NBER

NATIONAL BUREAU *of* ECONOMIC RESEARCH

THE ECONOMICS OF ARTIFICIAL INTELLIGENCE

Health Care Challenges

Edited by Ajay Agrawal, Joshua Gans, Avi Goldfarb, and Catherine E. Tucker

2

The Potential Impact of Artificial Intelligence on Health Care Spending

Nikhil R. Sahni, George Stein, Rodney Zemmel, and David Cutler

2.1 Introduction

Can artificial intelligence (AI) improve productivity in health care? That is a central question in the United States and the focus of this paper.

In the United States, health care is considered too expensive for what it delivers. Businesses are looking to manage their costs, and health care spending is a large and growing expense. As health care spending grows in the public sector, it crowds out other governmental budget priorities. Previous research has found that health care in the United States could be more productive—both costing less and delivering better care (Berwick and Hackbarth 2012; Sahni et al. 2019). AI is likely to be part of the solution.

The improvement in US health care productivity could manifest in several ways. Administrative costs are estimated to account for nearly 25 percent of all US health care spending (Sahni et al. 2021); AI could reduce this burden. Harnessing clinical knowledge to improve patient health is a second way. Medical knowledge is growing so rapidly that only 6 percent of what the average new physician is taught at medical school today will be relevant

Nikhil R. Sahni is a partner and leader of McKinsey's Center of US Healthcare Improvement at McKinsey & Company and a fellow in the Economics Department of Harvard University.

George Stein is an associate partner at McKinsey & Company.

Rodney Zemmel is a senior partner and global leader of McKinsey Digital at McKinsey & Company.

David Cutler is the Otto Eckstein Professor of Applied Economics at Harvard University and a research associate of the National Bureau of Economic Research.

This paper was prepared for the NBER Economics of Artificial Intelligence Conference, September 2022. The paper represents the views of the authors and not any entity with which they may be affiliated. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see https://www.nber.org/books-and-chapters/economics-artificial-intelligence-health-care-challenges/potential-impact-artificial-intelligence-healthcare-spending.

in ten years (Rajkomar, Dean, and Kohane 2019). Technology such as AI could provide valuable clinical data to the clinician at the time of diagnosis. Improving clinical operations is still another example. Operating rooms (ORs) are one of hospitals' most critical assets. Yet inefficient operations can result in wasted hours, leading to excessive building of space, hindering patient access, degrading the patient experience, and reducing hospitals' financial margins.

In this paper we focus on two questions about AI. First, how much might be saved by wider adoption of AI in health care? To answer this, we estimate potential savings by considering how AI might affect processes for three stakeholder groups—hospitals, physician groups, and private payers. For each stakeholder group, we illustrate AI-enabled use cases across both medical and administrative costs and review case studies. Using national health care spending data, we then scale the estimates to the entire US health care industry. We find that AI adoption within the next five years using today's technologies could result in savings of 5 to 10 percent of health care spending, or \$200 billion to \$360 billion annually in 2019 dollars, without sacrificing quality and access.[1] For hospitals, the savings come largely from use cases that improve clinical operations (for example, OR optimization) and quality and safety (for example, condition deterioration management or adverse event detection). For physician groups, the savings also mostly come from use cases that improve clinical operations (for example, capacity management) and continuity of care (for example, referral management). For private payers, the savings come largely from use cases that improve claims management (for example, auto-adjudication or prior authorization), health care management (for example, tailored care management or avoidable readmissions), and provider relationship management (for example, network design or provider directory management). While we only quantify cost savings in this paper, there are additional nonfinancial benefits from the adoption of AI, including improved health care quality, increased access, better patient experience, and greater clinician satisfaction.

The magnitude of these savings raises a second question: If AI in health care can be so valuable, why is it not in greater use? At the organizational level, in our experience, there are six factors for successful AI adoption. AI's limited uptake can be partly explained by the difficulty of addressing these factors, such as the failure to create "digital trust" with patients. In addition, we discuss industry-level challenges such as data heterogeneity, lack of patient confidence, and misaligned incentives. Recent market trends, such as

1. In this paper, we focus only on what is possible using existing technologies. The opportunity increases as more advanced approaches come to market, such as digital twins or generative AI. We also acknowledge the adoption of AI elsewhere in the health care value chain, from medical training to pharmaceutical discovery to medical device manufacturing, none of which are discussed in this paper.

increasing venture capital and private equity investments, may increase the rate of AI adoption in the near future.

The paper is organized as follows. Section 2.2 outlines the scope of potential uses of AI in health care. Section 2.3 lays out how AI might be used for the three specific stakeholder groups—hospitals, physician groups, and private payers—and presents case studies as examples. Section 2.4 estimates the annual net savings that might result from adopting AI across all of US health care. Section 2.5 considers the challenges to greater adoption of AI, and section 2.6 discusses how market trends may change the decisions that organizations make about adopting AI. The final section offers concluding thoughts.

We note that the authors of this paper are an unusual group compared with the authors of other economics papers. One of the authors is an academic, and three are consultants with extensive experience in health care. Thus, our insights draw upon a combination of academic and industry experience. In many cases, the insights are not based on randomized control trials or quasi-experimental evidence; rather, they are distillations of observations from a number of organizations, in and out of health care. Given this, the reader should understand that the evidence base underpinning some of our conclusions is less analytically rigorous than traditional economics papers.

2.2 The Scope of AI

We define AI as a machine or computing platform capable of making intelligent decisions. Health care has more often pursued two types of AI: machine learning (ML), which involves computational techniques that learn from examples rather than operating from predefined rules; and natural language processing (NLP), which is a computer's ability to understand human language and transform unstructured text into machine-readable, structured data. An example of ML is recommending additional purchases based on a consumer's current choices, such as a book or a shirt; an example of NLP is analyzing written customer feedback to identify trends in sentiment that can inform improvements in a product's features.

It is not hard to envision the application of these technologies to health care. ML examples include predicting whether a patient is likely to be readmitted to a hospital, using remote patient monitoring to predict whether a patient's condition may deteriorate, optimizing clinician staffing levels in a hospital to match patient demand, and assisting in interpreting images and scans. NLP examples include extracting words from clinician notes to complete a chart or assign codes; translating a clinician's spoken words into notes; filling the role of a virtual assistant to communicate with a patient, help them check their symptoms, and direct them to the right channel such as a telemedicine visit or a phone call; and analyzing calls to route members to the right resource and to identify the most common call inquiries. Sometimes combining ML and NLP can create greater value; for example, using NLP to extract clinician notes and then using ML to predict whether a prior authorization is needed.

In general, AI-enabled use cases address operational processes. One type of operational change is simplifying an existing process. In these situations, the ideal processes usually are repetitive in nature, are highly manual, or involve complex decision trees. For example, forecasting inventory, demand, and capacity in the manufacturing, retail, and hospitality industries was once a highly manual job, involving meticulous note taking and trend forecasts. AI can perform the same processes faster with more precision. Another type of operational change is the creation of new processes. These generally were not accessible to organizations until now, but AI has unlocked them. For example, some insurance companies allow customers to send a photo of an incident to initiate a claim, which is then automatically processed by AI.

The application of AI in these use cases allows value to be created in several ways. Labor productivity improvement is one of the most important levers in health care. Historically, labor productivity growth in health care has been negative; only education has performed worse among all US services industries over the past few decades (Sahni et al. 2019). For many health care organizations, labor represents the single largest variable-cost item. Value can also be created in nonfinancial ways. For example, furnishing clinicians with data at the point of service could improve the course of treatment selected for the patient based on clinical evidence. As a result, health outcomes may improve with no increase—or even a reduction—in costs.

Adopting AI to create this value would unlock multiple levels of potential automation in health care. We illustrate these by considering the current use of AI in autonomous cars (figure 2.1). The Society of Automotive Engineers defines five levels of automation (Society of Automotive Engineers 2021). Each increasing level involves a greater degree of autonomous input: no driving automation but automatic emergency procedures in level 0, driver assistance such as lane centering in level 1, partial automation such as adaptive cruise control in level 2, conditional driving automation such as in traffic jams in level 3, local driverless taxis in level 4, and anywhere driverless taxis in level 5. At level 3 and above, the technology is in greater control than the human.

It is difficult to align on a single level for all of health care because AI-enabled use cases may vary. For example, clinical decision making is likely to approach level 1: the clinician makes final decisions jointly with the patient, but AI acts as a "member of the team" to present possible courses of treatments. The interpretation of radiology images could exemplify level 2, with AI reviewing an MRI or X-ray and outputting an interpretation. Humans would make the final decision for quality control and ensure the AI algorithm is trained properly. AI-enabled use cases in which technology would

		Human-led			Technology-led		
	Level	0	1	2	3	4	5
Automotive	Description	No driving automation	Driver assistance	Partial driving automation	Conditional driving automation	High driving automation	Full driving automation
Automotive	Example driving features	Automatic emergency braking	Lane centering	Lane centering *and* adaptative cruise control at the same time	Traffic jam chauffeur	Local driverless taxi	Anywhere driverless taxi
Healthcare	Example AI-enabled use cases	N/A	Clinical decision making	Interpreting radiology images	Referral recommendations	Claims automation	

Most likely level of automation for healthcare AI-enabled use cases

Fig. 2.1 Society of Automotive Engineers levels of automation adapted to health care
Source: Society of Automotive Engineers 2021; authors' analysis

play the leading role could include referral recommendations (level 3) and claims automation (level 4).

2.3 Domains of AI in Health Care

To understand how AI might influence health care spending, we start by breaking down the industry into five stakeholder groups—hospitals, physician groups, private payers, public payers, and other sites of care, such as dentists and home health.[2] We focus primarily on the first three, which collectively represent 80 percent of total industry revenue (Singhal and Patel 2022).

For each of these stakeholder groups, we identify the key domains with underlying AI-enabled use cases. A "domain" is defined as a core functional focus area for an organization. A "use case" is a discrete process that is addressed within a domain. For example, hospitals have clinical operations teams (a domain) that specialize in operating room efficiency (a use case). For each domain, we consider whether the use of AI will affect medical or administrative costs. We also note the position of each domain along the adoption curve. We define this as a typical technology S-curve—first developing solutions, then piloting, followed by scaling and adapting, and finally reaching maturity. In addition, we identify whether the processes affected are existing or new.

In addition, we provide a measure of impact on "total mission value." Health Care involves many nonfinancial factors, such as quality outcomes,

2. We recognize that many hospitals are part of broader health systems. In this paper, we use the term *hospital* to reference just that portion of a broader health system when applicable.

patient safety, patient experience, clinician satisfaction, and access to care. The combination of financial and nonfinancial factors is what we term total mission value.

2.3.1 Hospitals

2.3.1.1 Domain Breakdown

In our experience, AI-enabled use cases are emerging in nine domains: continuity of care, network and market insights, clinical operations, clinical analytics, quality and safety, value-based care, reimbursement, corporate functions, and consumer (figure 2.2).[3] Within clinical operations, for example, hospitals are focusing on use cases such as improving the capacity of the operating room, freeing up clinical staff time, and optimizing the supply chain (Luo et al. 2020; Kilic et al. 2020). Clinical analytics, with AI-enabled use cases such as clinical decision making or treatment recommendations, is another area of focus for hospitals, usually within specialties such as radiology (Allen et al. 2021). A key domain, and the focus of much academic research, is quality and safety. This includes AI-enabled use cases such as predicting the likelihood of condition deterioration, an adverse event, or a readmission (Bates et al. 2021).

Some domains, such as reimbursement and corporate functions, are more advanced in AI adoption than others. Key reasons for the variation in uptake among hospitals include organizational priority and need, availability of data, and the share of AI deployment in the total budget.

Consider the quality and safety domain. There have been only a few successful use cases, such as identifying sepsis early or the prediction of adverse events (Bates et al. 2021; Nemati et al. 2018; Cooley-Rieders and Zheng 2021). This is due in part to the need for a strong business case to launch a pilot. When an organization considers only financial factors, AI-enabled use cases usually do not meet the threshold for investment. The business cases for adopting AI become more compelling when the focus shifts to total mission value, which includes nonfinancial factors such as experience and access.

Five domains have a greater impact on administrative costs than on medical costs: continuity of care, network and market insights, value-based care, reimbursement, and corporate functions. These are further along in part because administrative costs are generally associated with processes that that are manual and repetitive, which AI is well suited to address. However, the overall opportunity is likely lower given that administrative costs represent a smaller portion of the total than medical costs do.

3. The consumer domain is not included in our estimates because AI-enabled use cases in this domain often lead to a zero-sum outcome between hospitals. Revenue for one organization is generally taken from another organization.

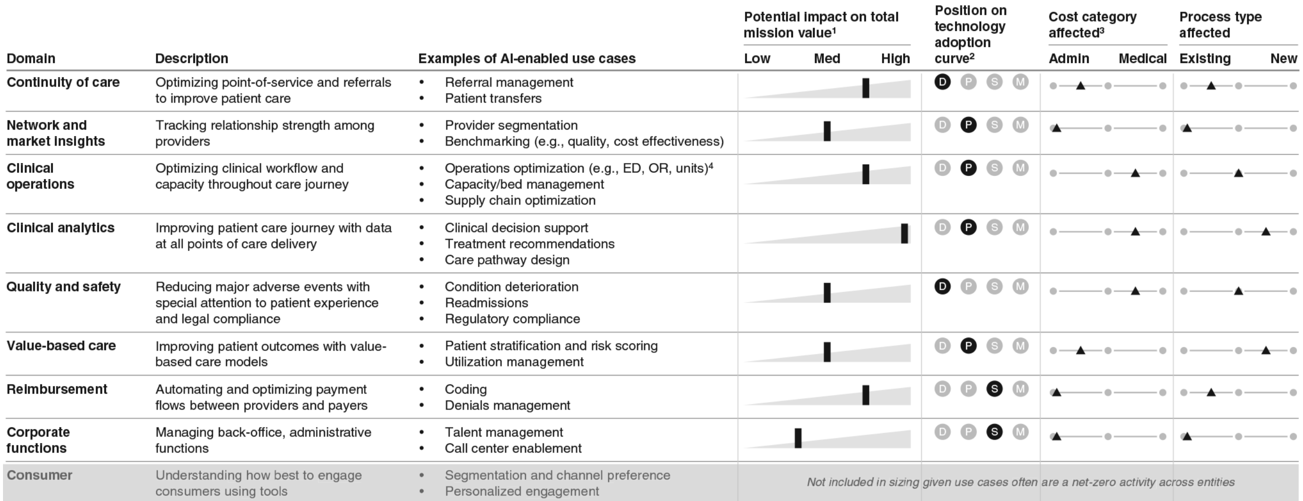


Fig. 2.2 Hospital AI domains and example use cases

1. We define "total mission value" as the combination of financial and nonfinancial factors, such as quality outcomes, patient safety, patient experience, clinician satisfaction, and access to care.

2. D = development of solutions; P = piloting; S = scaling and adapting; M = mature.

3. Positioning represents the direct cost category affected; second order effects may also reduce costs, but are not estimated.

4. ED = emergency department; OR = operating room.

Source: Authors' analysis

2.3.1.2 Case Study in Corporate Functions

During the COVID-19 pandemic, one multistate hospital's call center experienced a large increase in call volumes as patients sought more information on topics such as billing, COVID-19 tests, COVID-19 vaccines, scheduling and finding a clinician, searching for care, and getting a telemedicine visit. The hospital had not anticipated the increase in call volumes and was not staffed accordingly; nor did it have adequate existing call routing protocols. As a result, the hospital observed higher call wait times and dropped calls, both of which negatively affected patient experience and access to care.

Using NLP, which is well suited for tasks that have a consistent set of outcomes, the hospital created a virtual agent (a digital version of a customer service representative) for its mobile app and website. This virtual agent would answer common questions and route patient questions to a specific, prebuilt process with the appropriate supporting information. As a result of this rollout, call volume decreased by nearly 30 percent, the patient experience improved, and managers redeployed workers to less manual, more customized tasks such as answering calls related to upcoming and completed procedures.

2.3.1.3 Case Study in Clinical Operations

One large regional hospital was losing surgical volumes to other local hospitals. The surgical team investigated the OR schedule and found that while ORs appeared to be 100 percent blocked, actual utilization was about 60 percent. Reasons for this underuse included historical block-scheduling techniques that did not adjust time allocation based on surgeon demand, scheduled time slot durations that did not reflect actual surgical times (for example, some operations were scheduled for longer than they actually took, leaving unused time), and manual processes related to the release and reallocation of unused blocks.

Hospital leadership and frontline managers used AI to optimize the OR block scheduler—the system for assigning surgical time slots to surgeons—for more than 25 ORs and dozens of surgeons. An AI algorithm ingested historical block utilization and trends; forecast case hours by specialty, physician group, or surgeon; and rules related to procedural equipment needs, staffing, and surgeon availability. An optimization algorithm was then run to generate proposed schedules for a given week, using current schedules as a reference. To ensure acceptance of the results, the hospital worked with surgeons throughout the process, incorporating their insights into the algorithms. The solution has increased the amount of open time in the OR schedule by 30 percent, making it easier to treat patients with critical needs sooner. To sustain this progress, a data scientist was assigned responsibility for the algorithm and provides ongoing review of the output with the OR team for continued buy-in.

2.3.2 Physician Groups

2.3.2.1 Domain Breakdown

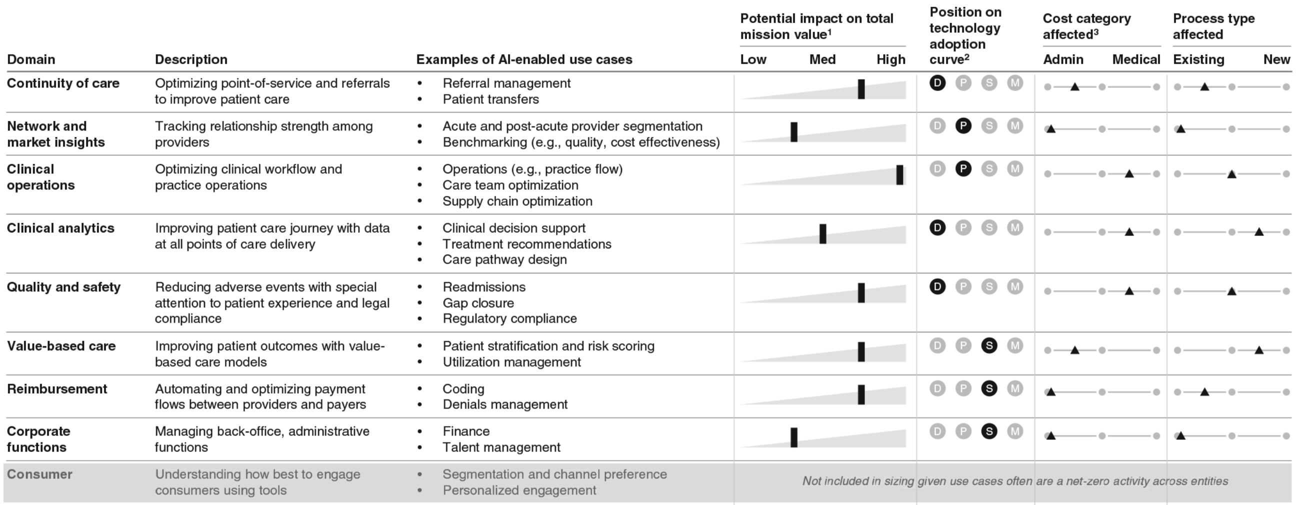
For physician groups, AI-enabled use cases are developing in the same nine domains as hospitals (figure 2.3). Within clinical operations, physician groups aim to reduce missed visits (patients failing to show up for a planned appointment) and ensure access to procedures by focusing on overall workflow, operations, access, and care team deployment. For example, understanding which patients might miss an appointment or need support with transportation influences how the clinical team conducts outreach to patients and the overall schedule of the physician group. Quality and safety is another domain of focus, especially for physician groups in value-based arrangements, where quality and safety outcomes directly affect financial performance. For example, AI supporting a value-based arrangement may predict which patients are at higher risk for readmission, therefore enabling care team members to intervene and address a patient's care needs to prevent deterioration in the condition. As these payment models grow in acceptance, particularly in primary care, physician groups are increasingly focusing on overall population health management use cases within the value-based care domain.

In terms of AI adoption, some domains are more mature than others. Those further along reflect the impact of market forces on physician group economics and the transition to value-based arrangements. For example, the clinical operations domain is more mature, given how central it is to a physician group's economics, patient access, and patient and clinician experience. In contrast, continuity of care, an important aspect of care management, is less mature given the fragmented nature of data across providers. However, new interoperability application programming interfaces (APIs), which enable the exchange of data between two organizations—such as two providers or a payer and a provider—are making it easier to exchange data in standard formats.

As with hospitals, five of the domains have a greater impact on administrative costs than on medical costs: continuity of care, network and market insights, value-based care, reimbursement, and corporate functions. These five domains are generally further along the adoption curve but likely have smaller total impact. Further, AI-enabled use cases in these domains tend to address existing processes. Future adoption in these domains is tied to market trends—such as the growth of value-based arrangements mentioned above—and the increase in vendors who can serve physician groups.

2.3.2.2 Case Study in Value-Based Care

One large physician group was in a value-based arrangement for a single chronic disease and searching for innovative ways to manage the total cost



Domain	Description	Examples of AI-enabled use cases
Continuity of care	Optimizing point-of-service and referrals to improve patient care	• Referral management • Patient transfers
Network and market insights	Tracking relationship strength among providers	• Acute and post-acute provider segmentation • Benchmarking (e.g., quality, cost effectiveness)
Clinical operations	Optimizing clinical workflow and practice operations	• Operations (e.g., practice flow) • Care team optimization • Supply chain optimization
Clinical analytics	Improving patient care journey with data at all points of care delivery	• Clinical decision support • Treatment recommendations • Care pathway design
Quality and safety	Reducing adverse events with special attention to patient experience and legal compliance	• Readmissions • Gap closure • Regulatory compliance
Value-based care	Improving patient outcomes with value-based care models	• Patient stratification and risk scoring • Utilization management
Reimbursement	Automating and optimizing payment flows between providers and payers	• Coding • Denials management
Corporate functions	Managing back-office, administrative functions	• Finance • Talent management
Consumer	Understanding how best to engage consumers using tools	• Segmentation and channel preference • Personalized engagement

Fig. 2.3 Physician group AI domains and example use cases

1. We define "total mission value" as the combination of financial and nonfinancial factors, such as quality outcomes, patient safety, patient experience, clinician satisfaction, and access to care.

2. D = development of solutions; P = piloting; S = scaling and adapting; M = mature.

3. Positioning represents the direct cost category affected; second-order effects may also reduce costs, but are not estimated.

Source: Authors' analysis

of care while improving outcomes for its patients. To meet those goals, the organization identified reducing preventable complications as an opportunity area. The organization observed that about 10 percent of patients were admitted to the hospital on a monthly basis. Using AI, the physician group developed a risk model to assess likelihood of unplanned admission. The AI application ingested data from several sources (for example, electronic health records, lab results, demographics, risk scores, and health information exchange admission, discharge, and transfer feeds) to develop the model and understand the main variables influencing unplanned admission. The initial model showed a potential decrease of several percentage points in inpatient spending due to better care management. The physician group is now planning to deploy the algorithm more broadly based on the prototype models. As a result, members of the care team will be able to better prioritize their outreach to patients, more efficiently using their time to improve patient outcomes. To operationalize this, the physician group is creating new clinical workflows that are helping the care team better focus their attention and resources.

2.3.3 Private Payers

2.3.3.1 Domain Breakdown

In our experience, AI-enabled use cases are emerging in six domains for private payers: health care management, provider relationship management, claims management, member services, corporate functions, and marketing and sales (figure 2.4).[4] Within health care management, private payers are focusing on care management, medical and clinical utilization and spending, and quality AI-enabled use cases. For example, private payers are attempting to predict behavioral health needs to better match patients with support resources and seeking to improve care management programs that help prevent avoidable readmissions. Another example is claims management, where private payers are using AI to improve auto-adjudication rates; predict and improve prior authorization outcomes to enable greater access to care; and prevent fraud, waste, and abuse. Further, the provider relationship management domain focuses on designing networks that enable better quality outcomes and access in a cost-effective way for members.

Adoption of AI varies across these domains. Claims management and corporate functions generally are more mature in their adoption of AI. Many use cases, such as processing a prior authorization or adjudicating a claim, are largely repetitive processes that are best suited for AI.

Three domains have more impact on administrative costs than on medi-

4. The marketing and sales domain is not included in our estimates because AI-enabled use cases in this domain often lead to a zero-sum outcome between private payers. Revenue for one organization is generally taken from another organization.

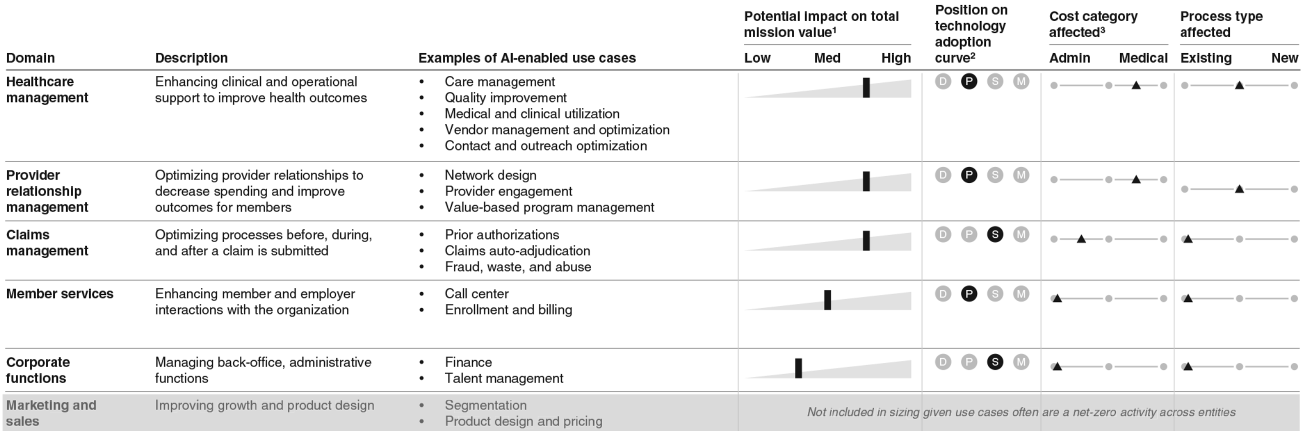


Fig. 2.4 Private payer AI domains and example use cases

1. We define "total mission value" as the combination of financial and nonfinancial factors, such as quality outcomes, patient safety, patient experience, clinician satisfaction, and access to care.

2. D = development of solutions; P = piloting; S = scaling and adapting; M = mature.

3. Positioning represents the direct cost category affected; second-order effects may also reduce costs, but are not estimated.

Source: Authors' analysis

cal costs: claims management, member services, and corporate functions. The opportunity in these domains is substantial but less than for medical costs, given that administrative costs are a smaller portion of total costs. For the domains that are focused on medical cost, there are also large non-financial opportunities, including improving health, quality, and member experience. In general, use cases tend to be focused on existing processes across all domains.

2.3.3.2 Case Study in Claims Management

One large private payer, experiencing high costs and conducting an overall effort to improve its financial position, assessed areas for improvement in claims management. The analysis concluded that the organization could replace existing manual processes with AI to address fraud, waste, and abuse (FWA) among providers. As a result, a team built an AI classification model to identify potential FWA based on prior patterns observed in several years of claims data. The output of the model was a list of providers for further investigation. The team could then manually validate the list and determine next steps. This AI-enabled identification model allowed the payer to streamline operations and inform efforts that resulted in the reduction of medical costs by about 50 basis points. The payer's FWA team maintains the AI model on an ongoing basis.

2.3.3.3 Case Study in Health Care Management

To improve patient outcomes, a private payer focused on how to reduce the readmissions rate for its most vulnerable members. To address these readmissions, the organization developed an ML model that ingested a variety of claims and member demographic information. The output identified which patients were most likely to have a readmission, quantified the differences between these patients and those who did not have a readmission, and identified which parts of the care journey were linked to the readmission. The private payer then used the output to inform core business processes such as care management outreach. For example, the organization created a specialized outreach team of care managers who used the output to prioritize tactics for these vulnerable patients. As a result of this ML model and associated personalized marketing techniques, about 70 percent more members connected with their care managers compared with previous efforts that did not use the model. Follow-up visits with primary care physicians within 30 days of discharge increased by about 40 percent, and the all-cause readmission rate decreased by about 55 percent for this cohort.

2.4 Opportunity Size

Based on the domains discussed above, we have estimated the annual net savings that AI could create for US health care in the next five years.

Net savings is defined as total gross savings less annual expenses to operate AI. We derive the savings estimates for each domain from our experience working with health care organizations; there are few experimental studies of the impact of AI on costs or outcomes to inform our analysis. All savings estimates are based on the use of technologies available today and assume that adoption reaches full scale.

To estimate the total AI opportunity, we first estimate the revenue for each stakeholder group from 2019 National Health Expenditure data. Using McKinsey's proprietary value pool data, we subtract each stakeholder group's total earnings before interest and taxes (EBIT), leaving total costs. For hospitals and physician groups, we estimate three cost categories: administrative costs, medical costs associated with labor (for example, clinicians), and nonlabor medical costs (for example, diagnostics and supplies). For private payers, we estimate two cost categories: administrative costs and medical costs.

With this baseline, each AI domain described in the previous section is then aligned to a cost category. Based on our experience, we estimate a gross savings percentage for each domain. We break down what portion of these savings will affect administrative or medical costs. One key adjustment is converting gross savings to net savings, which represents the expense needed to maintain AI. Based on our experience, we model labor and technology maintenance expenses for each stakeholder group. The total amount is then subtracted from gross savings to estimate a net savings range.

We then multiply these percentages by the dollar values in each cost category to estimate a net savings value for each domain. Summing the estimated savings for each domain results in the total net savings opportunity for a stakeholder group. Figure 2.5 shows an example of the quality and safety domain for hospitals. We begin with total hospital revenue as reported in

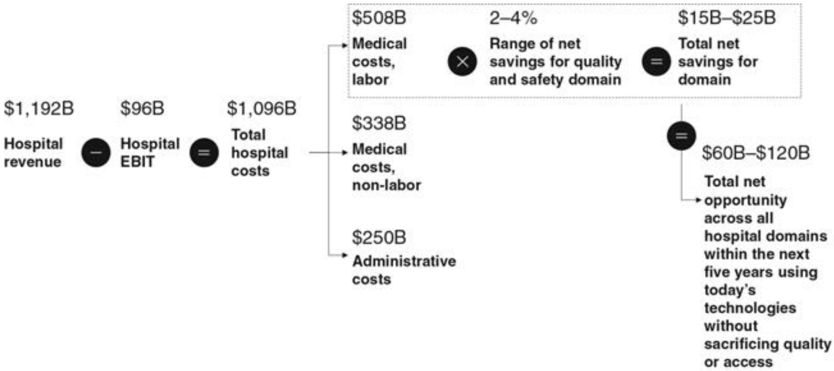


Fig. 2.5 Example of a hospital domain calculation: Quality and safety

Note: All data in 2019 dollars.

Source: National Health Expenditures data; authors' analysis the National Health Expenditure data of $1,192 billion in 2019. We subtract hospital EBIT to estimate a total for hospital costs. Total hospital costs are then broken into three cost categories. The quality and safety domain largely affects labor within medical costs, which we estimate to be $508 billion in 2019. Using the net savings rate for this domain (after accounting for the gross-to-net conversion), total net savings is $15 billion to $25 billion. Repeating this for all the domains, we estimate a total annual net savings opportunity for hospitals of $60 billion to $120 billion within the next five years using today's technologies without sacrificing quality or access.

To consider the full AI opportunity in health care, we also include public payers and other sites of care such as dentists and home health. For public payers, we begin with the AI opportunity estimate for private payers, which have several similar functions and operations. Referencing previous research, we estimate the total costs to be about 45 percent of those for private payers (Sahni et al. 2021). We further assume the savings opportunity would be about three-quarters that of private payers given that public payers do not undertake all the same functions to the same extent, such as provider relationship management and health care management. For other sites of care, we begin with the AI opportunity estimate for physician groups. Similarly referencing previous research, we estimate the total costs to be about 115 percent of those for physician groups (Sahni et al. 2021). We further assume the savings opportunity would be about half that of physician groups given differences in patient acuity, clinical staff mix (for example, less clinician time per clinical episode), and fewer applicable AI domains.

Our estimates do not include one-time implementation costs, which in our experience are 1.0 to 1.5 times the annual net savings. One-time implementation costs relate directly to building an AI-enabled use case, which includes hiring specialized talent, creating incremental infrastructure or computing power, and aggregating and cleaning the necessary data. One-time implementation costs do not include large investments such as new underlying core technology or last-mile change management, both of which could be necessary and can vary greatly by organization.

2.4.1 Hospitals

In 2019 dollars, total costs for hospitals are about $1,096 billion, of which 80 percent is medical and 20 percent is administrative. With about 6,000 hospitals nationally, this is a fragmented market. The top 10 hospital systems accounted for about 18 percent of admissions in 2017 (Sahni et al. 2019). Types of facilities include community hospitals and academic medical centers. The typical hospital has an "all-payer margin" of about 6 to 7 percent (Medicare Payment Advisory Commission 2022).

Based on our calculations, hospitals employing AI-enabled use cases could achieve total annual run-rate net savings of $60 billion to $120 billion (roughly 4 to 10 percent of total costs for hospitals) within the next five years using today's technologies without sacrificing quality or access. Clinical operations—encompassing emergency room and inpatient care, capacity and workflow, diagnostics, supply chain, and clinical workforce management—and quality and safety are the primary drivers of this opportunity. About 40 percent of total savings would come from reducing administrative costs (roughly 9 to 19 percent of this cost category), with the remaining 60 percent from reducing medical costs (roughly 4 to 8 percent of this cost category). About 45 percent of total savings would come from simplifying existing processes, with the remaining 55 percent from creating new processes.

2.4.2 Physician Groups

In 2019 dollars, total costs for physician groups are about $711 billion, of which 70 percent is medical and 30 percent administrative. The physician group landscape is fragmented, with about 125,000 groups nationally, including those employed by hospitals, owned by private organizations, or independent (Sahni et al. 2021).

Based on our calculations, physician groups employing AI-enabled use cases could achieve total annual run-rate net savings of $20 billion to $60 billion (roughly 3 to 8 percent of total costs for physician groups) within the next five years using today's technologies without sacrificing quality or access. The main domain of opportunity, similar to hospitals, is clinical operations, with a focus on outpatient operations and access, supply chain, and clinical workforce management. About 50 percent of total savings would come from reducing administrative costs (roughly 4 to 14 percent of this cost category), with the remaining 50 percent from reducing medical costs (roughly 2 to 6 percent of this cost category). About 45 percent of total savings would come from simplifying existing processes, with the remaining 55 percent from creating new processes.

2.4.3 Private Payers

In 2019 dollars, total costs for private payers are about $1,135 billion, of which 85 percent is medical and 15 percent administrative. In 2017, the top five private payers plus Medicare (Part A/B only) and Medicaid (fee-for-service only) accounted for about 58 percent of covered lives, and the 350-plus other private payers covered the remaining 42 percent (Sahni et al. 2019). Types of private payers include national, regional, and local for-profit and not-for-profit organizations.

Based on our calculations, private payers could achieve total annual run-rate net savings of $80 billion to $110 billion (roughly 7 to 9 percent of total costs for private payers) within the next five years using today's technologies without sacrificing quality or access. The primary domains of opportunity are health care management (including care management and avoidable readmissions), claims management (including FWA identification, prior authorizations, and adjudication), and provider relationship management (including network design, value-based care, and provider directory management). About 20 percent of total savings would come from reducing administrative costs (roughly 8 to 14 percent of this cost category), with the remaining 80 percent from reducing medical costs (roughly 6 to 9 percent of this cost category). About 55 percent of total savings would come from simplifying existing processes, with the remaining 45 percent from creating new processes.

2.4.4 Overall

With these estimates, we then scale the savings to the entire US health care industry (table 2.1). In 2019 dollars, we estimate the annual run-rate net savings to be $200 billion to $360 billion within the next five years using today's technologies without sacrificing quality or access. This would amount to a 5 to 10 percent overall reduction in US health care spending. AI adoption could also create nonfinancial benefits such as improved health care quality, increased access, better patient experience, and greater clinician satisfaction. (In this paper, we do not offer an estimate of these nonfinancial benefits.)

Administrative costs could be reduced by 7 to 14 percent, roughly $65 billion to $135 billion annually. This is about 35 percent of total savings. The remaining 65 percent could reduce medical costs by 5 to 8 percent, roughly $130 billion to $235 billion annually. The overall AI opportunity is divided nearly equally between simplifying existing processes and creating new processes.

2.5 Adoption Challenges

Despite the large opportunity, the AI adoption rate in health care has lagged behind that in other industries (Cam, Chui, and Hall 2019). Generally, technology adoption follows an S-curve—first developing solutions, then piloting, followed by scaling and adapting, and finally reaching maturity. Other industries have already reached the final stage of the S-curve; for example, financial services companies deploy sophisticated AI algorithms for fraud detection, credit assessments, and customer acquisition. Mining companies use AI to boost output, reduce costs, and manage the environmental impact of new projects. Retailers use AI to predict which goods will interest a customer based on the customer's shopping history.

Across nearly all the domains identified in section 2.2, AI adoption in health care is at an earlier stage of the S-curve. There are several possible reasons for this. Many economists believe that AI is underused because the health care payment system does not provide incentives for this type of innovation. Another view is that management barriers, both at the organizational and industry level, are responsible for slower adoption in health care.

Table 2.1 Breakdown of overall AI net savings opportunity within next five years using today's technology without sacrificing quality or access

Stakeholder group	Total costs (2019), $ billions	Net savings opportunity (2019), $ billions	Net savings opportunity as percent of stakeholder group's total costs	Percentage of net savings opportunity focused on administrative costs
Hospitals	$1,096	$60–$120	5–11%	~40%
Physician groups	$711	$20–$60	3–8%	~50%
Private payers	$1,135	$80–$110	7–10%	~20%
Public payers	$511	$30–$40	5–7%	~20%
Other sites of care	$817	$10–$30	1–4%	~50%
Total		**$200–$360**	**5–10%**[a]	**~35%**

Source: National Health Expenditures data; authors' analysis.

[a] This represents the percent of total national health spending in 2019.

In this paper, we do not settle the debate about whether better incentives will lead to greater adoption of AI. Rather, we discuss the managerial difficulties in bringing AI to bear in health care. Even if the right payment models were in place, organizations would still need to overcome challenges such as legacy technology, siloed data, nascent operating models, misaligned incentives, industry fragmentation, and talent attraction (Goldfarb and Teodoridis 2022; Henke et al. 2016).

In our experience, private payers are further along the AI adoption curve than other health care organizations, although larger national private payers with greater resources are more advanced in their use of AI compared with smaller regional private payers that may face resource and talent attraction challenges. Hospitals have piloted AI and are beginning to scale adoption in some domains, with larger hospitals having done more than smaller hospitals. Most physician groups are at the beginning of their journey (unless employed by hospitals).

In this section, we discuss specifics about what is needed for AI adoption in health care. We break these down into "within" and "between/seismic" factors. "Within" factors are those that can be controlled and implemented by individual organizations. "Between" factors require collaboration between organizations but not broader, industry-wide change, and "seismic" factors require broad, structural collaboration across the US health care industry (Sahni et al. 2021).

2.5.1 "Within" Challenges

In our experience, successful AI adoption depends on six factors (figure 2.6). These are the same for all organizations across industries, though some of the underlying challenges are specific to health care.

The first factor is a **mission-led roadmap**. The roadmap should offer a clear view of value, link to business objectives and mission, and be sequenced for implementation. A key challenge is ensuring the end state is a transformative view of the organization, not incremental. Each AI-enabled use case should be quantified, which presents additional hurdles for health care organizations because value extends into nonfinancial factors such as quality outcomes, patient safety, patient experience, clinician satisfaction, and access to care. As noted above, we refer to this combination of financial and nonfinancial factors as total mission value. In our experience, the most successful organizations rely on strong collaboration between business and technology leaders to develop and implement this roadmap.

A second factor is **talent**. Organizations must ensure that the right skills and capabilities are available across the organization. Talent shortages are common, especially in AI (Zwetsloot, Heston, and Arnold 2019). Many organizations have addressed these shortages by establishing talent hubs, sometimes in a different city with operations than headquarters, but many health care organizations face the additional challenge of being inherently

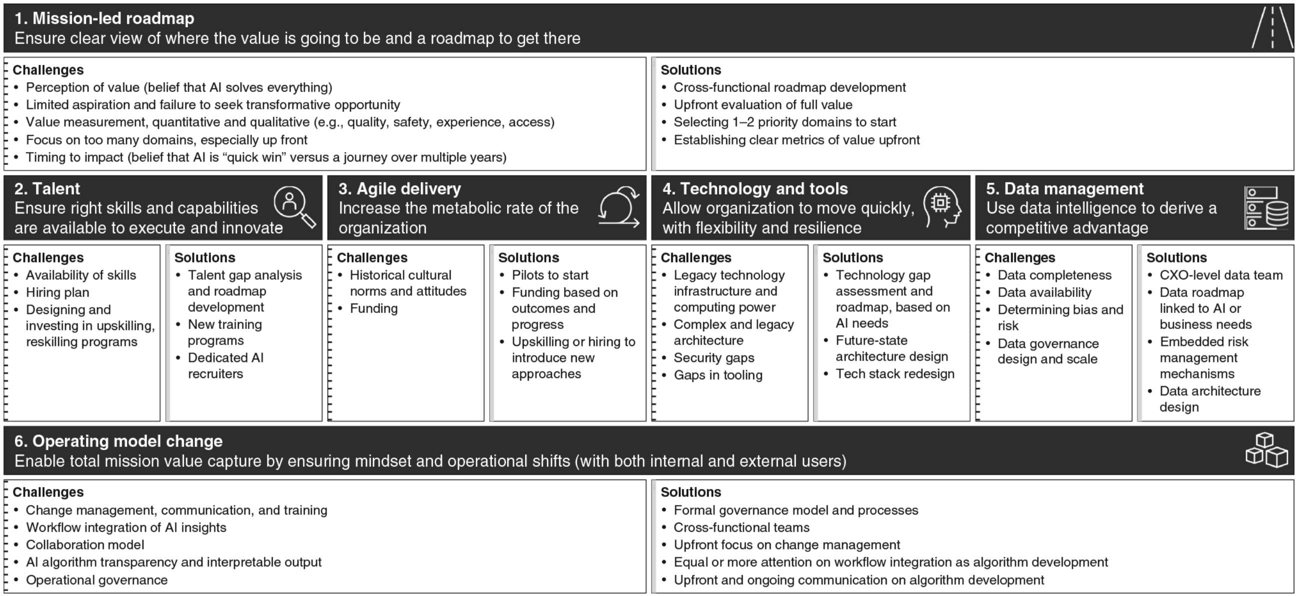


Fig. 2.6 Factors for successful AI adoption with associated challenges and solutions within health care organizations

Source: Carey et al. (2021); Rajkomar et al. (2019); Bates et al. (2020); Shaw et al. (2019); Singh et al. (2020); He et al. (2019); Authors' analysis.

local. Still, some are experimenting with ways to make this work—for example, by centralizing talent in a nearby location or using remote work options.

Agile delivery, or accelerating an organization's decision-making and delivery processes, is a new approach for many health care organizations. Changing the culture to move away from historical processes and ways of working is a challenge for organizations in all industries. It is an especially large hurdle in health care, where culture is often more deeply rooted than in other industries, and where clinicians are justifiably concerned that the process of change might harm patients. In our experience, organizations that empower small, integrated agile teams are more likely to have successful AI deployments.

Enabling agile delivery requires **technology and tools** that are flexible, scalable, secure, and resilient. Organizations in all industries confront complex legacy IT environments. This is particularly true in health care given the relatively low levels of investment in technology and high levels of customization. In our experience, successful deployments generally overinvest in the enablers of AI, such as core technology architecture and data systems.

Data management, or the use of data to derive a competitive advantage, is often overlooked in AI deployments, though it is one of the most critical factors. Organizations in all industries face key challenges with data fragmentation and quality. The challenge is even greater in health care given the large number of systems, general lack of interoperability, and data privacy and usage requirements. In our experience, the most successful AI deployments establish a dedicated function to manage all data at the beginning of any adoption journey.

Finally, establishing the right **operating model** is key.[5] Such a model enables an organization to capture full mission value by encouraging mindset and operational shifts among both internal and external users. Determining the right operating model is difficult in any industry, and the number of stakeholders, need for change management with providers, and heightened attention to security and model risk increase the challenge in health care. In our experience, organizations that deploy more central structures to build capabilities, consistency, and rigor from the beginning position themselves for more successful AI deployments, while setting up the operating model to work closely with their business partners.

While it is critical for organizations to pursue all six factors, it is just as important to foster "digital trust" among individuals—to inspire confidence that the organization effectively protects data, uses AI responsibly, and provides transparency. Building this trust requires organizations to establish the right controls, processes, and risk management. Without digital trust and

5. *Operating model* encompasses a number of components about an organization, including structure, governance, and processes.

the responsible use of AI, health care organizations may experience greater scrutiny and a slower pace of use-case scaling.

Investing in addressing these challenges is critical. Across industries, the highest performers spent 30 to 60 percent more than others when adopting technologies such as AI and expect to increase their budgets 10 to 15 percent over the following year. Meanwhile, lesser performers report small or no increases (D'Silva and Lawler 2022).

2.5.2 "Between" and "Seismic" Challenges

Even if a health care organization successfully deploys AI, it will face ongoing industry-level challenges—factors that are out of the organization's control and can hinder widespread adoption. These include data heterogeneity, lack of patient confidence, ongoing adaptability, the ability to capture productivity gains, and regulatory challenges (figure 2.7).

These industry-level challenges take two forms: social and technology. By *social challenge* we mean one in which the industry would need to encourage stakeholders such as physicians to adopt the same approach, process, or standard. By *technology challenge* we mean one in which the hurdle to adoption relates to the need for a technology solution.

Data heterogeneity in health care takes many forms. In industries with greater AI adoption, most data are structured. In health care, by contrast, large portions of key data are unstructured, existing in electronic health records. Clinical notes, the clinician's recording of a patient's response to a particular treatment, are one example. Further, these data exist in multiple sources, often with limited ways to connect disparate pieces of information for an individual patient (Kruse et al. 2016).

		Types of challenge[1]	
Challenge	**Description**	**Social**	**Technology**
Data heterogeneity	Large portions of needed healthcare data are unstructured, spread across multiple data sources, and stored in varying data structures		✓
Lack of patient confidence	Patients lack confidence in output due to concerns about the privacy of their data, potential biases in data affecting AI outputs, uncertainty about methodology, and clarity of the reports	✓	
Ongoing adaptability	Once launched, AI models may be slow to adapt or integrate new data when released		✓
Ability to capture productivity gains	Once AI frees up capacity of clinicians or assets such as operating rooms, it may not be applied to increase productivity	✓	
Regulatory challenges	Evolving regulations that could increase adoption may require approval processes for validation for organizations such as CMS or FDA	✓	✓

Fig. 2.7 Industry-level challenges by type

1. By *social challenge*, we mean one in which the industry would need to encourage stakeholders such as physicians to adopt the same approach, process, or standard. By *technology challenge*, we mean one in which the hurdle to adoption relates to the need for a technology solution.

Source: Authors' analysis

Patient confidence in AI output is also critical to the integration of information into the clinical workflow. One issue is privacy. Patients may worry about how their data are being used and prevent the application of AI for their medical needs. Another concern is whether AI output can be trusted. There are many examples of biases in algorithms, and patients may not trust AI-generated information even if a clinician validates it. There are also methodological concerns such as validation and communication of uncertainty, as well as reporting difficulties such as explanations of assumptions (Bates et al. 2020; Shaw et al. 2019; Singh et al. 2020; He et al. 2019).

In addition, questions arise about whether AI-enabled use cases would cement certain biases in existing data and be slow to respond to new types of data. For example, an organization using AI to help define clinical treatment pathways might need to control for biases in existing treatment recommendations and determine how to remove them. In addition to adding AI ethics to model development, organizations are addressing bias by creating synthetic data—manufactured data designed to train a model on a certain set of inputs, similar to real-world data. Further, as health care generates new data, these changes may require previously developed models to be refreshed.

Many clinicians and health care executives are optimistic that AI could address ongoing productivity challenges in health care. Historical analyses have shown negative labor productivity growth in health care and a likelihood that clinician shortages will continue (Sahni et al. 2019; Berlin et al. 2022). If adopted appropriately, AI could free up clinician capacity. The question arises, though, whether clinicians will use the excess capacity to see more patients or to complete nonclinical tasks.

Finally, in the United States, regulations generally focus on protecting the patient, given the private and sensitive nature of each person's data. But regulation also plays other roles in AI. For example, Medicare and Medicaid are beginning to reimburse for AI applications, though adoption is still in the early stages. This is unique to health care; organizations in other industries have to pay for AI themselves. Validation that algorithms are clinically robust and safe is another issue. For example, the Food and Drug Administration established standards for evaluating software as a medical device and AI-enabled medical devices. Dozens of AI products have since received approval, the majority in the past five years. Examples of digitally enabled therapeutics include those for treating type 2 diabetes and substance use disorder. This type of industry-level change could provide greater confidence for patients and clinicians using AI.

2.6 Changes That May Improve AI Adoption

Based on our experience, fewer than 10 percent of health care organizations today fully integrate AI technologies into their business processes. But the benefits of doing so are meaningful: in our experience, organizations that deploy AI have twice the five-year revenue compound annual growth rate compared with others that do not. With a $200 billion to $360 billion opportunity in health care and such a small subset of organizations capturing the potential, what might the future hold? Will AI adoption accelerate?

Several trends suggest the tide may soon turn. First, the COVID-19 pandemic, coupled with rising inflation and labor shortages, is straining the finances of health care organizations (Singhal and Patel 2022). For example, all seven of the largest publicly traded payers have announced productivity improvement programs in the past few years. Further, research shows that the most successful organizations coming out of a recession generally have run larger productivity improvement programs (Görner et al. 2022). This could be a boon for the adoption of AI-enabled use cases—especially use cases that focus on administrative costs, which are usually passed over in favor of a focus on medical costs.

A second trend is the flow of investment into AI technologies, even in today's uncertain macroeconomic climate. From 2014 to 2021, the overall number of venture capital–backed health care AI start-ups increased more than fivefold. Over the same period, the number of private equity deals for health care AI organizations increased more than threefold.

At the organizational level, there are indications that the C-suite's appreciation for the potential of technologies like AI is growing. For example, nearly all of the top 15 private payers have a designated chief analytics or chief data officer. Dozens of hospitals do as well, including most of the largest in the United States. This elevation of business importance suggests that more AI deployments may be on the way.

At the industry level, evolving regulations may enable the creation of new data sets that feed AI. For example, recently introduced medical price transparency regulations promise to increase the availability of hospital and private payer data. Alone, these data may not be good enough for AI algorithms to generate insights; however, if coupled with other data sets, such as member or census data, they could accelerate the adoption of AI. In addition, the Centers for Medicare and Medicaid Services has been developing interoperability rules and APIs that require data to be made available in a consistent structure to be exchanged across organizations.

2.7 Conclusions

The promise of AI in health care has been a topic of discussion in the industry for more than a decade. But its potential has not been quantified systematically, and adoption has been lacking. We estimate that AI in health care offers a $200 billion to $360 billion annual run-rate net savings opportunity that can be achieved within the next five years using today's technologies without sacrificing quality or access. These opportunities could also result in nonfinancial benefits such as improved health care quality, increased access, better patient experience, and greater clinician satisfaction. As our case studies highlight, both the challenges to adoption and actionable solutions are becoming better understood as more organizations pilot AI. Recent market trends also suggest that AI in health care may be at a tipping point.

Still, taking full advantage of the savings opportunity will require the deployment of many AI-enabled use cases across multiple domains. Ongoing research and validation of these use cases is needed. This could include conducting randomized control trials to prove the impact of AI in clinical domains to increase confidence for broader deployment. However, given that these studies will likely require a long timeline, additional work focused on case studies of successful deployments may provide greater evidence for organizations to overcome internal inertia in the near term. Finally, an independent third party could create a central data repository of AI deployments—both successful and unsuccessful—which would allow for more robust econometric analyses to inform rapid scaling.

As other industries have shown, AI as a technology could have an outsized financial and nonfinancial impact in health care, enabling patients to receive better care at a lower cost. The next few years will determine whether this promise becomes a reality.

References

Allen, B., S. Agarwal, L. Coombs, C. Wald, and K. Dreyer. 2021. "2020 ACR Data Science Institute Artificial Intelligence Survey." *Journal of the American College of Radiology* 18 (8): 1153–59.

Bates, D. W., A. Auerbach, P. Schulam, A. Wright, and S. Saria. 2020. "Reporting and Implementing Interventions Involving Machine Learning and Artificial Intelligence." *Annals of Internal Medicine* 172 (11 Suppl): S137–S144.

Bates, D. W., D. Levine, A. Syrowatka, M. Kuznetsova, K. J. T. Craig, A. Rui, G. Purcell Jackson, and K. Rhee. 2021. "The Potential of Artificial Intelligence to Improve Patient Safety: A Scoping Review." *NPJ Digital Medicine* 4 (1): 54.

Berlin, G., M. LaPointe, M. Murphy, and J. Wexler. 2022. "Assessing the Lingering Impact of COVID-19 on the Nursing Workforce." McKinsey & Company, May 11, 2022. https://www.mckinsey.com/industries/healthcare/our-insights/assessing-the-lingering-impact-of-covid-19-on-the-nursing-workforce.

Berwick, D. M., and A. D. Hackbarth. 2012. "Eliminating Waste in US Health Care." *Journal of the American Medical Association* 307 (14): 1513–16.

Cam, A., M. Chui, and B. Hall. 2019. "Global AI Survey: AI Proves Its Worth, but Few Scale Impact." McKinsey & Company, November 22, 2019. https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact.

Carey, D., R. Charan, E. Lamarre, K. Smaje, and R. Zemmel. 2021. "The CEO's Playbook for a Successful Digital Transformation." *Harvard Business Review*, December 20, 2021.

Cooley-Rieders, K., and K. Zheng. 2021. "Physician Documentation Matters. Using Natural Language Processing to Predict Mortality in Sepsis." *Intelligence-Based Medicine* 5: 100028.

D'Silva, V., and B. Lawler. 2022. "What Makes a Company Successful at Using AI?" *Harvard Business Review*, February 28, 2022.

Goldfarb, A., and F. Teodoridis. 2022. "Why Is AI Adoption in Health Care Lagging?" *Brookings*, March 9, 2022. https://www.brookings.edu/articles/why-is-ai-adoption-in-health-care-lagging/.

Görner, S., A. Govindarajan, A. Panas, E. Greenberg, A. Kelkar, J. Kelleher, I. Kristensen, L. Liu, A. Padhi, and Z. Silverman. 2022. "Something's Coming: How US Companies Can Build Resilience, Survive a Downturn, and Thrive in the Next Cycle." McKinsey & Company, September 16, 2022. https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/somethings-coming-how-us-companies-can-build-resilience-survive-a-downturn-and-thrive-in-the-next-cycle

He, J., S. L. Baxter, J. Xu, J. Xu, X. Zhou, and K. Zhang. 2019. "The Practical Implementation of Artificial Intelligence Technologies in Medicine." *Nature Medicine* 25 (1): 30–36.

Henke, N., J. Bughin, M. Chui, J. Manyika, T. Saleh, B. Wiseman, and G. Sethupathy. 2016. "The Age of Analytics: Competing in a Data-Driven World." McKinsey & Company, December 7, 2016. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-age-of-analytics-competing-in-a-data-driven-world

Kilic, A., A. Goyal, J. K. Miller, E. Gjekmarkaj, W. L. Tam, T. G. Gleason, I. Sultan, and A. Dubrawksi. 2020. "Predictive Utility of a Machine Learning Algorithm in Estimating Mortality Risk in Cardiac Surgery." *Annals of Thoracic Surgery* 109 (6): 1811–19.

Kruse, C. S., R. Goswamy, Y. Raval, and S. Marawi. 2016. "Challenges and Opportunities of Big Data in Health Care: A Systematic Review." *JMIR Medical Informatics* 4 (4): e38.

Luo, L., F. Zhang, Y. Yao, R. Gong, M. Fu, and J. Xiao. 2020. "Machine Learning for Identification of Surgeries with High Risks of Cancellation." *Health Informatics Journal* 26 (1): 141–55.

Medicare Payment Advisory Commission. 2022. *March 2022 Report to the Congress: Medicare Payment Policy*. March 15, 2022. https://www.medpac.gov/wp-content/uploads/2022/03/Mar22_MedPAC_ReportToCongress_v3_SEC.pdf.

Nemati, S., A. Holder, F. Razmi, M. D. Stanley, G. D. Clifford, and T. G. Buchman. 2018. "An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU." *Critical Care Medicine* 46 (4): 547–53.

Rajkomar, A., J. Dean, and I. Kohane. 2019. "Machine Learning in Medicine." *New England Journal of Medicine* 380 (14): 1347–58.

Sahni, N. R., P. Kumar, E. Levine, and S. Singhal. 2019. "The Productivity Imperative for Healthcare Delivery in the United States." McKinsey & Company, February 27, 2019. https://www.mckinsey.com/industries/healthcare/our-insights/the-productivity-imperative-for-healthcare-delivery-in-the-united-states.

Sahni, N. R., P. Mishra, B. Carrus, and D. M. Cutler. 2021. "Administrative Simplification: How to Save a Quarter-Trillion Dollars in US Healthcare." McKinsey & Company, October 20, 2021. https://www.mckinsey.com/industries/healthcare

/our-insights/administrative-simplification-how-to-save-a-quarter-trillion-dollars-in-us-healthcare.
Shaw, J., F. Rudzicz, T. Jamieson, and A. Goldfarb. 2019. "Artificial Intelligence and the Implementation Challenge." *Journal of Medical Internet Research* 21 (7): e13659.
Singh, R. P., G. L. Hom, M. D. Abramoff, J. P. Campbell, and M. F. Chiang. 2020. "Current Challenges and Barriers to Real-World Artificial Intelligence Adoption for the Healthcare System, Provider, and the Patient." *Translational Vision Science & Technology* 9 (2): 45.
Singhal, S., and N. Patel. 2022. "The Future of US Healthcare: What's Next for the Industry Post-COVID-19." McKinsey & Company, July 19, 2022. https://www.mckinsey.com/industries/healthcare/our-insights/the-future-of-us-healthcare-whats-next-for-the-industry-post-covid-19.
Society of Automotive Engineers. 2021. *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. https://www.sae.org/standards/content/j3016_202104/.
Zwetsloot, R., R. Heston, and Z. Arnold. 2019. *Strengthening the U.S. AI Workforce: A Policy and Research Agenda*. Georgetown Center for Security and Emerging Technology. September 2019. https://cset.georgetown.edu/wp-content/uploads/CSET_US_AI_Workforce.pdf.

Comment David C. Chan Jr.

In their very interesting piece, Sahni et al. estimate the potential impact of artificial intelligence (AI) on healthcare spending. As the authors note, AI has the potential to create more efficient processes and to improve decision making. These potential impacts could lead to productivity improvements, reducing the costs of delivering healthcare while improving outcomes.

The authors bring a unique mix of experience and perspectives from management consulting and economics. Collectively, they draw on industry knowledge and hands-on experience interacting with healthcare institutions seeking to implement AI to improve their processes. With this background, they conduct a costing analysis, breaking down the healthcare industry into five "stakeholder groups": hospitals, physician groups, private payers, public payers, and other sites of care (e.g., dental and home health care). Within each of these stakeholder groups, they further analyze nine domains—continuity of care, network and market insights, clinical operations, clinical analytics, quality and safety, value-based care, reimbursement, corporate

David C. Chan Jr. is an associate professor of health policy at Stanford University and a research associate of the National Bureau of Economic Research.

For acknowledgments, sources of research support, and disclosure of the author's material financial relationships, if any, please see https://www.nber.org/books-and-chapters/economics-artificial-intelligence-health-care-challenges/comment-potential-impact-artificial-intelligence-healthcare-spending-jr.

functions, and consumer—for hospitals and physician groups each and six domains for private payers—healthcare management, provider relationship management, claims management, member services, corporate functions, and marketing and sales.

As the authors state, they mostly draw on insights and experience without relying on experimental or quasi-experimental evidence that most economists would be more familiar with. Based on their analysis, they conclude that AI could lead to 5 to 10 percent lower US healthcare spending, about $200 billion to $360 billion annually in 2019 dollars, within five years and without reducing quality or access. Nonetheless, they note that AI adoption has lagged in the healthcare industry relative to other industries. To explain the lack of adoption, they focus on "managerial challenges," including "legacy technology, siloed data, nascent operating models, misaligned incentives, industry fragmentation, and talent attraction." They note market trends that suggest a mitigation of these challenges and an acceleration in the pace of AI adoption.

As a physician and health economist, the first question I have is the following: What makes technology adoption different in healthcare relative to other industries? Since Arrow (1963) and continuing with Cutler (2010), health economists have produced insights into differences in healthcare relative to other industries and the implications of these differences for productivity. Given the fee-for-service payment system and the high degree of market concentration in the industry, improving efficiency by reducing costs has not typically been the way for healthcare delivery systems to increase profits. As we know from efforts at healthcare reform, change in the industry will need to be filtered through stakeholder groups with powerful informational or institutional advantages. New technologies such as AI will need to be adopted by these stakeholder groups; if adopted, they will naturally be used for the benefit of these groups. If it is not in the best interests of these groups to reduce costs, then cost reduction may not come to fruition even with highly effective technologies.

The managerial challenges that the paper casts are somewhat generic—there seems to be little insight into why "legacy technology" and "talent challenges" should be a bigger barrier in healthcare relative to other industries. Is there a reason why healthcare should have talent challenges relative to other industries? To explain why AI adoption or AI impact has lagged in healthcare relative to other industries, it seems crucial to link these phenomena to underlying economic differences between healthcare and other industries. It may be instructive to review the string of technological tools that have come before AI in the past. For example, health IT has previously been cast as a technology with the potential to reduce costs, saving patients from unnecessary utilization and adverse events. However, despite the availability of health IT systems, less than 5 percent of hospitals adopted a health IT by 2008, when they were heavily incentivized by federal legislation to adopt health IT products (Jha et al. 2008). Healthcare systems rarely integrated data with other systems, again until legislated to do so (Adler-Milstein et al. 2014). In my view, health IT provides one of many cautionary examples of economic incentives imbedded in institutions and policies shaping the use and the features of a new technology.

Sahni et al. present a useful step forward in envisioning the potential impact of AI on healthcare spending. As they note, they lack citations to existing experimental and quasi-experimental evidence to form the basis of their opinions. The lack of existing evidence is a fine justification for using expert opinion to weigh in on an important question. However, in future work, I would be eager to see the gap filled by a more data-driven approaches, even if the data are simply correlational. Heterogeneity in adoption and in effect is the rule in healthcare rather than the exception. A closer look at the characteristics of healthcare systems that have adopted AI and the effects of adoption on spending and outcomes would likely yield significant insights into the intended and unintended consequences of AI on the healthcare industry as a whole.

References

Adler-Milstein, Julia, Catherine M. DesRoches, Michael F. Furukawa, Chantal Worzala, Dustin Charles, Peter Kralovec, Samantha Stalley, and Ashish K. Jha. 2014. "More Than Half of US Hospitals Have at Least a Basic EHR, but Stage 2 Criteria Remain Challenging for Most." *Health Affairs* 33 (9): 1664–71.

Arrow, Kenneth J. 1963. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review* 53 (5): 941–73.

Cutler, David. 2010. "Where Are the Health Care Entrepreneurs?" *Issues in Science and Technology* 27 (1): 49–56.

Jha, Ashish K., Catherine M. DesRoches, Eric G. Campbell, Karen Donelan, Sowmya R. Rao, Timothy G. Ferris, Alexandra Shields, Sara Rosenbaum, and David Blumenthal. 2009. "Use of Electronic Health Records in U.S. Hospitals." *New England Journal of Medicine* 360 (16): 1628–38.

Comment Mark Sendak, Freya Gulamali, and Suresh Balu

Introduction

While enthusiasm for the role of artificial intelligence (AI) in healthcare continues to mount, economic analyses demonstrating successful return on investment are scant. In their piece titled "The Potential Impact of Artificial Intelligence on Healthcare Administrative Spending," Sahni and colleagues estimate the total potential savings from AI in healthcare to be $200 billion to $360 billion annually. These estimates will likely spur further investment in the development and adoption of healthcare AI. However, unless stakeholders rapidly align on strategies to overcome barriers and achieve the required activation energy, the potential value of healthcare AI will remain beyond reach.

We represent the Duke Institute for Health Innovation (DIHI) at Duke Health, a multihospital health system with 67,000 inpatient admissions and 4.7 million outpatient visits annually (Duke Health 2023). Similar to Sahni and colleagues, we draw upon a combination of academic and industry experience. We have nearly a decade of experience working on internal innovation projects that design, develop, and integrate novel technologies and care delivery models within Duke Health. Through our work at DIHI, we have developed and implemented over 15 AI solutions internally and have multiple initiatives validating AI solutions in external health systems. We also launched the Health AI Partnership (HAIP) in 2021 to convene stakeholders from health systems across the United States to advance the ethical adoption of AI (Duke Institute for Health Innovation 2021). Through our work at HAIP, we have conducted 85 interviews with clinical, technical, and operational leaders across nearly a dozen health systems in the US to surface and disseminate AI adoption best practices. While we work across care delivery settings and medical conditions, our perspective is primarily grounded in the experience of health systems and physician practices.

In this comment, we present several analyses that complement the work of Sahni and colleagues. First, we describe concrete use cases that reinforce the hospital AI delivery domains and the need to capture both financial and nonfinancial benefits. Second, we present on-the-ground insights that identify gaps in evidence relied upon by Sahni and colleagues. Lastly, we

Mark Sendak is the population health & data science lead at the Duke Institute for Health Innovation.

Freya Gulamali is a research analyst at the Duke Institute for Health Innovation.

Suresh Balu is associate dean for innovation and partnership at the Duke University School of Medicine, and program director of the Duke Institute for Health Innovation.

For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see https://www.nber.org/books-and-chapters/economics-artificial-intelligence-health-care-challenges/overcoming-activation-energy-required-unlock-value-ai-healthcare-sendak.

identify specific organizational (within-health-system) and seismic (policy-level) interventions that could overcome the activation energy to unlock the value of healthcare AI.

AI Delivery Domains

In this section, we present use cases from DIHI and HAIP that illustrate the AI delivery domains described by Sahni and colleagues. We focus on the six domains related to direct patient care and not related to reimbursement or corporate functions.

The first AI delivery domain is continuity of care, described as "optimizing point-of-service and referrals to improve patient care." Within this domain, our team at DIHI used AI to predict hospital readmissions to optimize postdischarge transfers to skilled nursing facilities (SNFs). Geriatric patients discharged to SNFs are at increased risk of hospital readmission, and AI can prioritize patients for telemedicine support to ensure appropriate postacute care (Krol et al. 2019; Bellantoni et al. 2022). This use of AI can create financial value in value-based care programs by preventing hospital readmissions and nonfinancial value by improving the safety and quality of care provided within SNFs.

Second, network and market insight applications are described as "tracking relationship strength among providers." Within this domain, one of our HAIP sites, Parkland Center for Clinical Innovation, used AI to segment their patient population to design tailored clinical programs for clusters of patients (Tamer et al. 2022). This use of AI creates nonfinancial value by improving patient experience and addressing barriers to access.

Third, clinical operations applications are described as "optimizing clinical workflow and capacity throughout [the] care journey." Within this domain, our team at DIHI used AI to predict admissions to the hospital requiring either intermediate or intensive care unit level care (Fenn et al. 2021). This application of AI can improve patient flow within the emergency department, prompting timely inpatient transfers for patients requiring escalation of care and discharge for patients who can safely return home. This use of AI creates financial value by increasing emergency department throughput and nonfinancial value by improving patient experience.

Fourth, clinical analytics applications are described as "improving patient care journey with data at all points of care delivery." This domain overlaps heavily with clinical operations, especially when optimizations to health system operations align with patient care goals. For example, our team at DIHI used AI to identify patients at high risk of postsurgical complications as well as patients at high risk of inpatient mortality (Corey et al. 2018; Brajer et al. 2020). In both these cases, accurate risk stratification can ensure that invasive surgical and medical interventions align with patient goals of care. These uses of AI create nonfinancial value by improving patient experience, but financial value depends on the reimbursement model. In a fee-for-service model, these uses of AI can have a negative financial impact (i.e., by reducing procedures and treatments), whereas in a value-based care model, these uses of AI can create financial value.

Fifth, quality and safety applications are described as "reducing major adverse events with special attention to patient experience and legal compliance." This domain also overlaps heavily with clinical analytics and clinical operations, and the financial impact depends on reimbursement model. For example, our team at DIHI used AI to identify patients at high risk of sepsis as well as patients at high risk of incident HIV (Bedoya et al. 2020; Burns et al. 2022). In both these cases, infections and their resultant complications can be avoided with timely prevention and treatment. These uses of AI create nonfinancial value by improving patient safety and experience, but like other domains, the financial value depends on the reimbursement model. In a fee-for-service model, these uses of AI can have a negative impact, whereas in a value-based cased model, these uses of AI can create financial value.

The final AI delivery domain is value-based care, described as "improving patient outcomes with value-based care models." This domain resolves much of the tension in the prior domains by asserting the reimbursement model. Within this domain, our team at DIHI used AI to predict progression of chronic kidney disease within an accountable care organization population (Sendak, Balu, and Schulman 2017). Patients at high risk of end stage renal disease can be proactively referred to specialty care to initiate interventions that slow disease progression. These use cases create nonfinancial value by improving patient experience and create financial value by reducing costs associated with advanced chronic disease.

The examples above reveal the complexity of capturing value from AI and the role for total mission value metrics that combine financial and nonfinancial measures. In a fee-for-service reimbursement model, the only domain that consistently generates financial value is clinical operations. In a value-based care reimbursement model, a much broader variety of domains generate financial value. However, the efficient scaling and diffusion of AI in healthcare will ultimately be determined by how much total mission value creates real financial returns. In settings that are unable to fully align incentives across payer, hospital, and physician practice, only a limited scope of AI applications will achieve broad adoption.

On-the-Ground Insights

Three on-the-ground insights derived from our work with DIHI and HAIP reveal gaps in evidence relied upon by Sahni and colleagues. First, the benefits of AI integration presented by Sahni and colleagues are highly optimistic both in terms of timing (immediacy of returns) and magnitude (size of returns). Two quantitative estimates are, first, "In our experience, organizations that deploy AI have twice the five-year revenue compound annual growth rate [CAGR] compared with others that do not"; and second, "Our estimates do not include one-time implementation costs, which in our experience are 1.0 to 1.5 times the annual savings." As described above, most health system and provider practice AI use cases do not generate financial value and would not directly increase CAGR. In a recent McKinsey report, five-year annual CAGR was estimated at 3 percent, down from the prior estimate in July 2022 of 7 percent (Patel and Singhal 2023). All health systems face significant financial pressure in the current environment, due to inflation and high labor costs, which are not entirely addressable with AI. It's unclear how health systems that deploy AI would double their CAGR compared to health systems that don't deploy AI.

Existing evidence also does not support the claim that implementation costs for health AI are 1.0 to 1.5 times annual savings. In fact, health information technology (IT) is notorious for high implementation costs that yield minimal returns. For example, while interoperable health IT was estimated to yield $77.8 billion per year in 2005, despite a $30 billion investment by the US government, the impact of electronic health records (EHRs) on health system finances was minimal (Walker et al. 2005; Beauvais et al. 2021). Many health systems saw financial losses from EHR implementations (Adler-Milstein, Green, and Bates 2013). Without well-documented case studies of AI implementations leading to immediate financial value, it's unclear if health systems and physician practices will achieve the described results.

The second problematic gap in evidence relates to the scalability of current health information technology. The authors claim that "all savings estimates are based on the use of technologies available today and assume that adoption reaches full scale." Unfortunately, the authors do not describe how existing AI solutions can be fully scaled to achieve replicable results across settings. Two factors prevent the efficient scaling of current AI solutions across settings. First, current EHR system implementations are highly customized, and significant effort is required to normalize and harmonize data to conduct analyses across sites. Our team estimated the costs of implementing a single model at a single institution to be nearly $220,000 (Sendak, Balu, and Schulman 2017). Redundant effort to scale that single algorithm across all US hospitals would cost nearly $40 million. More recently, we described the significant effort required for interdisciplinary teams to conduct data quality assurance to develop new algorithms within Duke Health as well as externally validate existing algorithms in external settings (Sendak et al. 2022). Integrating AI systems into legacy IT systems in new settings remains a high-cost endeavor. Without infrastructure that normalizes, harmonizes, and monitors data across EHR systems, there are minimal efficiencies of scale for new settings to adopt AI solutions.

Even if the IT infrastructure were in place to scale an AI solution, organizations must adapt to effectively use and benefit from the technologies.

In 2018, our collaborator Madeleine Elish described Sepsis Watch, an AI-driven sepsis detection system, as *sociotechnical* to emphasize the ways in which the technology and social environment interacted to shape use of the AI system in practice (Elish 2018). Since that time, we regularly engage social scientists in our work to help surface change management opportunities and challenges to ensure successful AI integration (Elish and Watkins 2020; Kellogg, Sendak, and Balu 2022). Unfortunately, our experience building and integrating AI solutions across settings reveals that these technologies are not "turn-key," and significant effort is required from transdisciplinary teams to enable successful organizational adoption.

The final gap relates to organizational characteristics associated with AI software adoption. Sahni and colleagues claim, "Hospitals have piloted AI and are beginning to scale adoption in some domains, with larger hospitals having done more than smaller hospitals." Our own work reveals that health system size is not a factor driving AI adoption. Use of AI is highly concentrated within academic medical centers (AMCs), which only account for 35 percent of hospital admissions in the United States (Burke et al. 2019; Sendak et al. 2020; Price, Sachs, and Eisenberg 2022). Large health systems without internal AI expertise are also more likely to rely on EHR vendors for AI solutions, many of which perform poorly when used in new contexts (Wong et al. 2021). Furthermore, our work with HAIP sites has revealed the importance of centralized AI capabilities and organizational governance structures to ensure safe and effective adoption of AI. This best practice is most mature within AMCs that have significant internal AI development and integration expertise.

Overcoming the Activation Energy

To overcome the challenges listed above, we present multiple potential organizational and policy ("seismic") interventions. First, there are high returns to increasing investment in sociotechnical research of AI integrations in healthcare. There is value at both the policy level (i.e., increases in public sector research funding) and at the organizational level (i.e., sustained investment in social science roles). For example, three systematic reviews of randomized control trials (RCTs) evaluating AI products in healthcare were published between October 2021 and September 2022 (Plana et al. 2022; Lam et al. 2022; Zhou et al. 2021). The reviews included 95 studies across 29 countries. Only 15 AI products were validated in RCTs in the US leveraging broadly available data platforms, including EHR systems and radiology imaging data. Of those AI products, sociotechnical research was conducted for two. A team at PennMedicine conducted several studies examining clinician perspectives of an AI system used to prompt serious illness care conversations for patients with cancer, and multiple sites examined organizational factors related to adoption of an AI system to help triage patients with chest pain in the emergency department (Parikh et al. 2022a, 2022b; Gesell et al. 2018; Bean et al. 2021). Without including sociotechnical research as a standard component of AI development and validation, positive results are unlikely to be replicable in new organizational contexts.

Second, technical and regulatory structures could ensure quality control of AI used by health systems and physician practices. As described above, current EHR systems do not facilitate the efficient diffusion of AI across sites. A market failure currently incentivizes health systems to rely on AI solutions provided by EHR vendors, which often perform poorly (Sendak, Price, and Balu 2022; Wong et al. 2021). Even if a best-in-class solution emerges, integration costs prevent efficient scaling. National infrastructure investment could upgrade the current health IT ecosystem to enable rapid scaling across sites. Similarly, standards and regulation could ensure that AI solutions are validated within health systems and physician practices prior to use. Regulators such as the Office of the National Coordinator could require adoption of this best practice for health IT certification, and third-party accreditation organizations, such as the Joint Commission, can ensure that health systems adopt this best practice as part of organizational governance efforts.

Third, capacity-building programs could upskill the healthcare workforce to effectively use AI. Programs that target individual clinicians, such as our DIHI Clinical Research and Innovation Scholarship, can be scaled across clinical training sites to engage more clinicians in AI product development (Sendak et al. 2021). Similarly, programs that equip organizational leaders, such as HAIP, can equip teams of interdisciplinary professionals to rapidly enhance organizational governance of AI. Funding for this training from the public sector could ensure that the existing digital divide does not widen. Without public sector intervention, AI products will largely remain within the ivory tower of highly resourced AMCs.

Conclusion

In their analysis, Sahni and colleagues estimate the total potential savings from AI in healthcare to be $200 billion to $360 billion annually. While we agree that the opportunity to improve healthcare using AI is enormous, our experiences through DIHI and HAIP reveal a more complex picture. In this comment, we present gaps in evidence that must be addressed to ensure that AI solutions are scalable across sites. We also present policy and organizational interventions that could unlock the value of AI in healthcare. Without coordinated investments in sociotechnical research, technical and regulatory structures, and capacity-building programs, the potential benefits of AI in healthcare will remain out of reach for health systems and physician practices.

References

Adler-Milstein, J., C. E. Green, and D. W. Bates. 2013. "A Survey Analysis Suggests That Electronic Health Records Will Yield Revenue Gains for Some Practices and Losses for Many." *Health Affairs* (Project Hope) 32 (3): 562–70. http://content.healthaffairs.org/content/32/3/562.full.pdf.

Bean, G., U. Krishnan, J. R. Stone, M. Khan, and A. Silva. 2021. "Utilization of Chest Pain Decision Aids in a Community Hospital Emergency Department: A Mixed-Methods Implementation Study." *Critical Pathways in Cardiology* 20 (4): 192–207.

Beauvais, B., C. S. Kruse, L. Fulton, R. Shanmugam, Z. Ramamonjiarivelo, and M. Brooks. 2021. "Association of Electronic Health Record Vendors with Hospital Financial and Quality Performance: Retrospective Data Analysis." *Journal of Medical Internet Research* 23 (4): e23961.

Bedoya, A. D., J. Futoma, M. E. Clement, K. Corey, N. Brajer, A. Lin, M. G. Simons, M. Gao, M. Nichols, S. Balu, K. Heller, M. Sendak, and C. O'Brien. 2020. "Machine Learning for Early Detection of Sepsis: An Internal and Temporal Validation Study." *JAMIA Open* 3 (2): 252–60.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7382639/.

Bellantoni, J., E. Clark, J. Wilson, J. Pendergast, J. M. Pavon, H. K. White, D. Malone, W. Knechtle, and A. Jolly Graham. 2022. "Implementation of a Telehealth Videoconference to Improve Hospital-to-Skilled Nursing Care Transitions: Preliminary Data." *Journal of the American Geriatrics Society* 70 (6): 1828–37.

Brajer, N., B. Cozzi, M. Gao, M. Nichols, M. Revoir, S. Balu, J. Futoma, J. Bae, N. Setji, A. Hernandez, and M. Sendak. 2020. "Prospective and External Evaluation of a Machine Learning Model to Predict In-Hospital Mortality of Adults at Time of Admission." *JAMA Network Open* 3 (2): e1920733–14. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2760438.

Burke, L. G., D. Khullar, J. Zheng, A. B. Frakt, E. J. Orav, and A. K. Jha. 2019. "Comparison of Costs of Care for Medicare Patients Hospitalized in Teaching and Nonteaching Hospitals." *JAMA Network Open* 2 (6): e195229.

Burns, C. M., L. Pung, D. Witt, M. Gao, M. Sendak, S. Balu, D. Krakower, J. L. Marcus, N. Lance Okeke, and M. E. Clement. 2022. "Development of an Human Immunodeficiency Virus Risk Prediction Model Using Electronic Health Record Data from an Academic Health System in the Southern United States." *Clinical Infectious Diseases* 76 (2): 299–306. https://academic.oup.com/cid/article-abstract/76/2/299/6705475.

Corey, K. M., S. Kashyap, E. Lorenzi, S. A. Lagoo-Deenadayalan, K. Heller, K. Whalen, S. Balu, M. T. Heflin, S. R. McDonald, M. Swaminathan, and M. Sendak. 2018. "Development and Validation of Machine Learning Models to Identify High-Risk Surgical Patients Using Automatically Curated Electronic Health Record Data (Pythia): A Retrospective, Single-Site Study." *PLoS Medicine* 15 (11): e1002701. https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002701.

Duke Health. 2023. "Facts & Statistics." https://corporate.dukehealth.org/who-we-are/facts-statistics.

Duke Institute for Health Innovation. 2021. "Health AI Partnership: An Innovation and Learning Network for Health AI Software." December 23, 2021. https://dihi.org/health-ai-partnership-an-innovation-and-learning-network-to-facilitate-the-safe-effective-and-responsible-diffusion-of-health-ai-software-applied-to-health-care-delivery-settings/.

Elish, M. C. 2018. "The Stakes of Uncertainty: Developing and Integrating Machine Learning in Clinical Care." *Ethnographic Praxis in Industry Conference* 2018 (1):

364–80. https://anthrosource.onlinelibrary.wiley.com/doi/full/10.1111/1559-8918.2018.01213.

Elish, M. C., and E. A. Watkins. 2020. "Repairing Innovation: A Study of Integrating AI in Clinical Care." *Data & Society* Sep 25: 1–62. https://datasociety.net/wp-content/uploads/2020/09/Repairing-Innovation-DataSociety-20200930–1.pdf

Fenn, A., C. Davis, D. M. Buckland, N. Kapadia, M. Nichols, M. Gao, W. Knechtle, S. Balu, M. Sendak, and B. Jason Theiling. 2021. "Development and Validation of Machine Learning Models to Predict Admission from Emergency Department to Inpatient and Intensive Care Units." *Annals of Emergency Medicine* 78 (2): 290–302. https://www.annemergmed.com/article/S0196-0644(21)00161-X/fulltext

Gesell, S. B., S. L. Golden, A. T. Limkakeng, C. M. Carr, A. Matuskowitz, L. M. Smith, and S. A. Mahler. 2018. "Implementation of the HEART Pathway: Using the Consolidated Framework for Implementation Research." *Critical Pathways in Cardiology* 17 (4): 191–200.

Kellogg, K. C., M. Sendak, and S. Balu. 2022. "AI on the Frontlines." *MIT Sloan Management Review* 63 (4): 44–50. https://sloanreview.mit.edu/article/ai-on-the-front-lines/.

Krol, M. L., C. Allen, L. Matters, A. J. Graham, W. English, and H. K. White. 2019. "Health Optimization Program for Elders." *Journal of Nursing Care Quality* 34 (3): 217–22.

Lam, T. Y. T., M. F. K. Cheung, Y. L. Munro, K. M. Lim, D. Shung, and J. J. Y. Sung. 2022. "Randomized Controlled Trials of Artificial Intelligence in Clinical Practice: Systematic Review." *Journal of Medical Internet Research* 24 (8): e37188.

Parikh, R. B., C. R. Manz, M. N. Nelson, C. N. Evans, S. H. Regli, N. O'Connor, L. M. Schuchter, L. N. Shulman, M. S. Patel, J. Paladino, and J. A. Shea. 2022a. "Clinician Perspectives on Machine Learning Prognostic Algorithms in the Routine Care of Patients with Cancer: A Qualitative Study." *Supportive Care in Cancer* 30 (5): 4363–72.

Parikh, R. B., C. R. Manz, M. N. Nelson, W. Ferrell, Z. Belardo, J. S. Temel, M. S. Patel, and J. A. Shea. 2022b. "Oncologist Perceptions of Algorithm-Based Nudges to Prompt Early Serious Illness Communication: A Qualitative Study." *Journal of Palliative Medicine* 25 (11): 1702–7.

Patel, N., and S. Singhal. 2023. "What to Expect in US Healthcare in 2023 and Beyond." McKinsey & Company, January 9, 2023. https://www.mckinsey.com/industries/healthcare/our-insights/what-to-expect-in-us-healthcare-in-2023-and-beyond.

Plana, D., D. L. Shung, A. A. Grimshaw, A. Saraf, J. J. Y. Sung, and B. H. Kann. 2022. "Randomized Clinical Trials of Machine Learning Interventions in Health Care." *JAMA Network Open* 5 (9): e2233946.

Price, W. N., R. Sachs, and R. S. Eisenberg. 2022. "New Innovation Models in Medical AI." *99 Washington University Law Review* 1121. https://ssrn.com/abstract=3783879.

Sendak, M. P., S. Balu, and K. A. Schulman. 2017. "Barriers to Achieving Economies of Scale in Analysis of EHR Data. A Cautionary Tale." *Applied Clinical Informatics* 8 (3): 826–31. https://www.thieme-connect.com/products/ejournals/abstract/10.4338/ACI-2017-03-CR-0046.

Sendak, M. P., J. D'Arcy, S. Kashyap, M. Gao, M. Nichols, K. Corey, W. Ratliff, and S. Balu. 2020. "A Path for Translation of Machine Learning Products into Healthcare Delivery." *European Medical Journal Innovations*. January 27, 2020. https://www.emjreviews.com/innovations/article/a-path-for-translation-of-machine-learning-products-into-healthcare-delivery/.

Sendak, M. P., M. Gao, W. Ratliff, K. Whalen, M. Nichols, J. Futoma, and S. Balu.

2021. “Preliminary Results of a Clinical Research and Innovation Scholarship to Prepare Medical Students to Lead Innovations in Health Care.” *Healthcare* 9 (3):100555.

Sendak, M. P., N. Price, and S. Balu. 2022. “A Market Failure Blocks Efficient Diffusion of AI Software for Health.” *STAT News*, May 24, 2022. https://www.statnews.com/2022/05/24/market-failure-preventing-efficient-diffusion-health-care-ai-software/.

Sendak, M., G. Sirdeshmukh, T. Ochoa, H. Premo, L. Tang, K. Niederhoffer, et al. 2022. “Development and Validation of ML-DQA—A Machine Learning Data Quality Assurance Framework for Healthcare.” *Proceedings of Machine Learning Research* (182): 741–59. https://arxiv.org/abs/2208.02670.

Tamer, Y. T., A. Karam, T. Roderick, and S. Miff. 2022. “Know Thy Patient: A Novel Approach and Method for Patient Segmentation and Clustering Using Machine Learning to Develop Holistic, Patient-Centered Programs and Treatment Plans.” *NEJM Catalyst* August 23, 2022. https://catalyst.nejm.org/doi/full/10.1056/CAT.22.0084.

Walker, J., E. Pan, D. Johnston, J. Adler-Milstein, D. W. Bates, and B. Middleton. 2005. “The Value of Health Care Information Exchange and Interoperability.” *Health Affairs* 24 (Suppl1): W5-10–W5-18.

Wong, A., E. Otles, J. P. Donnelly, A. Krumm, J. McCullough, O. DeTroyer-Cooley, J. Pestrue, M. Phillips, J. Konye, C. Penoza, M. Ghous, and K. Singh. 2021. “External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients.” *JAMA Internal Medicine* 181 (8): 1065–70.

Zhou, Q., Z.-H. Chen, Y.-H. Cao, and S. Peng. 2021. “Clinical Impact and Quality of Randomized Controlled Trials Involving Interventions Evaluating Artificial Intelligence Prediction Tools: A Systematic Review.” *NPJ Digital Medicine* 4 (1): 154.