Stanford University
Human-Centered
Artificial Intelligence

Policy Brief
HAI Policy & Society
February 2026

Toward Responsible AI in Health Insurance Decision-Making

Michelle Mello, Artem Trotsyuk, Abdoul Jalil Djiberou Mahamadou, and Danton Char

Health insurers and healthcare providers are rapidly adopting AI tools to process prior authorization requests and adjudicate claims. According to one 2024 survey, 84% of large health insurers in 16 states were using AI for some operational purposes. This rapid uptake has been spurred by the hope that AI will streamline resource-intensive tasks like utilization review, reduce errors, and allow human reviewers to focus on more complex cases.

However, the widespread adoption of AI in health insurance processes has also caused public controversy and attracted policy attention, amid numerous reports of AI tool usage leading to wrongful claim denials. Prior authorization, in particular, has long been plagued by delays and wrongful coverage denials, and one fear is that by making prior authorization reviews cheaper to conduct, AI could supercharge a flawed process. Many insurers do not document the accuracy of the models they deploy or test them for biases. And many have not instituted governance mechanisms to ensure accountability.

Key Takeaways

Health insurers have attracted public controversy and policy attention amid reports that their use of AI tools may be contributing to wrongful coverage denials.

We examine why health insurance utilization review processes, especially prior authorization, have become a focal point for AI adoption by both insurers and healthcare providers and why ensuring responsible deployment is challenging.

AI has the potential to meaningfully reduce administrative burden and care delays by automating clearly approvable requests, improving documentation quality, and supporting appeals.

However, AI tools can also exacerbate existing flaws in already fraught processes or introduce new harms — for example, by reinforcing historically unjust denial patterns.

Insurers and providers must adopt stronger institutional governance mechanisms for vetting and monitoring AI tools to ensure they improve access to care rather than entrenching incentives to deny or delay treatment.

In our article in *Health Affairs*, "The AI Arms Race in Health Insurance Utilization Review," we explore why utilization review has become such a hot spot for AI applications and why ensuring responsible deployment remains challenging. Our analysis draws on empirical research into how insurers and healthcare providers use AI, including our own ethical evaluations of provider-facing tools within Stanford Health Care, a multi-hospital health system. We offer five recommendations for policymakers and healthcare organizations to consider as they decide what role AI will play in their future operations.

AI has the potential to dramatically streamline workflows in a field burdened by high administrative costs, wrongful claim denials, and worker burnout. Yet without safeguards, AI risks reinforcing existing incentives to delay or deny care.

AI has the potential to dramatically streamline workflows in a field burdened by high administrative costs, wrongful claim denials, and worker burnout.

Introduction

Health insurers and healthcare providers have been adopting AI tools so rapidly that it has been likened to an AI arms race. A 2024 survey of large health insurers conducted by the National Association of Insurance Commissioners found that 37% reported using AI for prior authorization, 44% for claims adjudication, and 56% for utilization management activities broadly defined.

Our own review of online offerings by AI vendors reveals a robust marketplace of generative and predictive AI tools that predominantly target either insurers or providers — though some emerging vendors sell collaborative solutions to both insurers and providers. Those products aim to bridge the payer-provider divide by standardizing data exchanges and creating shared decision frameworks.

Most commonly, AI developers market their tools to insurers to help conduct utilization review — the process insurers use to decide whether to approve payment for services recommended by an enrollee's physician. In this context, AI is primarily used to support prior authorization (the pre-approval of treatments); concurrent review (assessing the ongoing need for care); and decisions about claims after services have been provided. Insurer-facing AI tools often determine whether a patient meets prior authorization requirements and generate related recommendations and correspondence.

Beyond utilization review, AI tools can also support insurers with fraud detection, disease management, pricing, marketing, and risk adjustment. AI tools

geared toward healthcare providers primarily aim to help providers secure prior authorization and payment of claims by gathering clinical documentation and filling out insurance forms.

Concrete evidence of the extent to which AI tools are currently used and how effective or ineffective they are in insurance-related settings remains scarce.

AI's Potential to Improve Insurance Processes

Concrete evidence of the extent to which AI tools are currently used and how effective or ineffective they are in insurance-related settings remains scarce. If these tools are implemented responsibly, the hope is that they could meaningfully streamline burdensome insurance processes, especially for repetitive workflows like those involved in utilization review.

Prior authorization, in particular, is an area that would benefit from improved efficiency. The increased use of prior authorization has led to provider burnout, delayed access to care, and high administrative costs for providers and insurers alike. Longitudinal studies show that prior authorization is subject to both high denial rates and high overturn rates upon appeal — with one study of Medicare Advantage plans finding an overturn rate of nearly 82%. The massive volume of submitted health insurance claims presents another opportunity for automation. Many health plans deny a substantial share of claims, with evidence suggesting that many decisions — whether made by human reviewers or non-AI algorithms — are erroneous. Yet only a small fraction of denials are appealed, in part due to patients' difficulty in understanding the "Explanation of Benefits" letters they receive following a denial.

AI can help address these problems in three key ways:

1. **Automating prior authorization and claims approvals for clearly allowable requests:** Because the majority of prior authorization requests are approved, these cases are well-suited for automation. AI can extract straightforward information from electronic health records (EHR) or compare coverage rules with requests — thereby reducing delays and freeing medical reviewers to focus on more complex cases.

2. **Helping providers submit prior authorization requests:** AI tools can assist healthcare provider staff by pulling basic information from the EHR, providing explanations of medical necessity, linking supporting documentation, and checking submissions for completeness — thereby reducing denials caused by incomplete or poorly explained information.

3. **Supporting appeals of prior authorization denials:** AI tools can help providers identify which denials are most likely to be overturned and draft appeal

letters by compiling relevant EHR information, while helping insurers produce clearer Explanation of Benefits letters and specific denial rationales — thereby helping to rectify wrongful denials.

Problems Detected in Current Uses of AI Tools

Despite their potential benefits, the use of AI tools in these contexts has also raised a variety of concerns.

Toothless human-in-the-loop: Many insurance plans are legally required to ensure that a medical professional reviews every denial, but there are mounting concerns that these reviews are insufficiently thorough and lead to wrongful denials. A particular concern is that when AI tools assemble evidence for reviewers and generate summaries, they may give medical reviewers preconceived notions of the "right" answer, which could compromise how objectively they review a proposed denial. Another worry is that insurers are pressuring reviewers to closely follow AI-generated recommendations in order to cut costs or reduce prior authorization approvals — goals that predate the introduction of AI in utilization review.

Users' low familiarity with AI: In conducting ethical assessments of AI tools used at Stanford Health Care, we found that some administrative staff who prepare insurance coverage requests have a very limited understanding of how generative AI works and what its weaknesses are. Insurance company staff who work on claims — and often lack clinical expertise and work under extreme time pressure — may similarly lack AI literacy and struggle to detect errors when reviewing

When AI tools assemble evidence for reviewers and generate summaries, they may give medical reviewers preconceived notions of the "right" answer.

AI outputs. Their inability to correct hallucinations and other incorrect outputs could over time lead to a degrading of the performance of the tools themselves, since the underlying models learn from which outputs the user accepts.

Opacity of AI predictions: Predictive AI models offer little information about the underlying decision frameworks that drive a particular approval or denial, making it difficult to challenge determinations as unreasonable. This lack of transparency is compounded by the fact that less than a quarter of insurers disclose to providers when they have used AI, and only half of them have a process to determine when to disclose AI use to patients.

Underperformance issues: Insurance-facing AI tools may produce inaccurate predictions because because the EHR often lacks information on social determinants of health, such as social support outside clinical settings. Combined with training data that may not reflect timely updates made to coverage policies, this can lead to model underperformance — particularly for historically marginalized patients.

Unintended consequences: When AI tools are trained on insurers' past decisions, this can reinforce

problematic denial patterns. For instance, tools that predict the likelihood of a successful appeal may base predictions on how the insurer has behaved in the past, creating perverse incentives to keep denying coverage requests. As a result, instead of correcting past errors, AI systems can entrench them by treating insurers' historical denials as "true" signals of appropriateness.

Uneven governance practices: Insurers and healthcare organizations often lack strong governance mechanisms for AI. Many hospitals do not conduct local evaluations, and a significant share of insurers fail to document model accuracy, test for bias, or ensure accountability for tools used in prior authorization and claims decisions. Even insurers that claim to ensure the responsible use of AI have adopted so many tools so quickly — 1,000 tools at one large insurer — that meaningful oversight is unrealistic. Although federal regulations have started setting some standards for AI use in prior authorization for Medicare Advantage plans, these do not require plans to have a process for ensuring those standards are met.

Policy Discussion

Despite the challenges specific to AI use in insurance, these tools are likely to play a growing role in the industry going forward. Insurers and policymakers have acknowledged the need to improve the prior authorization process and pledged various streamlining efforts. At the same time, federal rules now mandate faster prior authorization decision-making. To increase the likelihood that AI tools will improve rather than degrade insurers' processes, we recommend policymakers adopt several measures.

This may ultimately mean that AI is not appropriate for certain utilization review tasks until the technology improves, especially for more complex determinations.

First, all insurers and provider organizations need stronger institutional governance mechanisms to vet AI tools prior to adopting them and to monitor how they perform after deployment. State insurance regulators have already begun pushing for stronger AI governance and risk-management frameworks. As part of their governance processes, providers should require that vendors disclose tool performance data, known limitations, risks, and how they will support ongoing monitoring. Insurers should apply the highest scrutiny to tools that could falsely signal that treatment requests are not approvable. This may ultimately mean that AI is not appropriate for certain utilization review tasks until the technology improves, especially for more complex determinations.

Second, insurers and providers must look beyond basic metrics like decision speed or approval and denial rates when monitoring the outcomes of AI tools. They should monitor for model biases, assess whether model training data is representative of their actual mix of patients and health plans, and work with vendors to ensure models are updated when coverage policies

change. Monitoring must also include gathering user reports of AI hallucinations and watching for other problematic model behavior — such as learning in perverse ways.

Training frontline users of healthcare AI tools is equally crucial. Insurance company and healthcare provider staff must gain at least a basic understanding of these tools' strengths, weaknesses, common errors, and biases so they know when and where vigilance is needed.

Ensuring that human medical reviewers at insurance companies are not unduly influenced by AI outputs is critically important, albeit challenging. Insurers often assert that final denial decisions are made by human medical professionals, yet insurers may be using AI to assemble the information that these professionals rely on to form their judgments. Insurers should be required to attest that AI is only being used in straightforward request approvals and that more complex requests are assessed by a medical professional *without* an AI tool having curated a summary of the file and its recommendation of a denial. Conducting thorough human reviews before coverage denial decisions are made is necessary, even if less time is saved .

Finally, insurers must report on their AI use more transparently. Insurers' recent pledges to offer clearer explanations for prior authorization denials is a step in the right direction. Insurers should also be more transparent with the public about how they use AI, avoid relying on AI to issue denials, and clearly describe how AI benefits enrollees and how its performance is monitored. Regulators should require disclosure of which tools were used, how they were applied, and what results they produced. Existing basic reporting requirements should be expanded to mandate disclosure of additional metrics that allow regulators and researchers to assess how AI affects utilization review outcomes.

AI could make approvals faster, communication clearer, and ensure more efficient human reviews; yet it could also worsen existing problems by making it cheaper and easier to deny or delay care. Because insurers have incentives that pull in both directions, we need more robust regulation and governance to steer the technology toward beneficial outcomes and prevent misuse.

Reference: The original article is Michelle M. Mello, Artem A. Trotsyuk, Abdoul Jalil Djiberou Mahamadou, and Danton Char, "**The AI Arms Race in Health Insurance Utilization Review**," Health Affairs, January 2026, accessible at https://www.healthaffairs.org/doi/10.1377/hlthaff.2025.00897.

Stanford University's Institute for Human-Centered Artificial Intelligence (HAI) applies rigorous analysis and research to pressing policy questions on artificial intelligence. A pillar of HAI is to inform policymakers, industry leaders, and civil society by disseminating scholarship to a wide audience. HAI is a nonpartisan research institute, representing a range of voices. The views expressed in this policy brief reflect the views of the authors. For further information, please contact **HAI-Policy@stanford.edu**.

Michelle Mello is a professor of law at Stanford Law School, professor of health policy at Stanford University School of Medicine, and a faculty affiliate at the Stanford Institute for Human-Centered Artificial Intelligence (HAI).

Artem Trotsyuk is a postdoctoral fellow in the department of genetics at Stanford University.

Abdoul Jalil Djiberou Mahamadou is a postdoctoral fellow at the Center for Biomedical Ethics at Stanford University.

Danton Char is an associate professor of anesthesiology at Stanford University School of Medicine and a faculty affiliate at Stanford HAI.

Stanford University
Human-Centered
Artificial Intelligence

Stanford HAI: 353 Jane Stanford Way, Stanford CA 94305-5008

T 650.725.4537 **F** 650.123.4567 **E** HAI-Policy@stanford.edu **hai.stanford.edu**