Advances in AI for Health in 2021: An Analysis of the Clinicaltrials.gov Registry

The role of artificial intelligence (AI) and machine learning (ML) in healthcare spans a vast range of potential applications. A view of AI development on clinicaltrials.gov helps to focus this landscape, giving a sense of trends and near-term applications. Following last year's review of AI development, we have updated our analysis to include registered studies starting in 2021.[1] The analysis revealed continued strong growth of studies across the field, especially in patient engagement and research. Within patient engagement, the tools studied (mostly chatbots) have increased in complexity, tracking with advancements in natural language processing (NLP). As capabilities evolve, we expect greater uptake of NLP-based tools in clinical workflows (vs simply administrative) in areas of high unmet need– such as behavioral health or low-acuity primary care. Within research, development of models is becoming increasingly formalized, with a greater focus on transparency in reporting. Thus, manufacturers developing AI-enabled tools–especially those hoping to publish their work–will need to meet a higher bar of disclosure of information about the validation of their products.

	Discover	Measure	Recommend	Execute
Patient engagement	N/A	Gauge status	Suggest action	Manage condition
Care delivery	N/A	Diagnose condition	Plan treatment	Deliver treatment
Population health	Reveal complex correlations	Predict risk	Determine intervention	Deliver intervention
Research	Discover new biomarkers	Analyze data	Determine approach	Carry out R&D

Table 1: Segmentation of AI applications

Importantly, this approach offers an incomplete view of usage of AI in healthcare. As requirements for CT.gov registration apply only to clinical trials (interventional studies with

---

[1] We mined clinicaltrials.gov by searching for studies with US site(s) with the following keywords: "artificial intelligence", "machine learning", and "deep learning" and excluded studies with the status "withdrawn". We reviewed each entry to ensure that AI use was germane to the study, and then classified them by domain and function (see Table 1). The results largely yielded studies covering the development, validation and performance of AI-enabled tools for healthcare. However, there were a handful of trials where AI or ML was used as a supporting data analysis tool (for example, to generate study outcomes); these were categorized under "Research / Measure".

human subjects), this analysis will omit use-cases such as AI-driven drug discovery or simplification of administrative functions such as coding and billing. However, the analysis still provides a valuable perspective as registering remains an important step to later publication in prestigious medical journals, and we also see registrations on CT.gov for several (but not all) FDA approved AI-enabled devices.
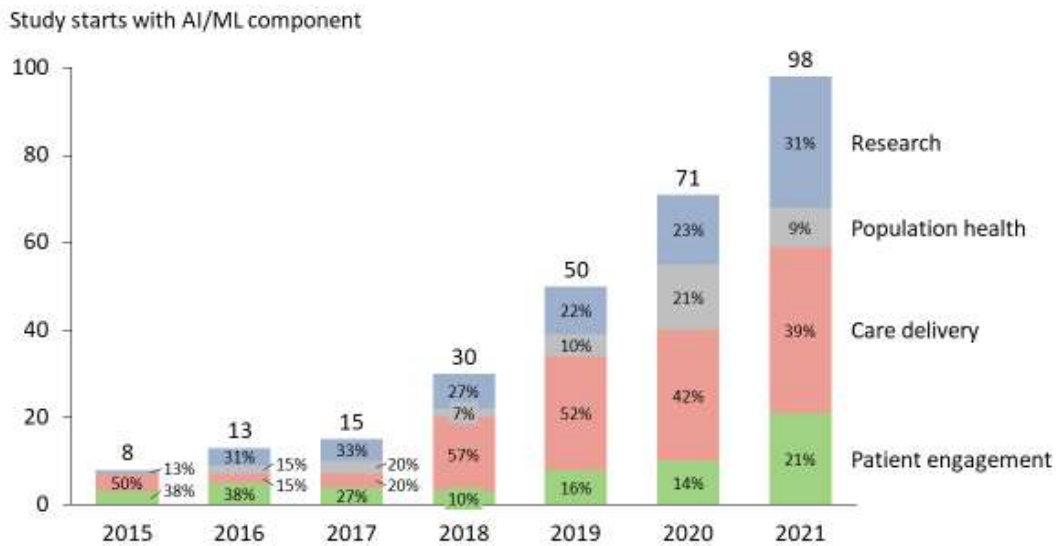

Figure 1: Growth of study starts using AI/ML

In 2021, 98 studies registered with clinicaltrials.gov that used AI/ML were initiated, marking a nearly 40% increase from 2020 (see Figure 1). Care delivery comprised the largest proportion of AI-related studies, with diagnostics remaining as a key focus area. Importantly, two trends emerged suggesting growing capabilities and standardization in the field of AI research: (1) the rise of next-generation patient engagement tools focused on condition management ("Patient engagement, Execute") and (2) the introduction of new guidelines governing reporting of AI trials, spanning development and validation studies to clinical research.

Patient engagement: Rise of next-generation tools
2021 saw the rise of advanced patient-facing tools that focus not only on delivering information to patients, but helping to manage their care as well (see figure 2). A particularly common model used to engage patients is the "digital coach" approach, which serves to augment clinician capacity by digitalizing certain activities. While in the past patient engagement tools

have focused on patient education and the delivery of personalized resources, they have evolved to simulate conversations with the user ("chatbots") and, in some cases, even deliver interventions such as CBT.

This evolution from one-way to two-way communication comes as capabilities in natural language processing (NLP) rapidly advance, thereby enabling these tools to engage in more complex interactions with their users (see Figure 3). Improvements in language understanding, natural language inference, and sentiment analysis, for instance, allow behavioral health chatbots to better interpret the user's struggles and identify appropriate solutions. These improvements are reflected in the growth of NLP research in 2021, with published health and life-sciences papers on Pubmed growing by 46% from 2020. In the same year, newly registered NLP trials on clinicaltrials.gov increased by 75% from 2020.[2] With ~80% of healthcare data estimated to be in unstructured forms, such as text and images, NLP is poised to impact care delivery in a significant way.[3]

Figure 2: Breakdown by segments of study starts using AI/ML 2018-2021

---

[2] https://pubmed.ncbi.nlm.nih.gov/; https://clinicaltrials.gov/
[3]https://www.accenture.com/us-en/blogs/insight-driven-health/unleashing-the-power-of-unstructured-data#_ftn1
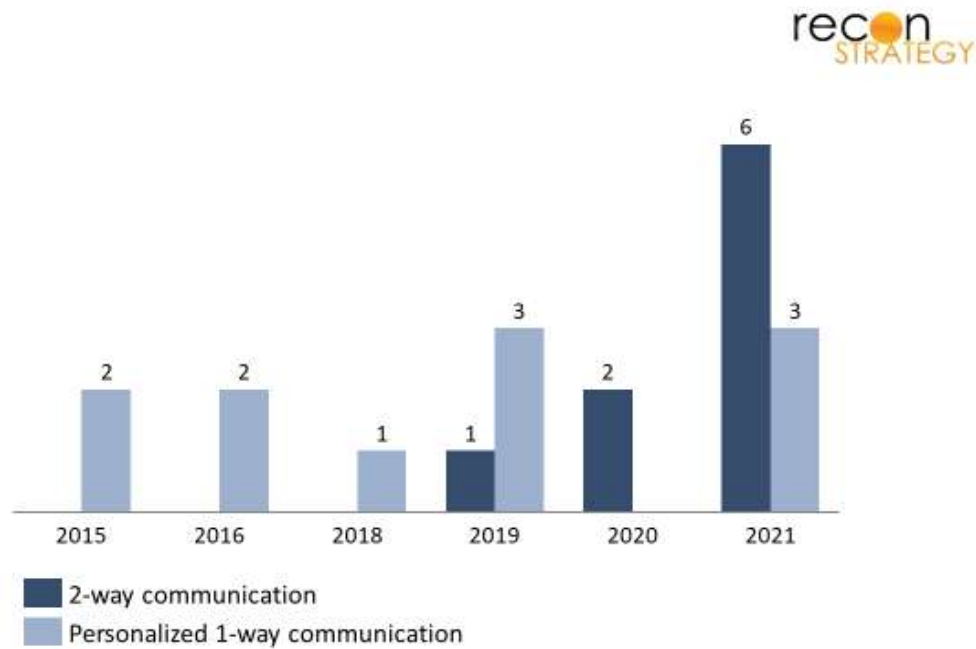
3

Figure 3: Registered studies using patient engagement tools with AI
component by type, 2015-2021

Patient engagement AI solutions have been deployed across the spectrum of clinical care, but with an enduring focus on behavioral health, and more recently, genetic counseling. The value proposition here is clear. These are specialties plagued by clinician shortages and yet require significant hands-on time by highly trained specialists. In behavioral health, a 13-15 week wait time for therapy is typical.[4] Similarly, the wait for a genetic counseling appointment can be up to 6 months. We heard from one geneticist: "For the most part, we're not a high volume specialty since we spend so much time– up to an hour and a half– with each case."[5] Digital counseling thus enables rapid scaling of a consistent intervention across many users, at a minimal incremental cost. These tools also provide on-demand access and the ability to incorporate diverse sources of data, from EHR data to patient-reported symptoms. As this field evolves, we can expect to see further proliferation of NLP-based patient solutions in areas with high unmet demand where activities can be feasibly digitalized, such as low-acuity primary care.

More rigorous AI studies on the horizon

---

[4] "Outpatient Mental Health Access and Workforce Crisis Issue Brief", Feb 2022
[5] Recon expert interview, Oct 2021

In 2021, we also observed a *doubling* in the number of registered "research / discover" studies–largely data-collection studies for the purpose of developing or validating new AI models (see figure 2). Major goals here are to assess patient risk and enable diagnosis, indicating that these will continue to be critical focus areas of the field in the near future. While this increase in registered AI research studies makes sense in the context of the rapid growth of the field, another potential driver is more rigorous documentation and transparency in AI research.

Historically, the level of pre-registration and public disclosure of AI studies has been lackluster. Pre-registration requires the researcher to publish his or her research plan prior to conducting a study, thereby preventing publication bias (failure to publish negative results) and p-hacking (conducting post-hoc statistical analyses that yield a positive result). Currently, pre-registration requirements do not apply to most AI studies, unless they involve human subjects and are planned to be published in a journal.[6] In some cases, however, AI enabled tools have reached approval without publishing clinical data at all, as there is no regulatory requirement to do so.[7] As a result, studies that *are* registered may reflect an overly optimistic view of AI tools.

Recently, however, pre-registration has become more prominent within scientific research as a whole, as well as AI research specifically. A review of the Open Science Framework registry using the keyword "artificial intelligence" reveals that pre-registrations have increased at a CAGR of 115% between 2018 and 2021.[8] Within healthcare, the importance of pre-registering AI trials has been discussed: one article proposing evaluation rules for AI models in healthcare stipulated that studies should be pre-registered in order to combat the biases discussed previously.[9] Thus, the observed uptick in research studies may not be related to a significant increase in AI clinical research, but rather a greater proportion of researchers formally registering these studies.

When trials *are* registered and reported, the quality of reporting has been low. One meta-analysis of deep learning models identified 122 articles, of which 40 were excluded, largely due to lack of comparison with healthcare professionals. Of the 82 remaining, only 25 were sufficiently validated with out of sample data for meta-analysis inclusion.[10] As a result, AI/ML

---

[6]https://www.icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html

[7] https://www.statnews.com/2021/02/03/fda-clearances-artificial-intelligence-data/

[8] OSF registry

[9]Abràmoff, M. D., Tobey, D., & Char, D. S. (2020). Lessons Learned About Autonomous AI: Finding a Safe, Efficacious, and Ethical Path Through the Development Process. In American Journal of Ophthalmology (Vol. 214, pp. 134–142). Elsevier BV. https://doi.org/10.1016/j.ajo.2020.02.022

[10] Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., Mahendiran, T., Moraes, G., Shamdas, M., Kern, C., Ledsam, J. R., Schmid, M. K., Balaskas, K., Topol, E. J., Bachmann, L. M., Keane, P. A., & Denniston, A. K.

study design has been somewhat opaque, making it difficult for others to fully understand these studies and assess their validity.

In response to these issues, the CONSORT and SPIRIT statements, guidelines governing reporting of RCTs and their protocols, were updated to include AI-specific addendums. This initiative aims to boost "transparency and completeness" of clinical trials involving AI interventions by clearly laying out items that should be reported.[11] For example, clear descriptions of the AI intervention should be reported, including the instructions and skills required for use, the setting in which the intervention is integrated, an analysis of error cases, among others.[12] This would apply to RCTs evaluating AI interventions with patient outcomes as the endpoint. Similarly, the forthcoming ML-specific addition to TRIPOD, the guideline that governs the reporting of prediction models, would establish comparable guidelines. This would apply to development and validation studies of ML-based prediction (diagnostic or prognostic) models.[13] This desire for transparency is reflected in the FDA's 2021 AI/ML action plan, which calls for manufacturers to clearly describe elements of the device from input data to evidence of performance.[14]

The establishment of preregistration as a norm and clear reporting guidelines for AI studies will standardize this area of research and "lift the veil". As adherence to these guidelines grows, we can expect greater clinician confidence in AI interventions, making it more likely for these tools to play a real role in clinical practice and improved patient outcomes.

2021 demonstrated the evolving capabilities of artificial intelligence and its ability to bridge gaps in the healthcare system, while also underscoring a need for this innovation to be documented in a rigorous, transparent manner. In 2022, we will revisit the state of AI research. Until then, a few predictions for the year ahead:
- There will be at least 150 AI/ML study starts in 2022, partially boosted by more rigorous reporting guidelines governing AI research

(2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. In The Lancet Digital Health (Vol. 1, Issue 6, pp. e271–e297). Elsevier BV. https://doi.org/10.1016/s2589-7500(19)30123-2

[11] https://www.clinical-trials.ai/spirit

[12] https://www.clinical-trials.ai/consort

[13] Collins GS, Dhiman P, Andaur Navarro CL, et alProtocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligenceBMJ Open 2021;11:e048008. doi: 10.1136/bmjopen-2020-048008

[14] https://www.fda.gov/media/145022/download

- There will be greater diversity in the use of NLP, with patient oriented tools extending to new areas of high unmet need (such as primary care) and leveraging diverse data sources
- Care delivery will continue to grow, as many of the de novo models in the research category begin to be evaluated in randomized controlled trials

As AI interventions and tools progress, we can expect to see greater uptake in clinical practice. As this occurs, AI ethics and more stringent regulatory oversight will come to the forefront to ensure these tools are being deployed in a safe, unbiased and responsible manner.