiol* 2020; 158: 9–15.
57. Jung M, Jonides J, Berman MG, et al. Construct validity of the multi-source interference task to examine attention in heart failure. *Nurs Res* 2018; 67(6): 465–472.
58. Kandul S and Naguleswaran A. Lying under self-control depletion and time pressure [Internet]. Report No.: 19–05, IRENE Working Paper, 2019. Available at: https://www.econstor.eu/handle/10419/213480 (accessed 18 March 2021).
59. *London South Essex, and South Hertfordshire to Move to Tier 3 Restrictions [Internet]*. GOV.UK. Available at: https://www.gov.uk/government/news/london-south-essex-and-south-hertfordshire-to-move-to-tier-3-restrictions (accessed 23 February 2021).

60. *Prime Minister Announces Tier 4: 'Stay at Home' Alert Level in Response to New COVID Variant [Internet]*. GOV.UK. Available at: https://www.gov.uk/government/news/prime-minister-announces-tier-4-stay-at-home-alert-level-in-response-to-new-covid-variant (accessed 23 February 2021).
61. Maslin-Prothero S. The role of the multidisciplinary team in recruiting to cancer clinical trials. *Eur J Cancer Care (Engl)* 2006; 15(2): 146–154.