INTRODUCTION

In 2012 Vinod Khosla, a famed Silicon Valley investor and someone prone to predicting the future, spoke at a Health Innovation Summit in San Francisco and famously said that machines will replace 80% of doctors in the future.1 We are now on the cusp of the massive adoption of digital health technologies. This talk aims to describe the current state of digital transformation in healthcare, to identify reasons for enthusiasm and caution, and to provide a framework for thinking about what is necessary for hospitals and health systems to be confident about incorporating these innovations into practice. I have three key recommendations. First, we should buy results, not claims. Those in positions that influence decisions about endorsing or purchasing digital products designed to improve care or outcomes ought to buy results, not claims or intermediate results. Moreover, although analytic validity and clinical validity are important, they sometimes do not reflect the impact of a product in its entirety. Ultimately, we need to know whether patients benefit. Second, we should insist on transparency. The performance of a product cannot be a secret. The basis on which developers make claims about their products should be open to all, including patients. Better yet, data on which experts reach a conclusion should be shared, just as many companies share research data on drugs and devices. Third, we should be aware of unintended adverse consequences. We should evaluate every intervention for unintended adverse consequences. Changes to systems, with all good intentions, can always go awry. In conclusion, insistence on good and evolving evidence is the best way to arrive at our destination: the use of innovations to improve outcomes.

Key Words: Digital health technology, data science, SaMD
ast of the future of digital health and I believe it will transform and improve outcomes, we have work to do before we get there.

Hello, I am Harlan Krumholz. I am a cardiologist and the Harold H. Hines, Jr. Professor of Medicine at Yale. I am the Director of the Yale New Haven Hospital Center for Outcomes Research and Evaluation. Our group principally seeks to obtain knowledge and insights from data that support actions to improve health and to promote more effective, efficient, equitable, patient-centered, timely, and safe healthcare. A central focus is leveraging technology and digital transformation. We also seek greater transparency in healthcare with better information to inform patients, clinicians, healthcare administrators, and policymakers. We measure our success by the impact we have on people’s lives.

This talk aims to describe the current state of digital transformation in healthcare, to identify reasons for enthusiasm and caution, and to provide a framework for thinking about what is necessary for hospitals and health systems to be confident about incorporating these innovations into practice.

DEFINING DIGITAL HEALTH

First, it seems essential to define digital healthcare. For this talk, I will use the approach used by the United States Food and Drug Administration (FDA) to define the scope broadly. They have written that “digital health includes categories such as mobile health, health information technology, wearable devices, telehealth and telemedicine, and personalized medicine.” They go on to write that “digital health technologies use computing platforms, connectivity, software, and sensors for health care and related uses.”

Digital health is part of the Third and Fourth Industrial Revolutions, as it applies to medicine and public health. The First Industrial Revolution used water and steam to fuel production. The Second used electricity to power mass production. The Third Industrial Revolution used electronic and information technology and produced intelligent and connected machines and systems.

Many industries have fully leveraged the Third Industrial Revolution and, according to Klaus Schwab, Founder and Executive Chairman of the World Economic Forum, are moving to the Fourth Industrial Revolution. This Fourth Revolution fuses advanced technologies across physical, digital, and biological domains. His point is that digital technologies will next become integrated into every system in ways that are “transforming societies and the global economy.”

In healthcare, we see technology adoption producing intelligent automation that transforms workflwos, data collection mechanisms that expand what we can see, and data analysis using digital data flows that augment what we know. According to the UK National Institute for Health and Care Excellence, digital solutions can be classified as solutions that 1) improve system efficiency, but with no measurable patient outcome benefit; 2) inform or deliver essential monitoring and provide two-way communication; and 3) intervene, including preventive behavioral change, self-management, treatment, active monitoring, calculations that influence treatment, and diagnosis.

This talk focuses on interventions intended to make a difference in patient outcomes, although we will also review the larger landscape.

SETTING THE STAGE

Almost everyone in healthcare knows that we are in a period of intense interest in digital health. According to IQVIA, consumer health-related mobile applications available globally in top app stores exceed 350000. Even more impressive is that they report that companies added more than 90000 apps in 2020. Although most of these apps are for general wellness, the number of apps designed for specific health conditions is growing. The leading areas are mental health, diabetes, and cardiovascular disease.

Among these apps, there are a few winners, but the majority have not received much attention. For example, IQVIA reports that 83% of the apps have fewer than 5000 downloads, accounting for less than 1% of the total downloads. Meanwhile, 110 apps have more than 10 million downloads, making up more than 50% of the downloads. The most popular apps include apps from across the globe: US-based WebMD, which focuses on patient education, and GoodRx, which helps people find a low drug price; Mobile JKN, an Indonesian government app that connects patients and providers; MHRS Mobile, a Turkish government app that helps with appointments; and Tata 1mg. Online Medical Store and Health app, an Indian product for appointments and home delivery of medication. As you can see, these most popular digital health apps are about improving system functions or informing patients.

In addition to consumer apps, the consumer wearable market is growing rapidly. These devices augment our ability to monitor patients in and out of traditional healthcare settings and introduce a new range of digital biomarkers. For example, how many people do you know who used a device or wearable to evaluate oxygen saturation at home? These devices are now almost as standard as thermometers and increasingly connect to the larger health ecosystem. These remote monitoring and digital biomarker capabilities are also becoming part of clinical care and clinical research.

Then there is a proliferation of digital therapeutics and digital care tools. These products use digital data and software to prevent or manage disease. This area is multiplying in size.

The COVID-19 pandemic has amplified interest in digital health, bringing a need for monitoring and care delivery outside of usual venues. In the US, the FDA relaxed regulatory re-
requirements relative to digital therapeutics, and the government made changes in payment policies to promote the use of telehealth. As a result, a big trend during the pandemic was the change toward patients receiving care at home. An entire constellation of products moved to fill this need. They included digital therapeutics, consumer wearables, connected biometric sensors, smartphone cameras, telemedicine visits, and web-based interactive programs.

Global investment in companies seeking to leverage digital transformation in health is growing every quarter. According to CB Insights, a market intelligence company, funding in the second quarter of 2021 exceeded $14 billion.

REASONS FOR OPTIMISM AND CAUTION

There are reasons for great enthusiasm. This revolution holds promise for collective wisdom in medicine: we will have the capacity to get smarter with every interaction and make that wisdom accessible for the next patient. Every great consumer tech company, from Google to Apple to Amazon to Samsung and many others, leverages digital data to improve its products. They are providing services and learning from real-world experience in every interaction. They do not just ask people what they want—they observe the behaviors and the responses and see the revealed preferences. They are collecting information that is more comprehensive and denser than has ever been previously available. The enormous computing power has made it possible to organize, manage, and analyze these data almost instantaneously and provide useful output.

They have built systems for convenience and to produce high levels of customer satisfaction. They reach people where it is convenient. They have obviated the need for unnecessary efforts and expenses and re-designed to serve people with information, services, and products. They have created more effective systems built on technology and data science that enhance effectiveness, convenience, and workflows.

But change is not easy, and medicine is a marketplace where it can be challenging to innovate, as the stakes are higher. Moreover, incumbent institutions and workforces that thrive in the current systems tend to resist change. And yet, the digital transformation will be irresistible. Already, data are primarily digital. The next stage will be leveraging the possibilities of digital data.

CAUTIONS

To be clear, the goal of digital transformation must be better health and healthcare. The goal is not to push digital innovations. These innovations are the means to a goal, not the goal itself. The attractiveness of the technology must not obscure the importance of focusing on what it has accomplished or not accomplished for patients, the public, and society. The audaciousness of any claims should not blind us to the need for proof of what they do.

A theme of this talk is that technology is necessary but not sufficient to ensure the improvements in the outcomes we seek. Moreover, all technologies are not equal, and there is significant variation in their products, suitability, and impact. And then, even with stellar technology, translation into benefits depends on actions people take.

The challenges involved in digital health do not deter companies from making claims. One website I saw the other day said that they deliver “improved clinical outcomes, faster reimbursement, lower overheads, and higher revenues.” Unfortunately, there is no evidence cited for this claim. In digital health, there is a surfeit of benefit claims and a lack of evidence of said benefits. While we are generally falling short of introducing innovations that fundamentally shift the experience and outcomes of patients, the promise is there.

We should be focused on results and not satisfied until we can demonstrate that what is new produces tangibly better results than the status quo. In this focus on the result, we need to ask: What does the innovation produce and for whom? Are there unintended adverse consequences? Has bias led to results that favor one group and penalizes another?

Also, the field is dynamic, and many products do not survive. In 2020, as stated, almost 100000 new apps appeared in top app stores. Many others disappeared. According to IQVIA, between July 2017 and June 2021, more than 350000 new apps appeared. Still, there was a net gain of only 32726 because so many had been discontinued, including many that developers made available for relatively brief periods before discontinuation. About half of the apps deleted from the app stores had fewer than 100 downloads. Many had never been updated, suggesting that developers abandoned them quickly. In addition, developers never updated many apps, rendering them unable to leverage new features from new operating systems.

I am avid for what is to come in the digital transformation of medicine, but there is still work to do before its promise is truly achieved. As stated in a recent Perspective in npj Digital Medicine by Guo and colleagues, “a huge challenge for end-users, including patients and providers (e.g., healthcare professionals, hospital administrators), is how to determine a new solution’s credibility and compliance with standards.” The key will be in the evidence.

FRAMEWORK

The FDA has developed a helpful framework to evaluate many new innovations. They use the term ‘software as a medical device (SaMD).’ The International Medical Device Regulators Forum defines SaMD as “software intended to be used for one or more medical purposes that perform these purposes with-
out being part of a hardware medical device.”

In clinical evaluation, they distinguish between a valid clinical association, analytic validation, and clinical validation. The valid clinical association focuses on whether the product’s output produces information relevant to the targeted clinical condition. In essence, do we care about what the result does? If not, we can stop there.

The next step is analytic validation. The question here is whether the software correctly processes the input to generate accurate, reliable, and precise output data—whether the software does what it claims to do. Does the algorithm “correctly and reliably process the input data and generate output data with the appropriate level of accuracy, repeatability, and reproducibility.” The focus here is on accuracy, reliability, and precision.

The next step is clinical validation, which is what I want to focus on and deem to be the critical question for these products. Clinical validation asks whether a product’s output achieves the intended purpose in the target population—whether the software produces clinically relevant results. The focus here might be on sensitivity and specificity.

These stages ensure that any output is relevant, that the software product gives accurate, reliable, and precise information, and that it achieves its intended purpose in the target population. There is another facet too, which is continuous evaluation. The performance of the product may change as it disseminates and as the health system changes around it. Thus, there is a need for continuous assessment and iteration to improve the system and resist degradation in performance.

In these last two steps, the validations must occur in diverse cohorts of patients that reflect those for whom the product is designed, as well as in venues and health systems that these patients are likely to use. For example, developers must validate digital dermatological products that facilitate diagnosis in people with various skin tones. Other products that sense signals through skin must do the same. For example, experts have found in the US that pulse oximeters had a bias such that they performed worse in Black patients than White patients. In a study in the New England Journal of Medicine, a device was three times more likely to miss the diagnosis of hypoxia in Black patients than in White patients. Presumably, the manufacturer tested the devices on White people and optimized its performance in that population, neglecting that the performance might change with skin tone.

Of note, government agencies do not regulate many apps. Moreover, the data supporting analytic and clinical validation are not publicly available, regardless of whether they are regulated. And the variation in quality is immense.

I think there is another level outside this framework—one that should concern people on the front lines of medicine—validation based on outcomes. Outcomes validation should make a difference in the marketplace. The question here is whether the product produces a meaningful impact.

There are many steps to deriving meaningful impact. First, will people use the product? Second, will people act on the information in ways that are likely to improve outcomes? Third, do outcomes improve? In this context, outcomes could be effectiveness, efficiency, equity, satisfaction, or safety.

The problem is the quality of the evidence.

1) There is no evidence of the clinical impact of many products.
2) The interpretation of published studies is marred by publication bias, which means that researchers only report some of the conducted studies.
3) Among those that do have published studies, methodological weaknesses often undermine the evidence.

In the Journal of the Comprehensive Cancer Network, Pawloski and colleagues did a systematic review of studies to evaluate the clinical effectiveness of clinical decision support tools—digital interventions designed to improve outcomes. They identified only 24 studies in oncology, and most had inferior scientific designs that would likely lead to bias and all but one reported a positive finding.

Publication bias is endemic in medicine. Among all clinical trials, our research group found that investigators only report about half. Even among NIH-funded trials, nearly a third of investigators never report results. In digital health, that is likely to be worse. All related studies need to be registered ahead of time and reported in a timely way, even if they fail.

Among published studies, there are indications of problems. For example, remote digital health studies have trouble keeping people engaged. Therefore, a study that makes claims on the very small number of people who stay in the study is providing a biased view of the intervention’s impact. Meanwhile, the dropout rate would tell you something about the willingness of people to use the intervention, even before we get to whether it can produce the desired effect. For example, among eight large remote digital health studies from 2014–2019, which included more than 100000 people, the median participant retention varied from 2 to 26 days; the median across all studies was 5 days. That figure means that the typical person enrolled in the study participated for only 5 days, and in one of the studies, it was only 2 days. After that, more than half the people no longer used the app. Even if the person’s clinician referred them to the study, the median participation retention was only 40 days. Paying people increased the retention time to only 22 days. Thus, people generally do not stick with these apps. Also, the studies did not have representative populations. In another example, researchers studied dropout rates in a clinical trial of smartphone apps for depressive symptoms. In 18 published studies, with data from 3336 participants, the dropout rate was estimated to be 48%. They also found evidence for substantial publication bias.

There is a need for new approaches. We need more and better evidence and evidence generated after the developers release the product. We may need new types of research meth-
ods that are adaptive and iterative and permit rapid learning. Google, Samsung, and Amazon do not run randomized trials as we do in medicine. They have found ways to do A/B testing that is rapid, efficient, and agile—and fits into their business model by shortening the distance between testing, learning, and applying. As we enter this digital age in medicine, we need to be adept at changing traditional models of knowledge generation and testing interventions. We need to use new digital tools to strengthen the testing and evaluation of these products in real-world settings and with real-world patients. We also need to appreciate the importance of optimizing the initial versions. Digital products are not like a drug, a fixed product evaluated for patients who meet the inclusion criteria. Digital products have features that developers can easily modify to improve how they engage people, support actions, and produce favorable effects. Evaluations should be geared toward testing iterations of an idea or product, rarely stopping after a failure and persisting until the effort is without merit. There may be negative trials that teach us something important that can help improve the product.

In a TechCrunch blog by Khosla, written just before his prediction of a healthcare future in which he predicted that we would not need 80% of the doctors, he wrote, “The best way to predict this future is not to extrapolate the past and what has or has not worked, but to invent the future we want, the one we believe possible!” I think this sentiment is as relevant to how we generate knowledge about medicine and healthcare as how we deliver healthcare.

**RECOMMENDATIONS**

So, where does that leave us? I think we are at a point where we should not drink the kimchi soup first. I mean that we should not be so overwhelmed by the exuberant excitement about a digital era that we neglect to do the work that ensures we are making wise choices for our patients and our society. The proliferation of products and the chatter of the marketplace can seduce us into the adoption of products that may, at the margin, hold little value and could even cause harm.

My foundational approach is to ask, what did we accomplish for patients at the end of the day? Were their lives tangibly improved? Did they live longer? Did they live better? Was their care made more affordable? Were their preferences and values respected? Did they have every opportunity to reach their personal health goals? We should apply these questions to these new products, building on the new capabilities made possible by technology and digitization.

I have three key recommendations to end this talk.

**Buy results, not claims**

Those in positions that influence decisions about endorsing or purchasing digital products intended to improve care or outcomes ought to buy results, not claims or intermediate results. I mean that claims are ephemeral, and we need to know the effects of the products. Moreover, although we need analytic and clinical validity, they can be insufficient in determining whether a product is beneficial to patients. We cannot consider a product a success if its results are not favorable.

For interventions intended to improve care and outcomes, we must create demand for the information about these products that demonstrates what they do. By doing so, we will create an innovation ecosystem that is accountable for claims it makes. To do so requires that we buy results.

**Insist on transparency**

The performance of any product cannot be a secret. The basis on which developers make claims about their products should be open to all, including patients. Better yet, the data on which experts reach a conclusion about a product should be shared, just as many companies share research data on drugs and devices.

**Beware of unintended adverse consequences**

We should evaluate every intervention for unintended adverse consequences. Changes to the system, with all good intentions, can always go awry. Thus, evaluations need to focus not only on the prospect of benefit but on the possibility of harm. Those harms may be more broad-based than the intended benefit, so safety evaluations should include a set of comprehensive outcomes.

**CONCLUSION**

We are entering an extraordinary era of transformation. As such, we need to reward innovations that make the most positive difference (not the best salespeople, the best results). In the end, the net result should be better health and healthcare outcomes for our patients.

I have a very pragmatic approach. Maybe I was influenced by where I was born. In the US, there is a state called Missouri, also known as the Show-Me State—a place where words were just words and people there wanted to see evidence for themselves before believing. Here, in a time of such great promise, we need to insist on seeing evidence.

I suggest that this moment requires that we all adopt that approach. Only with some skepticism and insistence on evidence (continuing through a product’s lifecycle) will the digital age achieve what it can. The pandemic has accelerated progress. Now we need to consolidate what we have learned and be disciplined in how we proceed, insisting on good and evolving evidence to arrive at our destination of not mere adoption of innovations, but the use of innovations to improve outcomes.
REFERENCES

1. Khosla V. Vinod Khosla says technology will replace 80 percent of doctors—sparks indignation [Internet] [accessed on 2022 January 3]. Available at: https://venturebeat.com/2012/09/02/vinod-khosla-says-technology-will-replace-80-percent-of-doctors-sparks-indignation/.

2. Krumholz HM. Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system. Health Aff (Millwood) 2014;33:1163-70.

3. Mori M, Khera R, Lin Z, Ross JS, Schulz W, Krumholz HM. The promise of big data and digital solutions in building a cardiovascular learning system: opportunities and barriers. Methodist DeBakey Cardiovasc J 2020;16:212-9.

4. Burns H. Germ theory: invisible killers revealed. BMJ 2007;334 suppl 1:s11.

5. Krumholz HM. Inflection point: ideas for accelerating breakthroughs and improving cardiovascular health. Circ Cardiovasc Qual Outcomes 2020;13:e007615.

6. Yale School of Medicine. Center for outcomes research & evaluation [Internet] [accessed on 2022 January 3]. Available at: https://medicine.yale.edu/core/.

7. U.S. FDA. What is digital health? [Internet] [accessed on 2022 January 3]. Available at: https://www.fda.gov/medical-devices/digital-health-center-excellence/what-digital-health.

8. Schwab K. The fourth industrial revolution [Internet] [accessed on 2022 January 3]. https://www.britannica.com/topic/The-Fourth-Industrial-Revolution-2119734.

9. National Institute for Health and Care Excellence. Evidence standards framework for digital health technologies [accessed on 2022 January 3]. Available at: https://www.nice.org.uk/corporate/ecd7/resources/evidence-standards-framework-for-digital-health-technologies-pdf-1124017457605.

10. IQVIA. Digital health trends 2021: innovation, evidence, regulation, and adoption [Internet] [accessed on 2022 January 3]. Available at: https://www.iqvia.com/insights/the-iqvia-institute/reports/digital-health-trends-2021.

11. WebMD. [Internet] [accessed on 2022 January 3]. Available at: https://www.webmd.com/.

12. GoodRx. [Internet] [accessed on 2022 January 3]. Available at: https://www.goodrx.com/.

13. Observatory of Public Sector Innovation. Mobile JKN, health insurance services in your hand [Internet] [accessed on 2022 January 3]. Available at: https://oecd-opsi.org/innovations/mobile-jkn-health-insurance-services-in-your-hand/.

14. Republic of Turkey Ministry of Health. MHRS appointment [Internet] [accessed on 2022 January 3]. Available at: https://ar-havrdh.saglik.gov.tr/EN,326412/mhrs-appointment.html.

15. Tata 1 mg. [Internet] [accessed on 2022 January 3]. Available at: https://www.1mg.com/.

16. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Devices and Radiological Health, Office of Product Evaluation and Quality. Enforcement policy for digital health devices for treating psychiatric disorders during the coronavirus disease 2019 (COVID-19) public health emergency: guidance for industry and Food and Drug Administration staff [accessed on 2022 January 3]. Available at: https://www.fda.gov/media/136939/download.

17. United States Department of Health and Human Services, Centers for Medicare & Medicaid Services. Final policy, payment, and quality provisions changes to the Medicare Physician Fee Schedule for calendar year 2021 [Internet] [accessed on 2022 January 3]. Available at: https://www.cms.gov/newsroom/fact-sheets/final-policy-payment-and-quality-provisions-changes-medicare-physician-fee-schedule-calendar-year-1.

18. CB Insights. State of healthcare Q2’21 report: investment & sector trends to watch [Internet] [accessed on 2022 January 3]. Available at: https://www.cbinsights.com/research/report/healthcare-trends-q2-2021/.

19. Guo C, Asgharian H, Ghafur S, Fontana G, Gardner C, Prime M. Challenges for the evaluation of digital health solutions—A call for innovative evidence generation approaches. NPJ Digit Med 2020;3:110.

20. U.S. FDA. Software as a medical device (SaMD) [Internet] [accessed on 2022 January 3]. Available at: https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd.

21. IMDRF SaMD Working Group. Software as a medical device (SaMD): key definitions. [accessed on 2022 January 3]. Available at: https://www.imdraf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf.

22. Sjöding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. N Engl J Med 2020;383:2477-8.

23. Pawlowski PA, Brooks GA, Nielsen-Bullis BA. A systematic review of clinical decision support systems for clinical oncology practice. J Natl Compr Canc Netw 2019;17:331-8.

24. Becker JE, Krumholz HM, Ben-Josef G, Ross JS. Reporting of results in ClinicalTrials.gov and high-impact journals. JAMA 2014;311:1063-5.

25. Ross JS, Tse T, Zarin DA, Xu H, Zhou L, Krumholz HM. Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. BMJ 2012;344:d7292.

26. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100000 participants. NPJ Digit Med 2020;3:21.

27. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. J Affect Disord 2020;263:413-9.

28. Khosla V. Do we need doctors or algorithms? [Internet] [accessed on 2022 January 3]. Available at: https://techcrunch.com/2012/01/10/doctors-or-algorithms/.