NEJM AI

DOI: 10.1056/AIpc2400083

POLICY CORNER

Scaling Adoption of Medical AI — Reimbursement from Value-Based Care and Fee-for-Service Perspectives

Michael D. Abramoff, M.D., Ph.D.,[1,2,3] Tinglong Dai, Ph.D.,[4,5,6] and James Zou, Ph.D.[7,8,9]

Received: January 22, 2024; Revised: February 14, 2024; Accepted: February 23, 2024; Published: April 12, 2024

Abstract

Sustainable reimbursement is key for medical artificial intelligence (AI) to benefit patients and populations at scale; however, achieving reimbursement is complex and requires the support of various stakeholders. We explain the roles of the different stakeholders and the extent to which reimbursement mechanisms, including fee-for-service and value-based care, align stakeholder interests and facilitate the scaling of medical AI adoption.

A key challenge facing medical artificial intelligence (AI) is translating it into real-world adoption to improve patient outcomes. Once rigorous evidence of safety, effectiveness, equity, and interoperability has been established and regulatory requirements have been met, financial sustainability is imperative. This, in turn, requires successfully aligning the incentives of stakeholders, including patients, providers, payers, regulators, and AI creators.[1] Although this alignment is needed worldwide, the U.S. health care system illustrates this point adeptly; in addition, U.S. reimbursement decisions are monitored closely by health systems outside the United States. The current article examines which payment or reimbursement models, including fee-for-service (FFS) payment models, value-based care (VBC) payment models, and hypothetical AI-specific revenue-sharing models, may best address the financial sustainability of medical AI in the U.S. health care system.

The recent *NEJM AI* article by Wu et al.[2] quantified real-world uptake of medical AI using commercial payer claims data rather than survey data and illustrates this point well; reimbursement, under appropriate guardrails,[3] is key to real-world adoption at scale. Their findings confirm anecdotal evidence that some form of financial return, typically under the FFS mechanism, is a key factor in the successful scaling of medical AI.

This perspective examines Food and Drug Administration (FDA)–regulated medical AI systems intended for use by clinicians; consumer-facing AI and administrative AI are beyond the current scope. Clinicians use AI when for a specific patient, it may provide actionable

The author affiliations are listed at the end of the article.

Dr. Abramoff can be contacted at michael-abramoff@uiowa.edu or at Retina Service, Department of Ophthalmology and Visual Sciences, University of Iowa, 200 Hawkins Drive, IA 52242.

clinical information; has the potential to improve clinical outcomes, patient satisfaction, or efficiency; and in many cases, meets the constraints set by payers and other stakeholders. Clinicians, as intermediaries, also receive reimbursement for the costs and resource usage they incur for the AI service.

For AI systems to be reimbursable, they must show real-world effectiveness in improving clinical outcomes, health equity, clinician productivity, and cost-effectiveness while adhering to rigorous ethical standards.[4-6] Collectively, such evidence provides the scientific basis for scaling safe, effective, and interoperable medical AI. Encouragingly, emerging evidence from randomized, controlled trials highlights the impact of autonomous AI for diabetic eye examinations of primary care patients to improve health equity,[7] physician productivity,[8] and adherence to appropriate care[7,9] at scale. Meeting these foundational criteria is a prerequisite for adopting AI in health care, but even among the FDA-authorized AI systems — 692 to date[10] — many do not meet all these criteria. As a result, clearing these hurdles does not guarantee the adoption of such AI at scale. Although attempts have been made to create self-regulation around these foundational criteria, such as FDA's precertification program,[11] the impact has yet to be shown, possibly because of the almost insurmountable hurdles they would create for smaller, often physician-led AI start-ups that have been leading medical AI innovation to date. As patient characteristics and standards of care shift, AI systems must undergo continuous updates and regulatory reassessment, which in turn, require funding, and that must come from somewhere.

Without financial sustainability, adoption at scale is not achievable, even for an AI system that meets the aforementioned foundational criteria. Consider the recent case of Pear Therapeutics; this well-funded developer of AI therapeutic software received numerous regulatory approvals, established distribution partnerships, and generated ample evidence of improved patient outcomes. However, all of this could not compensate for the lack of reimbursement, leading to the company's demise last year.[12] Patients will never be able to access the proven benefits of this AI technology.

Given the pivotal role of reimbursement in the widespread adoption of medical AI, it is critical to understand who determines reimbursement policies and the framework that guides those decisions. This perspective begins with an examination of potential reimbursement mechanisms. We primarily examine public reimbursement decisions and processes, including those of the U.S. Centers for Medicare and Medicaid Services (CMS), as commercial payers in the medical AI space are often quick to follow the CMS lead (in fact, the inflection point in figure 2 in the article by Wu et al.[2] for "diabetic retinopathy" followed immediately from nationwide CMS Medicare reimbursement). We consider beyond the scope of this article, legislative pathways, such as amending the Social Security Act,[13] to create AI reimbursement, as well as patient out-of-pocket payments, as self-evident or raising health equity concerns.

A sustainable reimbursement model must meet the following criteria as also documented by the CMS[3]: ensure trust and patient benefit, by meeting foundational criteria and an ethical framework, as outlined, and thus able to address stakeholder concerns proactively, including the mitigation of racial, ethnic, and other biases[14]; ensure appropriate utilization and avoidance of overutilization; ensure financial sustainability; ensure a relatively short path length; and allow informed deployment decisions by medical professionals, health systems, and payers through transparency in estimating financial impact.

We first consider the traditional FFS approach, which treats services provided by medical AI similarly to how new drugs or medical devices are reimbursed. Multiple AI systems are already reimbursable under the FFS model, as shown by Wu et al.[2] For the AI developer, this approach carries substantial risk in terms of time and resources; it may not be the optimal way to bring the promise of medical AI to the greatest number of patients, and some suggest that it may exacerbate health disparities.[15] Despite such limitations, FFS meets many of the aforementioned criteria and may be appropriate for AI because health care systems are adept at assessing the financial impact of new technologies under it. Such transparency makes it easier to understand the potential financial benefits of specific AI systems. In particular, FFS reimbursement through a Category I Current Procedural Terminology (CPT) code under the Physician Fee Schedule (PFS), the Outpatient Prospective Payment System (OPPS), and the Inpatient Prospective Payment System (IPPS) provides financial sustainability unlike temporary mechanisms, such as the New Technology Add-on Payments or Medicare Coverage of Innovative Technologies, which are time limited.[1]

Next, we consider VBC, in which providers are reimbursed on the basis of patient- or population-related

metrics.[15] Recently, there has been a significant shift from FFS to VBC,[16] and various forms of VBC accounted for approximately 60% of the total $4.3 trillion in U.S. health care spending in 2022.[17] Merit-Based Incentive Payment Systems (MIPS),[18] the Healthcare Effectiveness Data and Information Set (HEDIS),[19] Hierarchical Condition Category (HCC)/Risk Adjustment Factors (RAF),[20-23] and full risk capitation[24] are relevant prominent examples of VBC. These VBC mechanisms also meet many of the aforementioned criteria and may have significantly fewer statutory and regulatory constraints. As Abramoff et al.[1] documented, obtaining authorization for autonomous AI to "count" toward closing care gaps for MIPS[25] and HEDIS[26] was considerably more straightforward in a shorter time frame than achieving FFS. Under fully capitated care, the providers receive a fixed payment to cover all the health care services needed by their patients. The financial impact of adopting AI technology is relatively straightforward to model here as one can quantify the financial benefit of, for example, increasing physician productivity through the adoption of autonomous AI.[8] However, only a small number of providers and health systems currently operate under such a fully capitated model.

MIPS and HEDIS are process-based VBC models that reward the performance of specific care processes for a certain fraction of a patient population. Their financial impact, in the absence of a fully capitated model, is highly discontinuous and therefore, more difficult to ascertain. For example, for a health care system to meet MIPS measure 117, 80% or more of its population must receive an annual diabetic eye examination[25] — "closing the care gap." Below that 80%, the financial benefit of closing, for example, even three quarters of that care gap is typically zero. When the measure is met, its financial benefit depends on multiple system-wide factors, such as meeting care gap closures in entirely different specialties, although it can be substantially higher per patient than FFS. This discontinuity makes it more challenging to reliably forecast the benefit of introducing new technologies, potentially disincentivizing AI adoption.

The HCC/RAF financial impact is population risk based, where reimbursement depends on the expected care expenses in a population for specific diagnoses, and there is a similar difficulty in assessing long-term financial impact as with quality measures. To illustrate, incorporating diagnostic AI under HCC/RAF may initially increase the number of patients with diagnosed cases of diabetic retinopathy (HCC 18) and thereby, the expected expenses, leading to higher reimbursement. However, if AI improves clinical outcomes and reduces disease burden, this could lead to a lower average risk in that same population and a lower financial benefit over an uncertain time frame. Finally, it is important to keep in mind that thus far, despite its promise, the real-world impact of value-based reimbursement in general on both cost and quality of care has been mixed.[27]

A potential new financial approach could be derived from the Medicare Part B model, which pays for drugs administered in an outpatient setting on the basis of the average sales price plus a markup. Here, health systems or professionals could purchase the rights to use AI, either up front or on a subscription basis, and receive payment on the basis of the average market price of the service plus a specified add-on, contingent upon CMS coverage of a particular CPT code. This model, which essentially splits revenue between AI creators and users, could alleviate some of the tensions of the FFS model, but it would not meet all of the aforementioned criteria as it still carries the risk of overutilization and incurs the time needed to create a new category I CPT code for the AI service.

As a sample real-world case study of obtaining FFS and VBC reimbursement for an AI service, let us consider autonomous AI for primary care diabetic eye examinations. Given that this AI was FDA de novo authorized to make a diagnosis without physician supervision of that diagnosis,[28] it is particularly well suited for implementation in underserved regions, typically characterized by pronounced racial, ethnic, socioeconomic, and rural health disparities. In such under-resourced areas, a significant number of patient encounters, particularly for populations covered by state Medicaid, are billed as such because Medicaid coverage for it has become widely available,[29] under the FFS model. In addition, because this type of autonomous AI has been qualified for closing care gaps under MIPS[25] and HEDIS[26] metrics, reimbursement under a VBC framework can be achieved. In fact, many providers are currently using this type of autonomous AI primarily to improve their MIPS/HEDIS scores, foregoing FFS billing. Some report that per patient, the VBC reimbursement can be up to 10 times higher than the FFS reimbursement amount (personal communication to M.D.A., January 15, 2024, by FQHC in California). Clearly, both FFS and VBC billing are effective reimbursement mechanisms for this type of AI.

Thus, several strategies can facilitate adoption at scale and maximize patient and health equity benefits. As Wu et al.[2] reported, FFS utilization under Medicaid, and VBC utilization under MIPS/HEDIS (without FFS billing) remain unrecorded in commercial payer claims data.

Exploring alternative pathways is not only valuable but essential given the challenges of achieving reimbursement for either VBC or FFS — a process that can be lengthy, resource intensive, and subject to strict regulatory and legal constraints. In the diabetic eye examination example, as Abramoff et al.[1] have shown, FFS may require broad stakeholder support under an AI ethical framework,[14] and the creation of sustainable national reimbursement for PFS and OPPS by the CMS[3,30] as well as state-based Medicaid.[29] For VBC, it may require generation of evidence of outcome improvement[5] and updating HEDIS and MIPS quality measure language to support the use of the AI under consideration.[31] A lengthy path delays large-scale adoption and more importantly, delays the realization of improved patient outcomes and health equity. In the worst case, it can completely thwart these advances as AI creators exhaust their resources.

Patients in desperate need of medical AI services rarely have the luxury of time. So, how can adoption at scale be accelerated? For example, the creator of an autonomous AI to increase access to breast cancer screening (assuming it has met the above guardrails, received regulatory clearance and exhibits extensive scientific evidence of patient benefit) — where no form of reimbursement is currently available — may pursue both pathways: for FFS, devoting resources to establishing a CPT code[30] and then pursuing CMS reimbursement and for VBC, creating evidence to fit into the existing MIPS/HEDIS quality measure and HCC category, even though pursuing all pathways in parallel may be resource intensive for smaller start-ups. To our knowledge, no other country's health care system, including single-payer systems, has worked out transparent, sustainable reimbursement for medical AI, although some Southeast Asian countries have used a procurement style approach at the national level.[32]

In conclusion, as Wu et al.[2] highlight, we are starting to see widespread adoption of medical AI in health care. We illustrate how efficient reimbursement is critical to the scaled adoption of AI. As novel reimbursement frameworks are being discussed, we explain how, working with all health care stakeholders, the existing pathways in FFS and VBC can be leveraged by specific medical AI to achieve sustainable reimbursement. We also discuss the requirements for improving existing and new reimbursement frameworks. Ultimately, sustainable reimbursement will be key to scaling adoption — and thereby, realizing the potential benefits of medical AI to better health outcomes for all.

Disclosures

Author disclosures are available at ai.nejm.org.

Author Affiliations

[1] Department of Ophthalmology and Visual Sciences, University of Iowa, Iowa City
[2] Department of Electrical and Computer Engineering, University of Iowa, Iowa City
[3] Digital Diagnostics, Coralville, IA
[4] Carey Business School, Johns Hopkins University, Baltimore
[5] Hopkins Business of Health Initiative, Johns Hopkins University, Baltimore
[6] School of Nursing, Johns Hopkins University, Baltimore
[7] Department of Biomedical Data Science, Stanford University, Stanford, CA
[8] Department of Electrical Engineering, Stanford University, Stanford, CA
[9] Department of Computer Science, Stanford University, Stanford, CA

References

1. Abramoff MD, Roehrenbeck C, Trujillo S, et al. A reimbursement framework for artificial intelligence in healthcare. NPJ Digit Med 2022;5:72. DOI: 10.1038/s41746-022-00621-w.
2. Wu K, Wu E, Theodorou B, et al. Characterizing the clinical adoption of medical AI devices through U.S. insurance claims. NEJM AI 2023;1(1). DOI: 10.1056/AIoa2300030.
3. Department of Health and Human Services, Centers for Medicare & Medicaid Services. Proposal to establish values for remote retinal imaging (CPT code 92229). September 13, 2021 (https://public-inspection.federalregister.gov/2021-14973.pdf).
4. Char DS, Abramoff MD, Feudtner C. Identifying ethical considerations for machine learning healthcare applications. Am J Bioeth 2020;20:7-17. DOI: 10.1080/15265161.2020.1819469.
5. Abramoff MD, Cunningham B, Patel B, et al. Foundational considerations for artificial intelligence using ophthalmic images. Ophthalmology 2022;129:e14-e32. DOI: 10.1016/j.ophtha.2021.08.023.
6. Saenz AD, Harned Z, Banerjee O, Abramoff MD, Rajpurkar P. Autonomous AI systems in the face of liability, regulations and costs. NPJ Digit Med 2023;6:185. DOI: 10.1038/s41746-023-00929-1.
7. Wolf RM, Channa R, Liu ATU, et al. Autonomous artificial intelligence increases screening and follow-up for diabetic retinopathy in youth: the ACCESS randomized control trial. Nat Commun 2024; 15:421. DOI: 10.1038/s41467-023-44676-z.
8. Abramoff MD, Whitestone N, Patnaik JL, et al. Autonomous artificial intelligence increases real-world specialist clinic productivity in

a cluster-randomized trial. NPJ Digit Med 2023;6:184. DOI: 10.1038/s41746-023-00931-7.

9. Huang J, Wang J, Channa R, Wolf R, Abramoff MD, Liu TYA. Autonomous artificial intelligence exams are associated with higher adherence to diabetic retinopathy testing in an integrated healthcare system. Invest Ophthalmol Vis Sci 2023;64:212.
10. U.S. Food and Drug Administration. Artificial intelligence and machine learning (AI/ML)-enabled medical devices. October 19, 2023 (https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices).
11. U.S. Food and Drug Administration CDRH. Digital health software precertification (pre-cert) pilot program. September 26, 2022 (https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-software-precertification-pre-cert-pilot-program).
12. Aguilar M. Digital health pioneer Pear Therapeutics files for bankruptcy. STAT, April 7, 2023 (https://www.statnews.com/2023/04/07/pear-therapeutics-bankruptcy-stock-sale/).
13. U.S. Congress. Social Security Act. Title XVIII — health insurance for the aged and disabled. January 19, 2024 (https://www.govinfo.gov/content/pkg/COMPS-8768/uslm/COMPS-8768.xml).
14. Abramoff MD, Tarver ME, Loyo-Berrios N, et al. Considerations for addressing bias in artificial intelligence for health equity. NPJ Digit Med 2023;6:170. DOI: 10.1038/s41746-023-00913-9.
15. Horstman C. Promoting health equity by changing how we pay for care. The Commonwealth Fund, August 15, 2023 (https://www.commonwealthfund.org/blog/2023/promoting-health-equity-changing-how-we-pay-care).
16. Porter ME, Lee TE. The strategy that will fix health care. Harvard Business Review, October 2013 (https://hbr.org/2013/10/the-strategy-that-will-fix-health-care).
17. Wilson R. Value-based care growth stagnant in 2022. Becker's payer issues. November 2, 2023 (https://www.beckerspayer.com/payer/value-based-care-growth-stagnant-in-2022.html#1479737966515-d822014b-3f98).
18. Schneider EC, Hall CJ. Improve quality, control spending, maintain access — can the merit-based incentive payment system deliver? N Engl J Med 2017;376:708-710. DOI: 10.1056/NEJMp1613876.
19. Schneider EC, Riehl V, Courte-Wienecke S, Eddy DM, Sennett C. Enhancing performance measurement: NCQA's road map for a health information framework. JAMA 1999;282:1184-1190. DOI: 10.1001/jama.282.12.1184.
20. Burwell SM. Setting value-based payment goals — HHS efforts to improve U.S. health care. N Engl J Med 2015;372:897-899. DOI: 10.1056/NEJMp1500445.
21. Centers for Medicare & Medicaid Services. Medicare Risk Adjustment Eligible CPT/HCPCS Codes. 2024 (https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors-Items/CPT-HCPCS).
22. Centers for Medicare & Medicaid Services. Note to: Medicare Advantage Organizations, Prescription Drug Plan Sponsors, and Other Interested Parties. Announcement of calendar year (CY) 2024 Medicare Advantage (MA) capitation rates and Part C and Part D payment policies. March 31, 2023 (https://www.cms.gov/files/document/2024-announcement-pdf.pdf).
23. Mandal AK, Tagomori GK, Felix RV, Howell SC. Value-based contracting innovated Medicare advantage healthcare delivery and improved survival. Am J Manag Care 2017;23:e41-e49.
24. Kumar W, Adashi EY, Kocher B. Making care primary: Medicare's latest attempt at value-based primary care. Health Affairs Scholar 2023;1:qxad072. DOI: 10.1093/haschl/qxad072.
25. Department of Health and Human Services, Centers for Medicare & Medicaid Services. Medicare and Medicaid Programs. CY 2024 payment policies under the physician fee schedule and other changes to Part B payment and coverage policies; Medicare Shared Savings Program requirements; Medicare Advantage; Medicare and Medicaid provider and supplier enrollment policies; and basic health program. Fed Regist 2023;88:78818-80047.
26. National Committee for Quality Assurance. Healthcare Effectiveness Data and Information Set (HEDIS). Eye exam for patients with diabetes. Healthcare Effectiveness Data and Information Set ("HEDIS®") Volume 2: technical specifications for health plans. Washington, DC: National Committee for Quality Assurance, 2024.
27. Burns LR, Pauly MV. Transformation of the health care industry: curb your enthusiasm? Milbank Q 2018;96:57-109. DOI: 10.1111/1468-0009.12312.
28. Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med 2018;1:39. DOI: 10.1038/s41746-018-0040-6.
29. Alabama Medicaid Agency. New coverage for procedure code 92229: imaging of retina for disease detection. October 31, 2023 (https://medicaid.alabama.gov/alert_detail.aspx?ID=16254).
30. Frank RA, Jarrin R, Pritzker J, et al. Developing current procedural terminology codes that describe the work performed by machines. NPJ Digit Med 2022;5:177. DOI: 10.1038/s41746-022-00723-5.
31. National Committee for Quality Assurance. HEDIS measurement year 2020 and measurement year 2021. Volume 2L technical specifications for health plans. Washington, DC: National Committee for Quality Assurance, 2020.
32. Ta AWA, Goh HL, Ang C, Koh LY, Poon K, Miller SM. Two Singapore public healthcare AI applications for national screening programs and other examples. Health Care Science 2022;1:41-57. DOI: 10.1002/hcs2.10.