Reporting directly to the White House, the FPMA could prioritize, coordinate and oversee efforts to contain a pandemic, while evaluating therapeutics and vaccines. There exist more than 6,000 US acute-care hospitals, and over half have more than 100 beds. Over the past 2 years, most hospitals did not enroll patients in COVID-19 trials due to lack of infrastructure, whereas many academic and high-volume sites were overwhelmed by multiple competing trials. Moreover, most studies were too complex to seamlessly integrate into the workflow of an already stressed health system. A coordinated national clinical trial network would increase efficiency, cost effectiveness and enrollment of diverse populations, while avoiding duplication of efforts. The solutions below build on existing efforts, while others will require longer-term planning and investment.

First, clinical trials should be simplified so that they can be conducted during a pandemic. This would include the use of off-the-shelf adaptable and simplified multi-arm master protocols, similar to that of the UK RECOVERY trial, with a focus on safety and significant clinical endpoints, and with multiple master protocols run in parallel. The simplified protocols should be randomized but unblinded and designed to allow ease of enrollment and minimal disruption of patient care. A centralized Institutional Review Board (IRB) and contracting process will facilitate study start-up.

Document templates, including protocol, investigator brochure, informed consent and case report forms, can be prepared in advance and then tailored to each specific study. A minimal dataset regarding patient outcomes should be entered into a centralized study data-management system to answer safety and efficacy questions as expeditiously as possible. Such real-world master protocol studies will lower costs, reduce duplicative efforts, increase the rate of patient enrollment, improve quality of evidence and provide results in a compressed timeframe.

Second, the US clinical trial infrastructure should be increased. All acute-care hospitals with over 100 beds and accountable healthcare organizations that receive Centers for Medicare & Medicaid Services (CMS) funding should be required to have a clinical trials office prepared to conduct inpatient and outpatient studies, with funding and training for staff provided by the federal government. This will provide added benefit to the national clinical research enterprise, when not in pandemic mode, to support real-world studies of new therapies and comparative effectiveness studies.

Third, multiple clinical trials should not compete for the same patients and resources. All CMS-supported healthcare facilities and plans should prioritize enrollment of patients into master protocols. Study drugs can then be assigned to the master protocols by an FPMA expert panel based upon scientific merit, expected clinical impact and ability to scale up manufacturing of the drug quickly.

Lessons from the COVID-19 pandemic must inform national efforts in the United States to develop effective therapies when the next pandemic occurs. The current US approach with fragmented, overly complex, expensive and often underpowered clinical trials competing for a limited number of patients has largely failed to yield actionable results despite substantial investments. A coordinated national response is therefore required to safeguard health and security. ❐

Rieko Yajima[1,2], Alexander F. More[3,4], Cynthia Garvan[5], Corina Harper[6] and Kevin V. Grimes[1,2] ✉

[1]SPARK Program in Translational Research, Stanford University School of Medicine, Stanford, CA, USA. [2]Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA, USA. [3]Department of Public Health, Long Island University, New York, NY, USA. [4]Initiative for the Science of the Human Past, Harvard University, Cambridge, MA, USA. [5]Department of Anesthesiology, University of Florida, Gainesville, FL, USA. [6]Alexion AstraZeneca Rare Disease, Boston, MA, USA.

✉*e-mail: kgrimes@stanford.edu*

Published online: 31 May 2022
https://doi.org/10.1038/s41591-022-01831-1

References

1. Bugin, K. et al. *Nat. Rev. Drug Discov.* **20**, 254–255 (2021).
2. LaVange, L. et al. *Ann. Intern. Med.* **174**, 1293–1300 (2021).
3. Wise, J. et al. *BMJ* **370**, m2800 (2020).
4. Wilkinson, E. *BMJ* **369**, m1626 (2020).
5. RECOVERY Collaborative Group. *N. Engl. J. Med.* **384**, 693–704 (2021).
6. National Health Service. https://www.england.nhs.uk/2021/03/covid-treatment-developed-in-the-nhs-saves-a-million-lives/ (23 March 2021).

Author contributions

All authors contributed to the manuscript's conception, material preparation and design. The first draft of the manuscript was written by R.K. and K.G., and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Check for updates

Evaluating and reducing cognitive load should be a priority for machine learning in healthcare

To the Editor—The promise of machine learning (ML) to augment medical decision-making in dynamic care environments has yet to be fully realized because of a gap in how algorithms are translated to the bedside, sometimes known as the 'AI chasm'[1]. The drivers of this gap are numerous and complex, but a central challenge relates to the integration of ML into complex decision-making processes and clinical workflows. The ML field has developed a more nuanced appreciation of the importance of having the "human in the loop"[2], but has yet to identify precisely how to optimize the human–ML interface to achieve maximal impact on key outcomes[3].

Cognitive load is a term used to reflect the mental effort required to perform a task, which can be immense in care environments where clinicians collate, integrate, filter, weigh and reason about patient data in real time. Cognitive overload is associated with medical errors and burnout and contributes to suboptimal care outcomes[4,5]. ML is capable of decreasing the mental effort required to process immense amounts of biomedical data, yet its ability to do so is rarely, if ever, evaluated. We argue that cognitive load can and should be measured throughout the ML development cycle to maximize the potential for integration of ML into medicine and to improve patient and provider outcomes.

Table 1 | Rationale and roadmap for measuring cognitive load in machine learning projects in healthcare

Why measure cognitive load?	Excessive cognitive load is associated with suboptimal outcomes for patients
	Machine learning that decreases cognitive load may improve integration potential into challenging workflows
How to measure cognitive load	Psychometric rating scales
	Dual-task procedures
	Physiological measures
When to measure cognitive load	Project inception: assess which problems to target
	Pre-model development: obtain useful baselines
	Pre-deployment: evaluate machine learning model in lifelike contexts
	Deployment: utilize during initial small-scale feasibility studies

The main memory system that supports information processing is working memory. Cognitive load represents the amount of working memory resources consumed by a task. Intrinsic cognitive load refers to working memory resources consumed by the inherent difficulty of a task. Extraneous cognitive load describes working memory resources consumed by unnecessary or distracting details from the environment within which the task is performed. Cognitive load theory is based on the understanding that working memory resources are finite, and when the sum of intrinsic (non-modifiable) and extraneous (modifiable) cognitive load overwhelms working memory, task performance suffers as outlined above. ML that decreases cognitive load should therefore be valued, and ML that increases cognitive load, for example through excessive false alarms[6] or complicated explainability visualizations[7], may actually present unintended additional risk to clinicians, patients and the success of the ML project overall.

Cognitive load should be measured before and after implementation of ML in medicine. Checklists for 'ideal' ML in healthcare have been developed[8], and we argue that an additional requirement should be that the cognitive load of the targeted clinical task(s) decreases with the addition of an ML system. There are qualitative and quantitative techniques that can be employed to make this determination. Psychometric rating scales (such as the 9-point Paas Scale and NASA Task Load Index), dual-task procedures (measurement of performance on a primary task while a distracting secondary task must be performed) and/or physiological measures (such as pupillary dilation, heart rate, galvanic skin response or electroencephalography data) can be used alone or in combination to increase the validity of cognitive load estimation[9].

Such cognitive load tools should be employed throughout the ML development cycle (Table 1). Stakeholders should add assessment of cognitive load to the list of considerations when they decide which medical decision-making processes to target with ML. As a supplement to discussions with clinicians, candidate tasks can be broken down into subtasks and the cognitive load associated with each end user subtask can be measured using the above techniques at the bedside, in the simulation lab and/or with end-user recall surveys. Cognitive load analysis can guide precisely where ML might be maximally impactful.

After the ML model has been trained, validated and tested, various user interfaces and mechanisms for conveying predictions, explainability and uncertainty can be trialed using techniques such as A versus B testing in end users[10]. Visualization techniques that maximally decrease the cognitive load associated with the previously identified subtasks should be prioritized for future study and potential roll-out to the point of care.

Simulation might then be used to recreate the task and its associated clinical ecology, ideally using an exact replica of the care environment in which the study team can trial the intended deployment strategy in a more life-like clinical context. Realistic sources of extraneous cognitive load, such as multi-source data streams and alarming monitors, can be introduced. Paired clinical scenarios might require the clinician to perform the relevant clinical tasks without and with the help of the ML model. The variables associated with ML inference presentation can be validated and/or re-tested at this phase. The psychometric and physiologic metrics of cognitive load described above can be measured and compared between simulations, in addition to traditional metrics of ML and task performance. Pre and post evaluations can be repeated for a predefined cohort of clinical users with varying roles, experience and task expertise to gain a holistic understanding of the potential impacts of the ML model on cognitive load for the entire care team.

We recommend that an ML model proceed to a trial of bedside implementation only if it performs well in traditional metrics of model evaluation, improves task performance and decreases cognitive load during realistic simulations. This will optimize prospects for achieving return on investment for stakeholders and patients, given the substantial investment required to develop and deploy ML in clinical environments. Once the ML model is introduced to the bedside, qualitative and quantitative cognitive load assessments should continue, as part of small-scale feasibility studies to garner valuable feedback about how the model is integrating into clinical workflows[3]. Ethnographic evaluations of the impact of ML on overall unit workload dynamics will determine the intended and unintended consequences of ML deployment. Any initial increases in cognitive load due to unanticipated changes in workflows should ideally be followed by decreases in cognitive load as end users integrate the ML model into their clinical practice. Sustained suboptimal performance on cognitive load metrics during the feasibility study should prompt study teams to hold discussions with clinician end users to inform possible improvements, especially if desired clinical outcomes are unrealized. This is especially important in the context of the COVID-19 pandemic, during which cognitive overload has reached crisis levels in overburdened care environments. ❐

Daniel E. Ehrmann[1,2,8] ✉, Sara N. Gallant[1,8], Sujay Nagaraj[2,3,4], Sebastian D. Goodfellow[1,5], Danny Eytan[1,6], Anna Goldenberg[4] and Mjaye L. Mazwi[1,2,7]

[1]Department of Critical Care Medicine and Labatt Family Heart Centre, The Hospital for Sick Children, Toronto, Ontario, Canada. [2]University of Toronto Temerty Centre for Artificial Intelligence Research and Education in Medicine, Toronto, Ontario, Canada. [3]University of Toronto Temerty Faculty of Medicine, Toronto, Ontario, Canada. [4]Vector Institute for Artificial Intelligence, Toronto, Ontario, Canada. [5]Faculty of Applied Science and Engineering, University of Toronto, Toronto, Ontario, Canada. [6]Department of Medicine, Technion, Haifa, Israel. [7]Department of Paediatrics, University of Toronto, Toronto, Ontario, Canada. [8]These authors contributed equally: Daniel E. Ehrmann, Sara N. Gallant.
✉*e-mail: dehrmann@umich.edu*

Published online: 31 May 2022
https://doi.org/10.1038/s41591-022-01833-z

References

1. Emanuel, E. J. et al. *JAMA* **321**, 2281–2282 (2019).
2. Xin, D. et al. in *Proceedings of the Second Workshop on Data Management for End-To-End Machine Learning* 1–4 (Association for Computing Machinery, 2018).
3. DECIDE-AI Steering Group. *Nat. Med.* **27**, 186–187 (2021).
4. Rothschild, J. M. et al. *Crit. Care Med.* **33**, 1694–1700 (2005).
5. Graber, M. L. et al. *BMJ Qual. Saf.* **21**, 535–557 (2012).
6. Mitka, M. *JAMA* **309**, 2315–2316 (2013).
7. Ghassemi, M. et al. *BMJ Qual. Saf.* **3**, e745–e750 (2021).
8. Loftus, T. J. et al. *PLoS Digit. Health* **1**, e0000006 (2022).
9. Paas, F. et al. *Educ. Psychol.* **38**, 63–71 (2003).
10. Austrian, J. et al. *J. Med. Internet Res.* **23**, e16651 (2021).

Acknowledgements

This work was supported by the William G. Williams Directorship at the Hospital for Sick Children to M.M.

Author contributions

D.E. and S.G. were responsible for the literature search, background and rationale, writing all or part of the manuscript, critical revision of the manuscript and editing of the manuscript. S.N. and A.G. were responsible for critical revision and editing of the manuscript. S.G., D.E. and M.M. were responsible for the background and rationale, critical revision of the manuscript and editing of the manuscript.

Competing interests

The authors declare no competing interests.

Public health law must never again be misused to expel asylum seekers: Title 42

To the Editor — On 1 April 2022, the US Centers for Disease Control and Prevention (CDC) moved to close a sorry chapter in the organization's history by terminating the inhumane immigration policy known colloquially as Title 42[1]. The original order, issued on 20 March 2020, invoked a rarely used provision of US health law — section 265 of US Code Title 42 — allowing the US federal government, on the grounds of public health, to immediately turn away and expel people arriving at the border seeking asylum protection[2]. The order was then revised and renewed multiple times over two years by the CDC under both the Trump and Biden administrations. The public health justification for Title 42 was spurious at its inception and at odds with the science on SARS-CoV-2 transmission and infection. In allowing public health to be weaponized and used as a means to subvert the internationally recognized right to seek asylum, the CDC established a disturbing global precedent that undermined trust in public health institutions at precisely the moment when the world needed that trust most.

Since the earliest days of the pandemic, public health experts have noted that tried and tested public health measures such as masking, social distancing, improved ventilation and testing can be layered together to allow the safe processing of asylum seekers at US borders[3,4]. The Title 42 order was enacted at the start of the pandemic despite the reported objections of senior CDC career scientists, and to this day, none have stepped forward to defend the action[5]. The tenuous public health basis for the order was underscored by the thousands of other travelers — including tourists — who were allowed to cross the border each day with little or no public health measures in place, while the order continued to be applied for the sole purpose of expelling asylum seekers. Notably, in their August 2021 renewal of the order, the CDC stated that safe processing is possible and is currently being implemented for unaccompanied minors. Yet, expanding these safety measures to families and adult asylum seekers took over seven months. Even as the science on COVID-19 evolved, the effective uses of testing and masking were demonstrated, and vaccines became widely available and encouraged, there was no commensurate change in the application of the ban. Indeed, the Title 42 order remained in effect even as masking and vaccination mandates eased around the United States as the Omicron BA.1 wave subsided.

The misuse and misappropriation of public health language by the administration to further immigration control objectives has caused untold damage to individuals and migrant populations. Many thousands of men, women and children have been returned to face threats to life and physical insecurity in Mexico and other countries[6]. Venezuelans have been expelled to Colombia, and thousands of Haitians expelled to Haiti, a country which the Biden administration itself has noted is "grappling with a deteriorating political crisis, violence, and a staggering increase in human rights abuses"[7]. The human costs of these policies have been profound, with over 9,800 kidnappings and other violent attacks (documented by Human Rights First) against migrants blocked in and/or expelled to Mexico since the Biden administration took office[8].

In March 2022, two contradictory court decisions highlighted the uncertainty and harm caused by utilizing public health as a stand-in for immigration policy. A federal court of appeals issued a ruling in a case involving migrant families, noting that they could not be expelled to countries where they may face persecution, and stating that "the CDC's § 265 order looks in certain respects like a relic from an era with no vaccines, scarce testing, few therapeutics, and little certainty"[9]. Yet later that same day, a federal judge in Texas ruled that the Biden administration could not exempt unaccompanied minors from the order despite the CDC's August 2021 update that outlined the government's ability and capacity to safely process this population. The dichotomy between the two rulings demonstrated the need for the CDC to provide clear leadership using the evidence that exists—to reclaim from politicians and lawyers its position as the science-based public health authority that it has always been—a move that finally came on 1 April, with its decision to revoke the order. However, the termination will only take effect on 23 May 2022, inflicting yet more harm on migrants and asylum-seekers in the ensuing weeks.

The developing humanitarian emergency in Ukraine provides a vivid illustration of the important role that asylum and international refugee protections play in protecting human life. Ukraine's neighboring states have waived COVID-19 travel restrictions to allow unfettered admission to safety, thereby living up to both the letter and spirit of international refugee law. European countries are balancing the present but manageable risks posed by COVID-19 with the need to safeguard the right to life and asylum of those seeking protection from violence and conflict.

In sharp contrast, the United States singled out asylum seekers to be banned and expelled. The Title 42 order not only was discriminatory and lacking a basis in evidence, but also fueled the xenophobic trope that migrants are vectors of disease, feeding the stigma and discrimination