

SEPTEMBER/OCTOBER 2026 VOL. 12, NO. 5

VALUE & OUTCOMES SPOTLIGHT

An HEOR news magazine

INNOVATION BOOST:

IMPROVING ACCESS,
OUTCOMES, AND
EFFICIENCY IN
HEALTHCARE



 ISPOR

VALUE & OUTCOMES
SPOTLIGHT

SEPTEMBER/OCTOBER 2026
VOL. 12, NO. 5

EDITOR-IN-CHIEF

Zeba M. Khan, RPh, PhD

EDITORIAL ADVISORY BOARD

Carl Asche, MBA, MSc, PhD
(United States)

Ingrid Cox, PhD
(Australia)

Koen Degeling, PhD, MSc
(The Netherlands)

Aakash Bipin Gandhi, BPharm
(United States)

Ramiro E. Gilardino, MD, MSc
(Switzerland)

Eman A. Hammad, PhD
(Jordan)

Eric Jutkowitz, PhD
(United States)

Steven E. Marx, PharmD, RPh
(United States)

Bertalan Németh, PhD
(Hungary)

Sandra Nestler-Parr, PhD, MPhil, MSc
(United Kingdom)

Raoh-Fang (Jasmine) Pwu, PhD
(Taiwan)

Diego Rosselli, MD
(Colombia)

Soham Shukla, DPharm, MSc
(United States)

D. Eldon Spackman, PhD
(Canada)

John Watkins, PharmD
(United States)

The mission of *Value & Outcomes Spotlight* is to foster dialogue within the global health economics and outcomes research (HEOR) community by reviewing the impact of HEOR methodologies on health policy and healthcare delivery to ultimately improve decision making for health globally.

VALUE & OUTCOMES SPOTLIGHT

An HEOR News Magazine

TABLE OF CONTENTS

ISPOR CENTRAL

- 3 From the Editor
Revealing the True Value of Healthcare Innovation
- 6 From the CEO
Stepping Onto the Frontier: Thoughts on the Importance of Innovation to the Future of HEOR
- 8 ISPOR Speaks
From Expected Value to Realized Value: An Agenda for Healthcare Transformation
- 10 HEOR News
- 11 ISPOR News
Innovating Evidence for Medical Device and Diagnostic Technologies: Connecting Data, Patients, and Decisions
- 14 Events & Education

COLUMNS

- 15 Policy Brief
Beyond Healthcare Spending: The Evolving Definition of Value in Health Policy
- 17 Beyond Drug Prices: What Does It Take to Foster Pharmaceutical Innovation?
- 19 From the Regions
ISPOR Chile National Ranking of Impactful HEOR Practices Expands Its Reach

FEATURE

- 21 Feature Article: Innovation Boost: Improving Access, Outcomes, and Efficiency in Healthcare
- 26 By the Numbers: Innovation in Healthcare

ARTICLES

- 27 AI-Assisted Evidence Generation and HTA: Why Transparency and Reproducibility Will Define Trust
- 30 Artificial Intelligence in HEOR: From Potential to Practice in Standards, Oversight, and Real-World Implementation
- 33 The Verification Era: Agentic AI and the Future of Evidence Generation in HEOR
- 36 Treating Multiple Sclerosis Early and Aggressively in Mexico: What the Numbers Say
- 39 Smartphone-Based Screening Could Change the Economics of Obstructive Sleep Apnea Care
- 42 AI-Estimated Depression Scores as Real-World Evidence of Treatment-Resistant Depression

INTERVIEW

- 45 Translating World-Class Research into Innovations With Real Impact in Germany

EDITORIAL STAFF

Lyn Beamesderfer

Senior Director, Publications
lbeamesderfer@ispor.org

Jordana Bieze Foster

Manager, Publications
jfoster@ispor.org

Yvonne Chan

Associate, Publications
ychan@ispor.org

Ashley Morgan

Manager, Publications
amorgan@ispor.org

ISPOR CORPORATE OFFICERS

Rob Abbott

Chief Executive Officer

Sue Capon

Chief Operations Officer

Laura T. Pizzi, PharmD, MPH

Chief Science Officer

Mitch Higashi, PhD

Chief Partnerships Officer

Betsy Lane

Director and Chief Marketing
& Communications Officer

ISPOR HEADQUARTERS

505 Lawrence Square Blvd, S
Lawrenceville, NJ 08648

Tel: 609-586-4981

info@ispor.org

www.ispor.org

VALUE & OUTCOMES SPOTLIGHT EDITORIAL OFFICE:

Value & Outcomes Spotlight

Online: ISSN 2375-8678

Published bimonthly by:

ISPOR

505 Lawrence Square Blvd, S
Lawrenceville, NJ 08648 USA

Direct photocopy permission and reprint
requests to Director, Publications.

Cover photo courtesy of
AdobeStock/Who is Danny

© 2026 International Society for
Pharmacoeconomics and Outcomes Research.

While Value & Outcomes Spotlight is designed
to provide accurate information regarding the
subject matters covered, the views, opinions, and
recommendations expressed are those of the
contributors and not of ISPOR—The Professional
Society for Health Economics and Outcomes
Research.

FROM THE EDITOR

Revealing the True Value of Healthcare Innovation

Innovation in healthcare is transforming the way we diagnose, treat, and prevent disease, driving better outcomes and improving quality of life. Advances in technology, from artificial intelligence (AI) to wearable devices, are enabling more personalized care and faster decision making. Meanwhile, new approaches to research, patient engagement, and care delivery are breaking down traditional barriers, making healthcare more efficient, accessible, and proactive than ever before.

The question for providers, payers, patients, manufacturers, and health economics and outcomes research (HEOR) professionals is no longer whether innovation is arriving, but how we recognize its full value.

Ultimately, healthcare innovation should serve a single, clear purpose. That purpose is not to deliver new gadgets or clever software, but to achieve better patient outcomes and more efficient health systems. The true measure of success in healthcare will never be the sophistication of a technology; it will be the quality of the care that technology delivers. That's why it's imperative for HEOR professionals to ensure we have the tools and the skills to document those successes.

Too often, a new drug, device, or digital tool is judged by its acquisition cost and its trial endpoints. Yet the improvements that benefit both health systems and patients—such as fewer admissions, shorter stays, and less clinician burden—rarely appear in a pivotal trial or on a single budget line. Capturing those system-level gains is where HEOR and real-world evidence (RWE) can make a difference.

It Starts With AI

The rampant proliferation of AI in healthcare exemplifies this challenge in action, because its promise is large and its evidence base is still taking shape. From imaging triage to ambient clinical documentation, AI tools are being deployed to speed diagnosis and return time to clinicians. A growing body of economic evaluations is mapping where AI creates value, and the infrastructure for judging that value is maturing as well. In the United Kingdom, [the National Institute for Health and Care Excellence \(NICE\) has set a clear expectation](#) that AI methods be used where they demonstrably add value and are transparently reported. ISPOR working groups have produced a roadmap for [generative AI in health technology assessment](#), along with the [ELEVATE-GenAI](#) reporting checklist for

large language models in HEOR. [ICER](#) and [the Peterson Health Technology Institute](#) have begun adapting value frameworks for digital health. Together, these standards and frameworks aim to provide a foundation for adopting AI with confidence.

Healthcare innovation should serve a single, clear purpose. That purpose is not to deliver new gadgets or clever software, but to achieve better patient outcomes and more efficient health systems.

Electronic health records (EHRs) make evidence gathering more achievable. When EHRs are integrated with other digital platforms, they capture utilization, outcomes, and workflow at scale. This data capture lets us evaluate innovations being used in routine practice rather than only in

the controlled conditions of a trial. To ease the burden of EHR documentation, ambient AI scribes listen to patient-clinician conversations and draft structured notes for clinicians to review and finalize in the EHR. They have quickly become the most widely adopted generative AI tool in healthcare, with studies reporting roughly [20% to 40% reductions](#) in documentation burden and burnout. Early evidence also shows [modest gains in physician productivity](#) without more claim denials, but the overall [return on investment remains uncertain](#). HEOR and RWE can help health systems determine whether faster, more complete documentation delivers real efficiency and savings.

Value-Based Care

Value-based care offers another example of how RWE can help capture an innovation's full value, as it requires health systems to pay for the outcomes that matter rather than the volume of services delivered. The Netherlands has long used [bundled payments](#) for chronic conditions such as diabetes, which have been associated with improved care coordination. [Singapore's Healthier SG initiative](#) enrolls residents with a family physician to emphasize prevention and continuity. In the United States, integrated systems like

[Kaiser Permanente](#) carry both insurance and delivery risk, giving them a strong incentive to keep people well, and Kaiser's nonprofit Risant Health is now extending its value-based care expertise to community-based systems such as Geisinger.

Hospital-at-home is an innovation that is associated with cost savings, lower risk of hospital-acquired infections, faster recovery, and better outcomes.

As highlighted in our feature article, Brazil's federally funded, municipally run Sistema Único de Saúde (SUS) delivers free preventive and primary care through local multidisciplinary teams, extending care beyond urban hospitals to underserved populations. It is credited with gains in infant mortality, maternal health, vaccine uptake, chronic-condition hospitalizations, and life expectancy, and new policy now makes primary care the hub for integrating services. In Chile, the Universidad de Valparaíso's SISTAM program uses AI and other low-cost digital tools to help about 100 public primary-care centers reconnect with patients lost to follow-up. It flags areas with frequent missed appointments, sends personalized phone reminders, and prompts teams to address barriers such as transportation and distance.

In each setting, HEOR defines what “value” means—total cost of care, quality, or patient-reported outcomes—and uses RWE to track whether health systems' performance on those measures improves in routine practice.

Strategic Shifts

Another type of innovation reshapes care delivery by changing where care is provided and who provides it. Hospital-at-home is an example that allows patients to receive acute-level care in their homes rather than in a hospital, leading to cost savings, lower risk of hospital-acquired infections, faster recovery, and better outcomes. A 2024 Cochrane Review found

[moderate-certainty evidence](#) that hospital-at-home reduces costs, even as the [Centers for Medicare and Medicaid Services report](#) to Congress found inconclusive evidence of cost savings. A May 2026 *JAMA Network Open* study of nearly 16,000 Medicare patients reported [striking mortality differences](#) favoring those receiving hospital-at-home care.

The promise of innovation is undeniable: treatment that is faster, more predictive, and more personalized than ever before. Yet promise alone heals no one.

Targeted outreach to high-need patients by community health workers has helped improve cost-effectiveness. Task shifting to nurses,

pharmacists, and other professionals has [improved blood pressure, HbA1c, and mental health outcomes](#). The next step is to use rigorous RWE and full cost accounting to test whether these early findings hold across broader populations and settings and to sharpen estimates of their true costs and benefits. Given that few economic evaluations to date have looked at task shifting, there is a clear opportunity for HEOR to fill this evidence gap.

Whether an innovation succeeds also depends on who is evaluating it, because different health systems value innovation differently. A commercial insurer such as Blue Cross may focus on budget impact and member retention. The Veterans Administration, as an integrated payer and provider, can capture downstream savings. The National Health Service, through NICE, weighs cost per quality-adjusted life year against an explicit threshold. An innovation that demonstrably improves system efficiency—for example, by freeing bed days or clinician hours while also improving patient health outcomes—speaks to all 3 approaches and strengthens the case for reimbursement and adoption.

Obstacles Ahead

Of course, realizing healthcare innovation's potential means confronting some real obstacles. These include fragmented data, uneven interoperability, short budget horizons that miss long-term returns, and evidence standards that have not kept pace with the technology. Addressing them requires investment in interoperable data infrastructure, agreement on system-level outcome measures, and evaluation plans built into implementation from day one rather than added afterward.

The question is no longer whether health systems can afford to invest in innovation, but whether they can afford not to. The promise is undeniable: treatment that is faster, more predictive, and more personalized than ever before. Yet promise alone heals no one.

By pairing forward-looking tools with deliberate process improvements and by keeping the front lines of care firmly in view, HEOR can help healthcare leaders build systems that are resilient, sustainable, and truly driven by value.

As always, I welcome input from our readers.
Please feel free to email me at zeba.m.khan@hotmail.com.



Zeba M. Khan, RPh, PhD
Editor-in-Chief,
Value & Outcomes Spotlight

FROM THE CEO

Stepping Onto the Frontier: Thoughts on the Importance of Innovation to the Future of HEOR

Rob Abbott, Chief Executive Officer, ISPOR

What is innovation and why does it matter? There is perhaps no more vexing—or important—question in building a successful business or organization. The more specific question of how innovation applies to health economics and outcomes research (HEOR) is the focus of this issue of *Value & Outcomes Spotlight*—and one that has been an important part of my own HEOR journey.

When I interviewed for the CEO position at ISPOR, I told the hiring committee that the ceiling for our field was much higher than anyone had dared to imagine, but it would take innovative thinking—and application—to realize our full potential and truly transform healthcare decision making. It was time to embrace new technologies, data sources, and analytical frameworks to measure, prove, and communicate the clinical and economic value of medical interventions. For HEOR to adequately assess therapeutic innovations, such as one-time cell and gene therapies and faster drug-development cycles, we would need to revisit our historical reliance on clinical trial data and basic cost-effectiveness models and implement innovations of our own.

The ISPOR vision of a world in which healthcare is accessible, effective, efficient, and affordable for all demands nothing less than our complete investment in a true innovation agenda.

Often confused with invention (the creation of a new concept or device), innovation centers on how that new idea can be put into action to create value for customers and other stakeholders and/or to solve a specific problem. Common types of innovation include a new product or service offering that conveys more benefit to users, or improvements to internal workflows that save time and money. What is common across these innovations is a shared focus on streamlining operations, cutting costs, adapting to market changes, or gaining a competitive advantage. At its core, innovation draws together novelty and value as well as strategic growth to achieve something new: a step change in performance.

As the articles in this issue of *Value & Outcomes Spotlight* make clear, the metaphorical revolution is well underway in HEOR. Generative AI, natural language processing, and AI agents are increasingly being used to speed up literature reviews, automate economic modeling, prepare draft reports, and translate natural language into code. At the same time, the acceleration and maturation of real-world data and real-world evidence (RWE) have enabled HEOR to move beyond sole reliance on clinical

trials to tap large, federated datasets such as [Darwin EU](#)

and [Sentinel](#) to consider how medical interventions perform in everyday settings across diverse populations. We are also seeing that HEOR professionals have been quick to adapt traditional cost metrics to account for the unique, long-term benefits of rare disease and curative treatments, including societal impacts, health equity, and patient preferences.

Still, we are closer to the beginning of this “next phase” of HEOR than the finish, and there is still much to do. Insights like those in the papers collected here can guide us as we step a little deeper into the frontier of our profession.

I feel a real sense of urgency about the need to accelerate our collective innovation capacity because traditional value assessment frameworks are breaking down under the weight of a modernized healthcare paradigm. HEOR is facing unprecedented market pressures, including the arrival of curative but high-cost gene therapies, accelerated regulatory approvals based on surrogate endpoints, and a growing demand for RWE. If we rely on static, legacy models and frameworks, we risk becoming an administrative bottleneck rather than a strategic enabler of patient access and sustainable healthcare delivery. The ISPOR vision of a world in which healthcare is accessible, effective, efficient, and affordable for all demands nothing less than our complete investment in a true innovation agenda.

With the above in mind, it seems to me that innovation in HEOR must be expressed across 4 pillars: (1) methodological and analytical innovation, (2) technological and AI innovation, (3) data ecosystem innovation, and (4) commercial and policy innovation. Let me briefly discuss each of these in turn.

1. Methodological and Analytical Innovation

Traditional metrics, like the [quality-adjusted life year](#), are increasingly viewed as too rigid or potentially discriminatory to capture the nuanced realities of modern medicine.

It therefore makes sense to me that alternative metrics—such as [Equal Value of Life Years Gained](#) or the [Generalized Risk Adjusted Cost-Effectiveness](#)—would be advanced to better quantify the value of complex interventions. And I would hope that additional new ideas will be forthcoming. Deploying advanced statistical tools such as [mixture-cure models](#) to accurately evaluate one-time, curative therapies that feature high upfront costs but delayed, lifetime patient benefits are also illustrative of the methodological and analytical innovation we need. So too with [distributional cost-effectiveness analysis](#). While



it may not be perfect, its attempt to incorporate health equity into the mathematical architecture of economic evaluations to see how resource allocation impacts underserved demographics is an important step forward.

2. Technological and AI Innovation

The sheer volume, speed, and fragmentation of modern medical data mean that human-only analysis is becoming a structural limitation. Using AI to automate the “fast science” components of HEOR—such as accelerating systematic literature reviews, structuring draft economic models, and auto-generating value narratives—is no longer a novel luxury but a day-to-day imperative. The same is true for the use of machine learning to construct high-fidelity synthetic patient cohorts. This enables researchers to simulate clinical scenarios and project outcomes without running afoul of privacy hurdles or data-scarcity issues.

3. Data Ecosystem Innovation

Evaluating long-term therapeutic value requires moving away from tightly controlled, localized data silos toward fluid, interconnected networks. This is why accelerating the adoption of decentralized data marketplaces (again, Darwin EU and Sentinel come to mind) is so important. It allows HEOR researchers to run analyses across massive, global real-world datasets without moving sensitive patient information out of its home jurisdiction. Of course, to achieve true data ecosystem innovation, we must address interoperability barriers by standardizing parameters that seamlessly link these diverse datasets.

4. Commercial and Policy Innovation

Over the past year, the most common question I've been asked as ISPOR CEO is why I'm placing such emphasis on policy innovation and impact (I can also add commercial innovation and impact here). My typical response is that it's hard to improve healthcare decision making if you're not intentional about engaging with the decision makers. The somewhat more nuanced response is this: proving value is only half the battle; HEOR must innovate to show how value translates into market dynamics and patient access.

This new approach opens up several exciting avenues for dialogue and research within our field, including:

- The design of risk-sharing reimbursement frameworks where a drug's ultimate price or payout is structurally tethered to how well patients perform on objective, real-world metrics.
- The expansion of traditional value definitions to integrate socioeconomic factors such as “value of hope,” “insurance value,” and reduction of caregiver burden.
- Synchronizing early stage clinical pipeline modeling with external investment risk-return parameters to ensure that viable, cost-effective innovations receive the funding they need to make it to market. This was a topic at ISPOR's recent [Healthcare Investment Summit](#)—a first of its kind for the Society, but by no means the last.

The great Austrian-American economist, Joseph Schumpeter, introduced the concept of “creative destruction” in his 1942 book, *Capitalism, Socialism and Democracy*. He made a convincing argument as to how new innovations, technologies, and businesses constantly replace and render obsolete older industries and economic structures. In particular, he noted that any business that wasn't engaged in a true “innovation agenda” would be swept from the market by the “winds of creative destruction.”

However grandiloquent that may sound to modern ears, I agree with Schumpeter. ISPOR members and our stakeholders can therefore count on me to champion work that helps to ensure we are continually stepping onto the HEOR frontier and, as our mission would have it, *advancing* novel innovations that have the potential to genuinely transform healthcare decision making.

ISPOR SPEAKS

From Expected Value to Realized Value: An Agenda for Healthcare Transformation

Kelly Saldaña, Executive Director, ISPOR Institute for Healthcare Transformation

Health economics and outcomes research (HEOR) has given healthcare decision makers a disciplined way to assess value. [Applying HEOR methods](#) helps estimate outcomes and costs, make trade-offs visible, characterize uncertainty, and clarify the opportunity costs of choosing one use of scarce resources over another. Healthcare transformation happens when the value identified through HEOR can be converted into sustained patient benefit without unrecognized losses elsewhere in the system. That's why the [ISPOR Institute for Healthcare Transformation](#) was created: to strengthen the path from evidence and decisions to real-world application and patient benefit.

As a new organization within ISPOR with a mandate for healthcare transformation, the ISPOR Institute can help connect HEOR's comparative rigor with other fields that focus on the conditions under which assessed value is actually achieved, translating those insights for the practitioners and decision makers responsible for implementation. The concept of "realized value" offers a way to ask, with scientific discipline, what happens after an assessed (or claimed) value proposition encounters real financing arrangements, delivery systems, implementation processes, and patient lives. It also offers a way to identify practical actions that can strengthen the pathway to impact.

Realized value offers a way to ask what happens after an assessed value proposition encounters real financing arrangements, delivery systems, implementation processes, and patient lives.

The Institute proposes using realized value as the next step in an ongoing evolution of HEOR and health technology assessment (HTA), reflected in the related concepts of life cycle HTA and living HTA. Life cycle HTA [extends assessment](#) beyond initial adoption to evidence development, monitoring, reassessment, optimization, and potential disinvestment as evidence, technologies, or context change. [Living HTA](#) emphasizes a responsive, up-to-date assessment process as new evidence emerges, drawing on methods such as continual evidence surveillance, iterative review, and potentially digital tools.

Together, these concepts represent important advances because they recognize that value propositions evolve over time. Realized value builds on that evolution by shifting the main unit of inquiry from the assessment process itself to the applied settings where decisions are financed, implemented, adapted, experienced, and sustained.

This framing emerged from what Institute leaders heard and synthesized at a one-day convening in Boston in August 2026. For the event, the Institute brought together researchers, decision makers, and implementers working across HEOR, HTA, health systems strengthening, implementation science, policy, academia, and industry. Across those perspectives, participants converged on a shared challenge: healthcare transformation requires more than better evidence or better decisions at a single point in time. It requires disciplined ways to follow value as its context progresses from evidence to policy, practice, and patient experience, and to identify where that value is enhanced, preserved, diluted, or lost along the way.

In the context of HEOR, moving from inquiry to application is central to the ambition of healthcare transformation. The realized-value frame draws on:

- [HEOR to specify expected value](#)
- [implementation and health policy research](#) to examine adoption and execution
- [health-systems analysis](#) to locate organizational and financing constraints
- [political economy](#) to examine authority, interests, incentives, and feasibility from the outset
- [equity-informed economic evaluation](#) to identify where inequities and inefficiencies intersect (particularly where access, diagnosis, or delivery barriers affect who benefits and how much value is realized)

These analytic disciplines are paired with applied disciplines that translate diagnosis into action: [HTA](#), [health systems strengthening](#), implementation practice, and thinking and working politically bring those insights into ongoing decisions and action.

Realized value gives all of these perspectives a common organizing frame: specify expected value, identify where it breaks down, understand why it breaks down, and translate that diagnosis into applied practice that supports implementation, adaptation, and learning in real systems. Gaps between expected value and realized value may reflect inaccurate assumptions, changing context, implementation constraints, or shifting objectives; distinguishing among them matters because each requires a different remedy.



Methodologically, this calls for a comparative, empirical, and [mixed-methods](#) approach that uses quantitative measures of outcomes, costs, access, equity, and implementation alongside qualitative evidence on mechanisms, context, stakeholder experience, authority, incentives, and political feasibility, understanding the limitations of making specific [causal claims](#) in complex settings.

Illustrative Example

A medicine demonstrates favorable expected value at the time of assessment, but diagnostic delays, restrictive coverage rules, high out-of-pocket costs, fragmented referral pathways, or limited provider capacity prevent the medicine's timely and sustained use by patients. A realized-value inquiry would examine where along that pathway expected benefit is preserved, diluted, or lost, and explore how policy, financing, delivery, or implementation changes could potentially close the gap.

The reason to trace these pathways is practical. Without understanding where and why value is preserved, diluted, or lost, decision makers can mistake an implementation failure for a failure of the therapy itself, overlook avoidable losses of access or benefit, or miss opportunities to improve impact through financing, delivery, accountability, or patient-facing changes.

To make realized value operational, a product, service, or intervention can be traced in 2 directions. A **patient-backward pathway** begins with a consequential moment in the care journey: delayed diagnosis, treatment abandonment, an avoidable complication, financial distress, or a transition that worked well, and examines how specific features of that journey shaped access to, use of, experience with, or benefit from the product, service, or intervention. It then traces backward to the policy, financing, organizational, implementation, and accountability conditions that helped produce that consequential experience. Existing research offers methods for doing this carefully, including [patient journey mapping](#), [patient-reported experience analysis](#), and [case-study approaches](#) that can identify barriers, facilitators, mechanisms, and competing explanations without assuming simple causality.

A **policy-forward pathway** begins with a specific policy or system decision affecting a product, service, or intervention (such as coverage, reimbursement, procurement, guideline adoption, benefit design, or service redesign) and examines how access, use, experience, or benefit changes such as financing arrangements, institutional responsibilities, delivery processes, implementation choices, and patient-facing conditions respond. [Implementation outcomes](#)—including acceptability, adoption, feasibility, fidelity, appropriateness, implementation cost, penetration, and sustainability—can be distinguished from service and patient outcomes to identify where expected value is realized, diluted, or lost.

Across both pathways, [political economy](#) provides foundational understanding and contextualization from the beginning. It

helps define what is being studied, why particular bottlenecks or successes arise, how authority, interests, incentives, and institutions shape the pathway from evidence to impact, and what kinds of action are feasible in a given setting. This understanding then informs case selection, pathway mapping, stakeholder engagement, and feasibility assessment. Tracing value backward from patient experience or forward from policy decisions can identify where value is lost, while [political-economy analysis](#) helps explain why breakdowns or successes occur and where coalitions for change might emerge.

Within this realized value frame, the Institute will work to distill the complex path from evidence to policy to practice through a clear approach for moving from inquiry to application. The Institute has a **Definition–Consensus–Implementation** approach:

- It begins by clarifying the problem, the value expected, the pathway through which that value should reach patients and systems, and the evidence, system, or political-economy questions that need to be addressed.
- It then brings affected stakeholders into [structured deliberation](#) to distinguish symptoms from root causes, weigh values and evidence, and identify feasible actions with clear owners.
- Finally, it supports bounded implementation in real settings, with measurement, adaptation, and learning including, where needed, [thinking and working politically](#) as an [applied approach](#) for navigating authority, incentives, coalitions, and changing context.

Using this approach, the Institute can build a realized value portfolio: a practical body of tracer cases, tested actions, and comparative lessons that help healthcare transformation move from evidence to impact.

ISPOR has helped healthcare identify and assess value with rigor. The Institute now has an opportunity to help the community follow that value through real systems: from evidence to decisions, from decisions to institutions and delivery, and from delivery to the outcomes and experiences that matter to patients.

To the ISPOR community, the Institute extends an invitation that is ambitious and practical: apply the profession's rigor not only to estimating value, but also to understanding whether and how that value becomes real.

AI Disclosure

Copilot (Microsoft) was used to assist with drafting selected sections of this piece, improving language and organization, and editing for clarity and conciseness. The concept, ideas, opinions expressed, and final content were developed and critically reviewed by the author. The author verified all factual statements, references, and citations and accepts full responsibility for the accuracy, originality, and integrity of the paper.

HEOR NEWS

1 Final Version of GLOBE Drug Payment Model Slashes Projected Savings (Healthcare Finance)

The Global Benchmark for Efficient Drug Pricing (GLOBE) Model, finalized by the US Centers for Medicare & Medicaid Services, will save Medicare beneficiaries an estimated \$400 million over 7 years—far less than the \$12 billion originally proposed.

[Read more.](#)

2 Comments Open on Uncertainty-Related Methods in Health Economic Evaluations (HEMA)

Comments are open through November 2 on a draft report from the Health Economics Methods Advisory, focused on approaches for health technology assessment bodies to evaluate, present, and communicate uncertainty in health economic evaluations. [Read more.](#)

3 Experts Weigh in on US Food and Drug Administration's Transparency Roadmap (Regulatory Focus)

The roadmap for FY2027 suggests the agency will initiate rules requiring public disclosure of submissions for new drug applications, biologics license applications, abbreviated new drug applications, and “related applications.” [Read more.](#)

4 German Claims Data Illustrate Consequences of Medicine Shortage Due to Recall (EJHE)

An analysis of claims data for more than 30,000 German patients found that the 2018 global recall of valsartan was associated with 5.3% higher costs for antihypertensives, along with increased general practitioner service utilization and care fragmentation. [Read more.](#)

5 US Government Launches \$6 Million Real-World Evidence Innovation Challenge (NIH)

The US National Institutes of Health is launching the Real-World Evidence Innovation Challenge, a national prize competition designed to establish validated benchmarks for analytic methods operating on real-world clinical data. Up to \$6 million in prizes will be awarded. [Read more.](#)

6 London's Clean Air Zone Initiative Associated With Better Lung Function in Children (Lancet)

Introduction of the Ultra Low Emission Zone in London was associated with improved lung function growth trajectories in children there, compared with children in a comparator site, suggesting the public health initiative helped restore previous deficits in lung development. [Read more.](#)

7 Study Finds Cancer Screening Rate Disparities Affecting Adults With Disabilities (American Cancer Society)

An American Cancer Society study found that adults with disabilities in the United States have lower rates of breast, cervical, and colorectal cancer screening than people with no disability-related difficulties. At-home stool testing rates, however, were similar. [Read more.](#)

8 Expanding Percutaneous Coronary Intervention Access Improves Outcomes (JAMA Network Open)

In a study of nearly 600,000 patients with acute myocardial infarction, expanded access to percutaneous coronary intervention was associated with improved outcomes in previously underserved communities but had limited benefit in areas already being served. [Read more.](#)

9 Mortality Data Favor Private Settings for Acute Lymphoblastic Leukemia Care (Lancet Regional Health – Americas)

In a study of more than 1100 patients treated for acute lymphoblastic leukemia in Brazil, Argentina, or Chile, mortality risk was significantly lower for people treated at private institutions than those treated at public institutions. [Read more.](#)

10 Human Papillomavirus Vaccine May Lower Risk of Adverse Pregnancy Outcomes (BMJ)

Swedish researchers found that quadrivalent human papillomavirus vaccination before pregnancy was associated with a lower risk of adverse pregnancy outcomes, extending benefits of human papillomavirus vaccination beyond cancer prevention. [Read more.](#)

11 African Leaders Commit to AI-Assisted Preparedness for Health Emergencies (WHO)

African leaders and global health representatives have adopted a call to action that calls for stronger cross-border collaboration and the responsible use of artificial intelligence to strengthen preparedness for public health emergencies caused by infectious diseases. [Read more.](#)

12 European Parliament Backs Stronger, Prevention-Focused Cardiovascular Strategy (Eureporter)

Members of the European Parliament overwhelmingly voted to support a new EU cardiovascular diseases strategy that prioritizes prevention, emphasizing risk factors involving alcohol, food, and tobacco and nicotine products. [Read more.](#)

ISPOR NEWS

Innovating Evidence for Medical Device and Diagnostic Technologies: Connecting Data, Patients, and Decisions

Arturo Cabra, MSc, GE HealthCare, Miami, FL, USA; Belinda A. Mohr, PhD, Medtronic, Phoenix, AZ, USA

Three Breaks in the Medical Technology Evidence Chain

ISPOR 2026 in Philadelphia put the medical devices and diagnostics (MD&D) community on full display: 3 scientific presentations, a dedicated [MD&D Special Interest Group \(SIG\)](#) meeting, a research podium, and a short course; one of those sessions was delivered jointly with the [Digital Health SIG](#). Together, these sessions showcased the field's growing methodological rigor and cross-stakeholder collaboration but also surfaced a shared throughline: good medical technology isn't enough without the evidence infrastructure to connect data, patients, and decisions.

This article looks at 3 such breaks:

- Making device and diagnostic data traceable across care settings,
- Building the infrastructure for patients to co-design the digital endpoints and biomarkers meant to represent them, and
- Making evidence decision-relevant for the payers, regulators, and health technology assessment (HTA) bodies who act on it.

The article—grounded in real examples pulled directly from the [ISPOR 2026](#) sessions—asks whether the digital infrastructure now emerging across MD&D can close these breaks, generating evidence that is continuous, patient-anchored, and decision-relevant in ways that weren't possible before. Viewed together, the sessions also point the field toward the concept of an integrated life-cycle evidence plan that keeps evidence traceable, meaningful, and actionable from product use through decision making.

Break 1: Traceable Evidence to Preserve Product Identity

The problem begins long before any analysis is run: although identification standards exist, device- and diagnostic-specific identifiers do not reliably survive the trip from point of care into the data sources used for real-world evidence. The evidence problem is not a lack of data, but loss of product identity. A stent may be scanned in the cath lab or a diagnostic assay may be recorded precisely in the lab, yet by the time those encounters reach electronic health record (EHR) extractions, claims files, and registries, the details that matter—brand, model, version, platform—have often vanished.

This erosion of identity creates a quiet but consequential distortion. When a device's model or a diagnostic's platform collapses

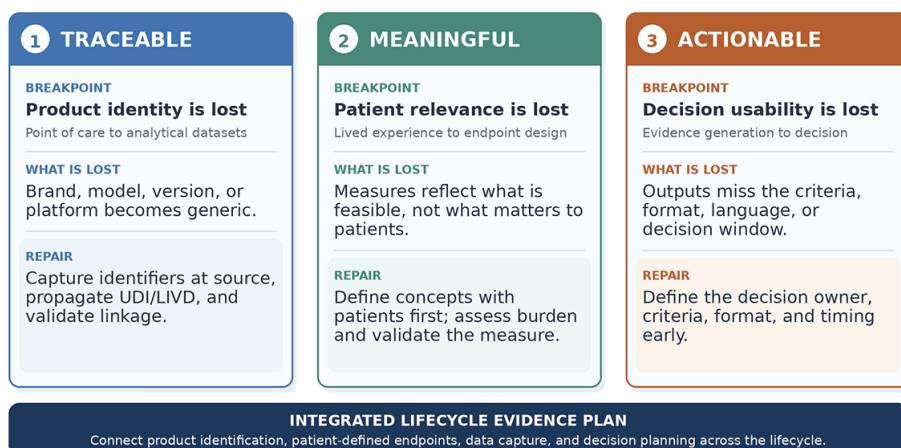
into a generic procedure or test code, the resulting evidence can obscure meaningful differences. Comparative effectiveness studies lose granularity. Postmarket surveillance becomes too coarse to be useful. Risks related to diagnostic real-world evidence (RWE) become nonactionable because the specificity needed to interpret results is buried in narrative notes or lost in translation between systems.

Identity breaks down in predictable ways across the systems that handle device and diagnostic data. Supply-chain tools may preserve unique device identifier (UDI) and lot numbers, EHR notes may preserve context, claims systems may preserve utilization, and laboratory information systems may preserve result values—but [each preserves only a fragment](#). No single system carries the full identity forward. The result is a data trail that begins with precision and ends with generalities, leaving decision makers with evidence that lacks the specificity needed for informed decisions.

Capturing identity at the source, maintaining robust master data, linking systems at the point of exposure, and abstracting residual narrative with modern language models all help rebuild what the data pipeline erodes. Several [pragmatic solutions](#) are already emerging, including linking device registration and device-generated data to downstream outcomes sources such as claims and EHRs, linking clinical trial and registry data to outcomes datasets for longitudinal follow-up, and using chargemaster data to identify specific products in comparative safety analyses. Still, even the most sophisticated linkage and

Figure 1. Three breaks in the medical technology evidence chain

An integrated lifecycle evidence plan keeps evidence traceable, meaningful, and actionable.



UDI indicates unique device identifier; LIVD, laboratory in vitro diagnostics.

extraction strategies do not eliminate [the need for validation, governance, and evidence](#) that is demonstrably fit for purpose.

In the end, the challenge is evident but hard to solve; RWE is only as trustworthy as the product identity it carries forward. If product-level identity is lost along the way, the resulting evidence may be too aggregated to distinguish meaningful differences across technologies.

Example in context: The US Food and Drug Administration's Systemic Harmonization and Interoperability Enhancement for Laboratory Data (SHIELD) pilot showed how quickly diagnostic identity drifts once tests enter real clinical systems. That pilot found that [only 59%](#) of Logical Observation Identifiers Names and Codes (LOINC codes) recommended by diagnostic test manufacturers matched the LOINC codes used in medical center laboratories for the same tests. Naming and coding practices used by local laboratory information systems led to mismatches in the remaining 41%. Without harmonized identity, diagnostic RWE cannot reliably compare tests or interpret results across sites.

Break 2: Meaningful Evidence – Start With What Matters to Patients

Patients live their lives between clinic visits—not inside them—and the signals that matter most to them often happen in those in-between spaces. As digital tools move into everyday use, from wearables to home sensors to passive monitoring, they promise to capture what patients experience in real time rather than at the next appointment. That promise is powerful, but it raises questions that most programs skip: who decides what is worth measuring, and how will those data connect to the broader clinical record?

The problem is not a lack of digital capability. It is the absence of patient input early enough to shape the endpoint itself, combined with infrastructure's limited ability to link patient-generated data to EHRs, claims, and other systems across care settings.

A stent may be scanned in the cath lab or a diagnostic assay may be recorded precisely in the lab, yet by the time those encounters reach electronic health record extractions, claims files, and registries, the details that matter—brand, model, version, platform—have often vanished.

Without linkage infrastructure, even well-designed measures may remain isolated signals rather than evidence that can inform decision makers. Instruments migrated from paper to a digital format can lose the meaning they were designed to capture, even when the conversion is technically sound.

Without patient input, digital endpoints risk measuring what is technically feasible rather than what is meaningful to the people living with the condition. Skipping patient input can add burden, reduce adherence, and leave technically sound measures disconnected from clinical workflows and reimbursement

decisions. Digital endpoints and biomarkers therefore need to be meaningful to patients and credible to the clinicians, regulators, and payers who will act on them.

Conceptual validity—defining what matters to patients before selecting a sensor or algorithm—[must be the first step](#). Without it, the technology may advance while the endpoint drifts from lived experience.

When identity, relevance, and usability stay connected from the moment a product is used to the moment a decision is made, real-world evidence becomes what it was always meant to be: a tool that improves care, rather than just describing it.

A digital endpoint built with patients from the start is also the one built to answer the questions patients, clinicians, payers, and regulators are already asking—the same audience Break 3 turns to next.

Break 3: Actionable Evidence, Designed for the Decision

Good evidence and good decisions do not always go hand-in-hand. Even when product identity is preserved upstream (Break 1) and patient-centered endpoints are built in collaboration with patients and backed by infrastructure to connect digital data to the broader care journey (Break 2), the evidence can still fail to influence the people who ultimately decide. Across medical devices and diagnostics, decision-readiness depends on 3 distinct elements: relevance and format, timing, and methodological credibility. Technologies often fall short on more than one of these elements.

[Procurement makes this tension easiest to see](#). Interviews across 6 European countries—part of an ongoing [MD&D SIG project](#) that includes a manuscript currently in preparation—showed that HTA and procurement are not integrated. Where they do connect, it is usually informal—for example, a procurement team reading an HTA report after the fact rather than shaping the decision together.

Two problems sit inside that disconnect. One is format: in interviews, procurement teams asked for shorter, more actionable outputs built around what a purchasing decision actually needs: budget impact, eligible population, implementation requirements, organizational effects, and timing aligned to the tender calendar. The other is criteria: Procurement teams weigh supply resilience, clinician preference, sustainability, and contract terms alongside affordability. Evidence built for HTA can be perfectly adequate for HTA's purpose and still arrive at procurement unusable, simply because it answers different questions.

Timing creates its own kind of failure. Evidence can be perfectly accurate and still arrive too late to matter, especially when economic questions are addressed only after pivotal data are locked in. [Early economic modeling](#) is more useful as an iterative part of development, helping teams clarify value drivers,

test scenarios, and reduce commercial risk long before major decisions become difficult to reverse. This is the case not just for earlier modeling, but also for modeling that evolves as the technology evolves. When that process is delayed, the evidence eventually generated may be exactly right, but its usefulness will be limited if it lands after the investment, design, and strategic decisions it could have shaped are already fixed in place.

Example in context: [The Device Economic Evaluation Methodological Quality \(DEEM-Q\) checklist](#), developed through a Delphi process with 40 health economists, illustrates what decision-ready evidence can look like. It provides a consensus framework for assessing the methodological quality and reporting of device economic studies—addressing the credibility component of decision readiness.

Reconnecting the Evidence Chain

Across identity, patients, and decisions, the pattern is the same: evidence breaks when continuity breaks. Product identity erodes as data move through systems. Patient relevance fades when endpoints are built around what can be measured rather than

what should be understood. Decision maker usability collapses when evidence arrives in formats, timelines, or languages that do not match the decision that must be made. These breaks are the result of fragmented processes.

One potential path forward is an [integrated evidence-generation plan](#), an approach used more often in pharma that could be adapted to medical devices and diagnostics. Rather than treating evidence development as a series of disconnected activities, it aligns product identification, data capture, endpoint development, and evidence planning across the technology life cycle. For medical devices and diagnostics, that means preserving product identity at the point of care, linking digital and clinical data across settings, engaging patients early to define meaningful endpoints, and generating evidence iteratively in formats decision makers can use.

When identity, relevance, and usability stay connected from the moment a product is used to the moment a decision is made, RWE becomes what it was always meant to be: a tool that improves care, rather than just describing it.

ISPOR Events & Education

Use these quick links to learn more about ISPOR and the HEOR issues that matter most to you.

Resources for This Issue's Primary Theme

[2026-2027 ISPOR Top 10 Trends: Artificial Intelligence, Innovative Therapies, Digital Health](#)
[Health Service Delivery and Process of Care: Topic at ISPOR Europe 2026](#)
[Artificial Intelligence: Topic Track at ISPOR Europe 2026](#)
[Emerging and Orphan Therapies Special Interest Group](#)
[Medical Devices and Diagnostics Special Interest Group](#)
[Digital Health Special Interest Group](#)
[Innovation: ISPOR Presentations Database](#)
[Digital Health: ISPOR Presentations Database](#)

HEOR Resources

Upcoming Conferences & Events

[ISPOR Europe 2026](#)
[ISPOR 2027](#)

Education

[Webinars](#)
[Short Courses](#)
[ISPOR Education Center](#)

Information

[HEOR News Desk](#)
[HEOR by Topic](#)
[ISPOR Institute](#)

More Topics In This Issue

[Realized Value](#)
[Evidence Generation](#)
[Value-Based Healthcare](#)
[Real-World Evidence](#)
[Health Technology Assessment](#)
[Transparency](#)
[Depression](#)
[Multiple Sclerosis](#)
[Disease-Modifying Therapies](#)
[Obstructive Sleep Apnea](#)
[Smartphones](#)

Thank you to our 2026 ISPOR Summit sponsors!



Analysis Group
 Charles River Associates



Premier Applied Sciences
 Truveta
 Vizient

POLICY BRIEF



The ISPOR Monthly Policy Brief provides concise insights into emerging developments shaping the global health policy landscape. Each edition highlights timely issues relevant to the health economics and outcomes research (HEOR) community and beyond, examining how evolving policy decisions influence healthcare systems, evidence generation, innovation, and patient access. View all Policy Briefs [here](#).

Beyond Healthcare Spending: The Evolving Definition of Value in Health Policy

Ana Amaris, MD, MPH, Director, Health Policy Initiatives, ISPOR, Lawrenceville, NJ, USA

Overview: This September Policy Brief picks up where the August installment left off: with the question of how to evaluate the success of a pharmaceutical policy that is expected to achieve multiple objectives. This Policy Brief looks at how that question is emerging across health systems in different geographical regions, and poses a new question: What evidence is needed to support increasingly complex policy decisions?

Quantifying Complex Policies

The August edition of the ISPOR Policy Brief explored how the objectives of pharmaceutical policy are becoming increasingly interconnected, with policy makers seeking to balance affordability and access alongside innovation, health system sustainability, supply chain resilience, and broader economic priorities. This evolution raises a related question: As the objectives of health and pharmaceutical policy expand, how should value—and, ultimately, policy success—be defined and measured?

This question is emerging across health systems. In Canada, policy makers are considering broader approaches to assessing pharmaceutical innovation. In the United States, new drug pricing agreements combine affordability objectives with commitments related to domestic manufacturing and supply chain resilience. Discussions in South Korea illustrate potential trade-offs involving price, access, and the timing of new medicine launches, while analysis from Latin America highlights the economic and societal consequences of health and disease.

These examples illustrate a concept of value that extends beyond immediate healthcare expenditure. For HEOR, they also raise questions about evidence.

Expanding What Counts as Value

In Canada, the [Pharmaceutical and Life Sciences Sector Task Force](#) recommended a multidimensional framework for assessing pharmaceutical innovation that considers clinical and patient value, health system value, and societal value, including productivity. The Task Force also recommended that Canada's Drug Agency continue evolving its health technology assessment framework to incorporate these dimensions.

A similar conversation is taking place in Latin America from a health system perspective. During the forum "Health as the Best Investment a Society Can Make" at the [IX ISPOR Colombia Congress](#), an [analysis of the socioeconomic burden of 7 diseases](#)

[across 8 Latin American countries](#) highlighted productivity losses associated with morbidity and premature mortality. Under one of the approaches used, these losses represented an estimated 3.9% of GDP on average across the countries studied in 2022.

While the contexts differ, both examples illustrate efforts to understand health and healthcare beyond immediate expenditure. Health outcomes can affect workforce participation, productivity, families, and economic and social development, while pharmaceutical innovation may generate effects across patients, health systems, and society that are not captured through a single measure of clinical benefit or healthcare spending.

When Policy Success Depends on What We Measure

This changing concept of value is also reflected in pharmaceutical policy design. On August 31, 2026, the US Administration announced [Most-Favored-Nation \(MFN\) pricing agreements with 9 additional pharmaceutical manufacturers](#), bringing the total number of participating companies to 26. Beyond commitments related to drug prices and Medicaid access, the agreements also commit to at least \$19.6 billion in US manufacturing investment and contributions of active pharmaceutical ingredients (APIs) to the [Strategic Active Pharmaceutical Ingredients Reserve](#). The agreements combine affordability objectives with domestic manufacturing capacity and pharmaceutical supply resilience.

South Korea illustrates how these objectives can generate complex trade-offs. Concerns have emerged that the country's relatively low medicine prices, combined with growing international price referencing, could influence the sequencing of new medicine launches. An analysis presented at the [2026 ISPOR Asia Pacific Summit](#) found that South Korea ranked among the 3 lowest priced countries for 12 of 15 newly reimbursed medicines examined within the 19-country reference group proposed under the US MFN models. The authors concluded that South Korea's position within these

reference price distributions could delay market entry, noting that the analysis assessed potential risk rather than observed launch delays. [Korea has also introduced a flexible contract system that separates publicly visible list prices from confidential negotiated prices](#), an approach that may help manage international pricing exposure while supporting patient access.

These examples show why the way we measure policy success matters. Lower prices or payer spending may be important outcomes, but they do not necessarily capture effects on access, innovation, or supply resilience. Understanding these trade-offs requires evidence that looks beyond a single policy objective and considers its broader effects across the health system.

From Expected Value to Realized Value

Identifying the potential clinical, economic, health system, or societal value of an intervention or policy is only part of the equation. An equally important question is whether that expected value is ultimately realized in practice.

This distinction is particularly relevant when policies pursue multiple objectives. An intervention may demonstrate value but face implementation barriers that limit its impact, while a policy may achieve some intended objectives but produce different outcomes across diverse populations. Evaluating value requires understanding what happens after implementation and the context in which outcomes are achieved.

The concept of **realized value** emerged from insights gathered and synthesized by the [ISPOR Institute for Healthcare Transformation](#) during a one-day conference in Boston, where discussions highlighted the need to strengthen the connection between evidence generation and the questions facing health policy decision makers. This includes identifying evidence gaps, determining which outcomes matter in specific decision contexts, and assessing whether policies and interventions ultimately deliver their intended value, for whom, and under what conditions.

For HEOR, realized value creates an opportunity to strengthen the connection between evidence generation, implementation, and policy evaluation. The objective is to generate evidence that helps decision makers understand not only what could create value, but what does create value in practice.

Looking Ahead: Policy Implications and Emerging Signals

As health and pharmaceutical policies pursue a broader range of objectives, how value is defined and evaluated will become increasingly important for evidence-informed decision making.

- **Definitions of value are broadening.** Clinical outcomes and healthcare expenditures remain central, but policy makers are also considering implications for access, productivity, innovation, resilience, and societal and economic outcomes.
- **Multiple objectives create complex trade-offs.** Evaluating policy success against a single outcome may overlook effects elsewhere in the health system or pharmaceutical ecosystem, reinforcing the need to consider both intended and unintended consequences.
- **Realized value is becoming an important consideration.** HEOR can help bridge the gap between what policies and interventions are expected to achieve and what they ultimately deliver in practice, including for whom those outcomes are achieved, under what conditions, and across which dimensions of value.

POLICY BRIEF



Beyond Drug Prices: What Does It Take to Foster Pharmaceutical Innovation?

Ana Amaris, MD, MPH, Director, Health Policy Initiatives, ISPOR, Lawrenceville, NJ, USA

Overview: The September edition of the [ISPOR Policy Brief](#) explored how the expanding objectives of pharmaceutical policy raise important questions about how value and policy success should be defined and measured. Among these objectives, fostering pharmaceutical innovation remains a central policy priority, particularly as governments seek to balance incentives for research and development (R&D) with affordability, access, and health system sustainability. While drug pricing often dominates discussions about innovation, the conditions that influence R&D investment extend well beyond pricing alone.

Recent policy developments in the United States, Europe, and China illustrate how governments are addressing different elements of the pharmaceutical innovation ecosystem, raising a broader question: How are governments rethinking the policy conditions needed to support pharmaceutical innovation?

A Broader Innovation Agenda

Recent discussions at [ISPOR's Healthcare Investment Summit](#) reinforced the importance of looking beyond drug prices when considering the conditions that influence pharmaceutical innovation. Participants highlighted how regulatory uncertainty, evidence requirements, market uptake expectations, and investment conditions can shape decisions about which technologies move forward.

These considerations are increasingly reflected in policy initiatives across major pharmaceutical markets, with governments taking steps to strengthen clinical research, streamline regulatory processes, and incentivize pharmaceutical development and investment. The policy debate is expanding beyond how innovation is rewarded to include the conditions under which it is developed.

Different Approaches to Strengthening the Innovation Environment

Governments are pursuing different strategies to strengthen their pharmaceutical innovation ecosystems, with recent initiatives placing particular emphasis on regulatory efficiency, research capacity, investment, and competitiveness.

In the United States, the [Food and Drug Administration \(FDA\)](#) launched its [Expedited Investigational New Drug \(IND\) Pilot Program](#) in September 2026, seeking to accelerate the transition from early drug development to first-in-human clinical trials. The initiative introduces closer collaboration between sponsors and research institutions, earlier regulatory engagement, and rolling review of application components. By addressing potential delays during early clinical development, the FDA is seeking to make the development process more efficient while maintaining regulatory oversight.

In the European Union, the proposed [European Biotech Act](#) seeks to strengthen the region's biotechnology sector and its position in global pharmaceutical research and development. The 2025 proposal addresses challenges related to clinical trial

processes, access to financing, research and manufacturing infrastructure, and regulatory complexity. These priorities reflect growing concerns about Europe's ability to attract investment and clinical research activity amid increasing international competition.

In China, [authorities announced new measures in September 2026](#) aimed at accelerating pharmaceutical innovation and strengthening domestic industry. These include enhanced regulatory consultation during development, protections for clinical trial data, market exclusivity provisions for certain pediatric and rare-disease medicines, and support for advanced manufacturing and investment. China's strategy places pharmaceutical innovation within a broader industrial and economic development agenda, connecting regulatory policy with research capacity and domestic competitiveness.

Although these initiatives reflect different national priorities, they raise a common policy consideration: how governments can create conditions that attract investment and strengthen research capacity while encouraging the development of medicines that respond to health needs.

The Policy Challenge: Balancing Innovation and Other Priorities

Policies that improve regulatory efficiency, expand research capacity, or strengthen commercial incentives may encourage innovation, but their effects will depend on how they interact with other policies and broader economic conditions. A favorable investment environment does not necessarily ensure that research will address the greatest unmet medical needs. Similarly, faster development processes must preserve the scientific rigor and patient protections necessary for generating trustworthy evidence.

Supporting innovation also involves important trade-offs related to affordability and access. Intellectual property protections and market exclusivity can strengthen incentives for investment while influencing competition and medicine prices. Public

investment can expand the scientific foundation for innovation, but governments must still consider how the benefits of that investment translate into improvements in health.

As governments pursue different approaches to supporting pharmaceutical innovation, it is important to determine whether these policies achieve their intended objectives and how they affect the broader innovation environment.

Looking Ahead

The emerging policy landscape suggests that supporting pharmaceutical innovation requires more than adjusting prices or introducing individual incentives. It calls for attention to the relationships among research investment, regulatory predictability, scientific capacity, clinical development, and market conditions.

For policy makers, the challenge is to create an environment that supports pharmaceutical innovation while balancing affordability, access, and health system sustainability.

FROM THE REGIONS

ISPOR Chile National Ranking of Impactful HEOR Practices Expands Its Reach

Daniela Paredes-Fernández, MPH, ISPOR Chile Chapter President

Why a National Ranking?

Across the Chilean healthcare system, providers, patient organizations, startups, and clinical teams make decisions grounded in healthcare economics every day without exception, whether it's measuring outcomes, navigating pharmaceutical budgets, or optimizing care pathways. Yet this work regularly takes place without receiving any kind of formal recognition. Today, Chile has no central body to coordinate health technology assessment at the national level, and structured health economics and outcomes research (HEOR) capacity remains concentrated in universities, widening the gap between academic production and frontline practice.

To close that gap, the ISPOR Chile Chapter developed its National Ranking of Impactful Practices in HEOR. By documenting and rewarding impactful HEOR practices across the health system, the initiative recognizes what has long gone unacknowledged. [Launched in 2024](#), the program's second edition in 2026 continues this mission with expanded reach and 2 new categories, reinforcing a culture of evidence-based healthcare management and bringing new stakeholders closer to the values and methods that define HEOR at the international level.

Highlights from the Second Edition

The 2025-2026 Ranking engaged more than 800 participants—a significant increase from its inaugural edition—distributed across 9 categories, including 2 new additions: **Projects for Resolving Waiting Lists Driven by the Network** (supported by Medtronic) and **Vanguard Models of Care**.

Submissions were evaluated through a double-blinded process and scored by a distinguished panel of national and international judges, including academics, researchers, and leaders from

ISPOR's Latin American chapters. The **Health Innovation** category attracted the highest number of applications, underscoring a growing community interest in disruptive and process-centered approaches to healthcare. Notably, a patient-led foundation earned second place in the **General Ranking**—affirming that patient-generated evidence is an increasingly legitimate driver of health policy in Chile.

What Impactful Practices Were Honored?

Across 31 award positions in 9 categories, provider organizations (public and private) represented the largest share of recognized stakeholders (48.4%), followed by industry (19.4%), patient associations (16.1%), startups (9.7%), and academia (6.5%).

First-Place Winners by Category

★ Diagnosis-Related Groups (DRGs)

Hospital San José de Melipilla was recognized for implementing same-day total knee arthroplasty under Chile's DRG framework. By adopting a methodology from FONASA (Chile's National Health Fund), the hospital identified a coding gap that classified ambulatory orthopedic procedures as ungroupable, and they established a reliquidation mechanism with FONASA. The hospital delivered this care pathway for 61 patients—reducing average hospital stays from 4 days to hours, enabling discharge from Day 1, and lowering infection risks.

★ Health Technology Assessment

Roche Chile earned first place for a local cost-effectiveness analysis of risdiplam for spinal muscular atrophy (Types I, II, and III) in Chile's public healthcare system. With approximately 130 patients currently relying on litigation as their only access pathway—reaching up to \$48 billion in unplanned state spending in 2024—this analysis provides locally grounded evidence to structure a sustainable, equity-based coverage model.

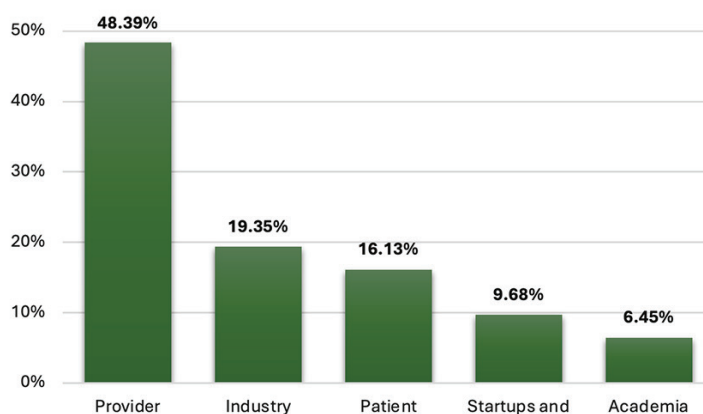
★ Health Innovation

SISTAM, developed by Gerópolis and the Universidad de Valparaíso, earned first place in this popular category as well as third place in the **General Ranking**. This co-designed primary care platform integrates electronic healthcare records, geographic information systems, and a patient reachability module across 26 family health centers in 6 municipalities, reaching approximately 600,000 people and reducing appointment no-show rates from 35% to 10%. (Learn more about SISTAM in this issue's feature story.)

★ Projects for Resolving Waiting Lists Driven by the Network

Hospital Santiago Oriente and Dr Luis Tisné Brousse were recognized for implementing a bariatric and metabolic surgery program in a public hospital specifically designed for

Figure 1. 2025-2026 Awards Sorted by Type of Stakeholder



the FONASA A Section patients, excluded from copayment mechanisms. The practice addresses an estimated 170,000 eligible patients currently without access, establishing surgical equity as a measurable health economics outcome within the public system.

★ Management of Medicines and Medical Devices

Neuro Group XR Inc, in alliance with Clínica Alemana de Santiago and Universidad del Desarrollo, earned recognition for generating real-world evidence supporting immersive virtual reality (VR) for rehabilitation. Among 21 clinicians and 44 patients, the system achieved a mean usability score of 82.5% (USEQ), with 84% of patients reporting perceived added rehabilitative value—building the evidence base for institutional adoption of this cost-effective neurotechnology.

★ Health Outcomes Evaluation

The same consortium (Neuro Group XR Inc., Clínica Alemana de Santiago, and Universidad del Desarrollo) along with Mutua de Seguridad CCHC and the University of Manchester, earned both the category prize and first place in the **General Ranking**. Their practice documented a 57% reduction in opioid expenditure in a severe burn patient and statistically significant functional improvements across all validated scales in a 12-patient acute stroke cohort treated with early immersive VR rehabilitation.

★ Patient-Led Projects

Fundación Mujeres por un Lazo organized Chile's first intersectoral forum on triple-negative breast cancer—a subtype

affecting 600 to 800 women annually, with no coverage at the time for innovative therapies in the public sector and a 5-year survival rate (approximately 7 times worse than other subtypes). The forum convened 19 stakeholders from clinical, policy, and patient communities, producing 7 concrete proposals for systemic change, and earned second place in the **General Ranking**.

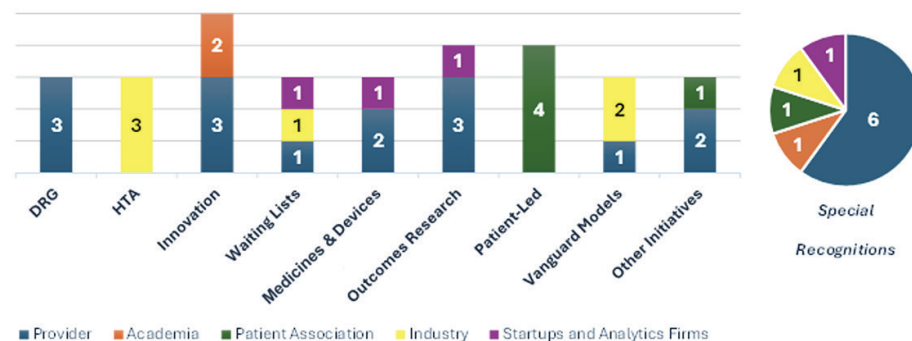
★ Vanguard Models of Care

CESFAM Dr Salvador Allende Gossens of Huechuraba Municipality was recognized for implementing a community hockey school as a salutogenic prescription for children and adolescents with attention deficit/hyperactivity disorder, anxiety, and depression. After 1 year, 85% of participants reported feeling more emotionally secure, with an average 20% reduction in anxiety scores—evidence that sport-based prescribing is an effective mental health strategy at the primary care level.

★ Other Health Impact Initiatives

The Banco Nacional de Tejidos de Chile, in alliance with FK Economics, was recognized for a quality-adjusted life years-based social and economic evaluation of Chile's national tissue bank. By correcting a market failure that had made locally produced tissue inaccessible due to high costs and limited availability, the practice established the evidence base for a policy that ensures equitable access to tissue therapies across Chile's public health system.

Figure 2. Classification of 2025-2026 Winning Stakeholders by Category



DRG indicates diagnosis-related group; HTA, health technology assessment

Advancing ISPOR's Mission

The second edition of the ISPOR Chile National Ranking has deepened the Chapter's engagement with a broadening stakeholder ecosystem: from primary care teams co-designing digital tools, to neurotechnology startups generating real-world evidence, to patient foundations driving intersectoral dialogue.

The inclusion of 2 new categories, growth in participation, and particularly, the presence of a patient association on the General Ranking podium signal that Chile's HEOR community is maturing. Evidence is being generated at every level of the health system, and the National Ranking is making that work visible.

INNOVATION BOOST

Improving **Access, Outcomes,**
and **Efficiency** in Healthcare

As health systems look for care-delivery innovations that provide value at both system and patient levels, health economics and outcomes research (HEOR) principles are informing their strategies.

By Beth Fand Incollingo

With so many advances emerging across the pharmaceutical and medical device arenas, stakeholders might expect corresponding improvements in health outcomes, along with some financial efficiencies. But in many cases, those results never materialize because their benefits are offset by [complex and cumbersome models](#) for healthcare delivery.

That's why experts in health economics and outcomes research (HEOR) are working to support innovations that can reshape the way healthcare is provided. In small private health systems and larger national ones, these approaches are being leveraged to improve patient and physician experiences while getting the most out of every payer dollar.

"I believe that the interplay between HEOR researchers and health systems is an important one to continue to nurture," said Michael Simonov, MD, Chief Product Officer at [Truveta](#), a healthcare data and artificial intelligence (AI) company. "It's important for health systems to understand whether they're choosing therapies that are beneficial both clinically and from an economic standpoint."

Health systems might innovate by heightening their emphasis on prevention, introducing technology to manage healthcare populations, or changing the setting of care. Many such approaches align not only with ISPOR's [Science Strategy and translation-to-practice agenda](#) but also with the [Top 10 HEOR Trends](#) it has identified for 2026 and 2027, which include real-world evidence (#2), value-based healthcare (#3), artificial intelligence (#1), patient-centricity (#6), and digital health (#10).

"There's a profound complexity to how healthcare systems can provide the best possible care, so there's a deep need to innovate and learn together about which treatments are the most impactful."

— Michael Simonov, MD

As a nonprofit, integrated organization in the United States, [Kaiser Permanente](#) has joined countries like Singapore in demonstrating the innovations that can arise out of value-based care—a model encouraged by ISPOR in a July 2026 task force report that introduced a new [framework](#) for applying HEOR tools to value-based healthcare implementation.

In value-based systems, doctors are paid based on the populations they treat and the health outcomes they generate. Compared with fee-for-service care, this removes incentives for unnecessary treatment, encourages commitment to prevention, and promotes health system effectiveness and sustainability.

"Every dollar invested in preventing a hip fracture can save patients significant pain and disability while potentially reducing treatment costs twenty-fold or more and produces

better economic and health outcomes," said Tad Funahashi, MD, an Orthopedic Surgeon and the Chief Innovation Officer, Emeritus, for Kaiser Permanente Southern California (KPSC).

Kaiser Permanente is also innovating by collaborating system-wide to collect, analyze, and share real-world data (RWD) about outcomes, resource use, and costs—data that allow member practitioners and facilities to learn from their own performances and those of their peers. The health system also uses that pipeline to maintain a [registry of implanted medical devices](#), such as artificial knee joints, and outcomes associated with the implantations. This database provides insights that can guide practitioners in reducing revision and readmission rates, Funahashi said.

"We can even do comparative analysis at the case level, giving reporting on revision, reoperation, and infection rates," he said. "If there is an outlier, they can try to figure out why and make changes that improve their results—for instance, by choosing specific artificial joints that are associated with better clinical outcomes."

Truveta is another organization that uses continuously updated, de-identified RWD to provide health delivery insights, but the company serves a broader array of stakeholders—the 30 US healthcare system members that govern it, as well as a host of outside subscribers, from governments to drug developers. It describes AI as central to the way it collects, aggregates, synthesizes, and compares the information its members gather from electronic health records, clinical notes, medical images, and insurance claims.

Truveta supports the ability of health systems to learn from their own performances by enabling its partners to quickly transform RWD into [evidence](#) about comparative treatment effectiveness, safety, clinical research, health disparities, care quality, and population health, Simonov said.

"There's a profound complexity to how healthcare systems can provide the best possible care, so there's a deep need to innovate and learn together about which treatments are the most impactful," he said. "That kind of learning needs to be democratized beyond the 4 walls of a health system, and the emergence of AI is opening the door for how that can happen at a national and international scale."

Exploring Value-Based Care

Innovations are needed because health systems around the world are looking for ways to reduce the stress placed on their limited resources by an aging population with growing medical needs.

In America, a lack of universal access is also challenging the effectiveness of the healthcare system, Funahashi said. While healthcare comprises [18% of the US gross domestic product](#), a [far higher percentage](#) than in other high-income countries, America's system [generates poorer health outcomes than its peers](#). To make matters worse, he said, [rising utilization](#)

rates are increasing financial pressure on Medicare, the public health-coverage program for older Americans.

“For the people who have commercial insurance and all the access they need, I do believe America still offers the most state-of-the-art healthcare you can get in the world,” Funahashi said. “But there are many people who don’t have access or can only access healthcare in less-than-ideal ways, and because of that, the overarching population is not doing that well.”

One response to those issues is value-based care, which is in the early stages of adoption around the world, according to a 2025 literature review. Among the countries testing the strategy is the Netherlands, where the idea is considered a national health-policy priority and has been incorporated by numerous hospitals. Over a decade, one pioneering hospital there has incorporated value-based care by blending deep changes to specific local practices with larger-scale, hospital-wide policy updates.

“Every dollar invested in preventing a hip fracture can save patients significant pain and disability while potentially reducing treatment costs twenty-fold or more and produces better economic and health outcomes.”

—Tad Funahashi, MD

Singapore has been shifting toward value-based care because of its “super-aged” society, where 21% of people are now aged 65 years or older. Between 2017 and 2023, across 19 medical and surgical conditions, the country’s value-based innovations were associated with the saving of 320 additional lives, the avoidance of 1753 repeat surgeries, the prevention of 589 hospital readmissions, and 22,000 fewer hospital bed days.

Value-based care is utilized within numerous health systems in the United States, too, where Medicare runs a Shared Savings Program that allows providers to form accountable care organizations and, through another initiative, provides bundled payments to practitioners for treating entire medical episodes, such as all the services involved in a joint replacement.

Kaiser Permanente’s model groups a single payer and all its member facilities and doctors under 1 cooperative umbrella, with the goal of optimizing the value of every dollar spent. The model’s ability to compile and analyze large amounts of RWD has led to several innovations that boost convenience and conserve resources by reducing on-site patient visits—including a project that has streamlined the diagnostic process for certain acute disease states. One study used electronic health records data to establish processes that have abbreviated how long it takes for patients with metastatic non-small cell lung cancer to receive pathology reports describing the molecular characteristics of their disease.

“What is the improvement in cost if we can cut diagnostic time from 6 weeks to 2 weeks by reducing the number of appointments and tests a patient needs?” Funahashi asked. “It’s such a rich field of study that we need to be doing in the United States to demonstrate how to improve the economics of healthcare.”

Through Risant Health, a company it created, Kaiser is working to spread value-based care throughout the United States. Risant’s nonprofit, value-based health systems include Geisinger, headquartered in Pennsylvania, and Cone Health, based in North Carolina.

Transforming Primary Care

Health systems around the world are also finding other innovative ways to target unmet medical needs, among them a reimaging of primary care provision.

In 2025, the director of the Pan American Health Organization called for the redesigning of primary care to close gaps in the management of noncommunicable diseases and mental health issues.

According to an Organisation for Economic Co-operation and Development (OECD) report, Brazil had already been working to meet that challenge through its Sistema Único de Saúde, or “unified health system,” which introduced free, local preventive and primary care provided by multidisciplinary teams. Funded federally but organized by municipal governments, this system shifted care away from urban hospitals with the goal of supporting health equity by more broadly reaching underserved populations. Since its origins in 1990, the program’s expansion has led to significant improvements in infant mortality rates, maternal health, vaccine uptake, hospitalization incidence for chronic conditions, and life expectancy.

Most recently, Brazil adopted a policy designed to integrate its healthcare services, with primary care as its axis.

Meanwhile, healthcare leaders in Chile are using digital technology at a population level to expand the primary care net. Working with clinics, experts from the Universidad de Valparaíso are using the technology to engage patients who frequently miss appointments.

So far, the program, known as SISTAM, is helping about 100 public primary care centers use data they already have to identify and contact patients who have been lost to follow-up, said Keiko Sahori Veintimilla Carrión, a biomedical engineer and health sciences doctoral student who is part of the university team developing the program. SISTAM, which uses AI but also simpler technologies when those are more cost-effective, identifies regions where missed appointments are common, determines which patients live in those areas, and works to improve engagement by sending personalized digital reminders to their phones.

“If we see that many people from the same area are not attending their appointments, we can ask if there are

problems with transportation, distance, or other social factors,” Veintimilla Carrión said. “With SISTAM, teams can consider information like age, health conditions, and use of services, and that can help them see the situation in a broader way.”

In 2023, [SISTAM announced the 6-month results](#) of its use at the family health center of Dr Carlos Godoy in Maipú, which included a 91% increase in attendance at health outreach events and a 14% reduction in missed appointments for health issues including electrocardiograms, during a month when patients received phone and text reminders about their appointments.

And in [a pilot program](#) from March through June 2026, SISTAM helped increase vaccine uptake by sending digital messages to target populations within 3 towns. Vaccination rates for influenza, COVID-19, and respiratory syncytial virus increased by 8.1 percentage points in El Quisco and 3.42 percentage points in Padre Hurtado, although they did not improve uptake in Curacaví, where only about half of eligible patients were reachable electronically.

“The context of each community is very different,” Veintimilla Carrión said, “so the type of messages we send or the number of reminders may need to change.”

SISTAM leaders also support a more person-centered model of care.

“The idea is to stay connected with people over time, not just when they are sick or when they have an appointment,” she said. “We want them to have more information, take a more active role in their health, and feel supported by their health centers.”

Shifting the Setting of Care

Another key theme driving health-system innovation involves shifting the setting of care.

Care administered near or at patients’ homes is encouraged by the [National Health Service in England](#), which, along with [Israel](#), [has demonstrated](#) that the hospital-at-home model can be up to 50% less costly than inpatient stays. Across the members of the Organisation for Economic Co-operation and Development, [at least 22 countries](#) offer hospital-at-home programs, including Australia, Canada, Chile, France, Japan, Mexico, Switzerland, and Spain.

At KPSC, Funahashi said, a hospital-at-home program treats 100 patients a day.

“For many patients, the reason they need to be in the hospital is because multiple teams need to closely monitor their symptomology until we can make them better,” he said. “Interestingly, if we really leverage technology, we can take many of those measurements in the patient’s home, where patients are probably more comfortable.”

In [a 2026 study](#), KPSC demonstrated that patients with heart failure had comparable levels of readmissions, mortality, and

days alive and out of the hospital when treated at home versus in a Kaiser Permanente facility.

KPSC has also moved cardiac rehabilitation out of its facilities and into patients’ homes by using smart watches to connect them with nurse practitioners. By easing barriers such as a lack of transportation, childcare, or free time, [the program](#) has significantly increased the percentage of eligible patients who participate, [reduced hospital readmission rates](#), and equalized health outcomes across [ethnicities and genders](#).

Now the concept should serve as a model for healthcare systems everywhere, Funahashi said.

“If cardiac rehabilitation was a drug, it would be a blockbuster because of how dramatically it improves health outcomes. Yet, in the United States, most eligible patients never get it,” he said. “We would like to see government agencies and academic centers doing refined health economics analysis on this to demonstrate its beneficial impact.”

Incorporating Technology

AI is becoming increasingly central to innovation across every aspect of healthcare, and that includes changes to the way systems deliver services.

A [report by the World Health Organization](#) in April 2026 showed that 74% of the 27 members of the European Union use AI-assisted diagnostics and 63% employ chatbots to engage with patients. Currently, the region is adding workforce training to ensure that AI will continue to support safety, equity, and accountability in healthcare.

“Increasingly, AI is landing in virtually every workstream within a health system. It’s appearing at the bedside, where clinicians use AI for decision support, in revenue cycle management, and in quality improvement,” said Simonov, who practices medicine within the US Department of Veterans Affairs.

“With SISTAM, teams can consider information like age, health conditions, and use of services, and that can help them see the situation in a broader way.”

— Keiko Sahori Veintimilla Carrión

At Kaiser Permanente, leaders have tasked [ambient AI](#) with listening during patient visits and creating most of the necessary written notes, so that doctors don’t have to do it afterward.

“As wonderful a gift as the electronic medical record is in terms of the comprehensiveness of the notes,” Funahashi said, “it has also turned us into file clerks and stenographers, so ambient AI is a huge burden saver.”

Truveta deploys AI to look for patterns and trends in RWD that

shed light on issues such as which treatments are associated with the best outcomes, how to reduce emergency-room waiting time, and whether revenue generation is meeting expectations.

"If you want to study whether patients with access to transportation do better or worse on a therapy, it's typically hard to get that information, because it's often only documented in a clinical note or in the deep bowels of the electronic health record," Simonov said. "Getting it out, cleaning it, and making it usable for research with AI is important to understanding all the facets of patient care."

Truveta's database is among [a growing group of large-scale platforms](#) that aggregate data generated by millions of patients, and the company also uses it to support research into public health issues—not only on behalf of members, but for subscribers including pharmaceutical companies, life sciences firms, government agencies, and academic and research institutions.

The KPSC Innovation Studio uses HEOR principles to solve problems, choosing only projects that are expected to impact numerous lives, have feasible and scalable solutions, and generate outcomes that add measurable value.

That work has helped identify patients who were eligible for clinical trials and [explored issues](#) ranging from GLP-1 prescription patterns to uptake of vitamin K for newborns to oncology drug use. Now, Truveta plans to use AI to delve more deeply into the [human genome](#), potentially helping to pave the way for the development of new treatments for rare diseases.

And soon, it expects to offer an additional service.

"We're investing in proactive intelligence, a set of capabilities where customers can learn about things they didn't ask for," Simonov said. "We can leverage our AI engines to identify signals that our members and clients should consider, because

they could suggest a meaningful impact and, potentially, opportunities to improve patient care or outcomes."

Relying on HEOR

So, how are HEOR experts contributing to these health system innovations?

The multidisciplinary team in the KPSC Innovation Studio that Funahashi oversees uses HEOR principles to solve problems, choosing only projects that are expected to impact numerous lives, have feasible and scalable solutions, and generate outcomes that add measurable value.

"We're actually measuring outcomes beyond whether a project saves lives, reduces readmissions, or improves quality, services, or costs in any way," he said. "We're also measuring the unintended consequences, because sometimes when you come up with a great solution for 1 problem, it creates a different problem for somebody else."

With AI-driven innovations, it's also important to keep humans in the loop to deploy technology in a way that keeps an organization's ideals and desired outcomes at the forefront, Simonov said.

HEOR professionals can guide that process by "demanding that the systems we use are scientifically rigorous and founded in rational and scientific thought," he said. "There are infinite questions that need to be weighed methodologically, and as a community, that's something we should all be invested in."

Finally, Simonov said, continuing to gather RWD and use it for HEOR will be essential for filling gaps in the literature about conditions like rare diseases.

"The answers AI gives will only be as good as the publications that are out there," he said, "so being able to capture that human experience and bring it to these learning systems is absolutely crucial to delivering exceptional and personalized care."

***Beth Fand Incollingo** is a freelance writer who reports on scientific, medical, and university issues.*

By the Numbers: Innovation in Healthcare

Section Editor: The ISPOR Student Network

Contributors: Ivy Nga-Weng Leong, University of Mississippi, Oxford, MS, USA; Cynthia C. Egbuemike, The University of Texas at Austin, Austin, TX, USA

Low-Tech Innovation Improves Early Detection of Postpartum Hemorrhage



Percentage of postpartum hemorrhages detected early using a calibrated drape in a randomized trial conducted at 80 hospitals in 4 African nations.

- By comparison, in hospitals where blood loss was estimated visually, only **51% of postpartum hemorrhages** later diagnosed by birth attendants were detected early.
- Severe postpartum complications or maternal death from bleeding occurred in **1.6% of patients** where the drape was part of a bundled care strategy, versus 4.3% where usual care was provided.

TAKEAWAY

A low-tech innovation for early detection of postpartum hemorrhage, combined with bundled treatment, significantly lowered the risk of severe postpartum complications.

Citation: Gallos I, Devall A, Martin J, et al. Randomized trial of early detection and treatment of postpartum hemorrhage. *N Engl J Med.* 2023;389(1):11-21. doi:10.1056/nejmoa2303966

Quality of Evidence Supporting AI Use for Prior Authorization Decisions



Percentage of studies with high overall risk of bias in a systematic review focused on the use of AI in prior authorization and coverage decisions.

- For applicability, **60% of the analyzed studies** were rated as having low concerns overall.
- AI use was most concentrated in payer review (**62.5%**), followed by postdenial appeals (**18.8%**), initial prior authorization submission (**12.5%**), and early prior authorization initiation (**6.3%**).

TAKEAWAY

Applicability of AI for prior authorization may be acceptable, but high risk of bias warrants caution when implementing AI-based models for this purpose.

Citation: Mousavi T, Tohme Shaya F, Cooke CE. Artificial intelligence in prior authorization and coverage decisions: a systematic review of methods, evidence gaps, and future implications for patient access. *J Manag Care Spec Pharm.* 2026;32(9):1135-1153. doi:10.18553/jmcp.2026.32.9.1135

Relevance of Trust in Relation to AI Implementation in Healthcare



The number of studies in which AI's perceived ability to tailor care to patient needs led to enhanced trust, out of 8 studies that assessed trust in AI in relation to implementation in healthcare.

- In **5 of 8 studies**, AI's self-learning and nontransparent and autonomous characteristics created uncertainty and threatened trust.
- AI was commonly used for diagnosis in **4 studies**. Other applications included brain modeling, image recognition, smart heart care services, treatment, surgery, and communication.

TAKEAWAY

Individual characteristics, AI characteristics, and contextual characteristics all can influence trust in AI implementation, with trust mediating actual AI use

Citation: Steerling E, Siira E, Nilsen P, Svedberg P, Nygren J. Implementing AI in healthcare—the relevance of trust: a scoping review. *Front Health Serv.* 2023;3:1211150. doi:10.3389/frhs.2023.1211150

AI-Assisted Evidence Generation and HTA: Why Transparency and Reproducibility Will Define Trust

Mallik Greene, PhD, DBA, BPharm, Abbott, Boston, Massachusetts, USA

KEY TAKEAWAYS

Generative artificial intelligence (AI) is moving into health economics and outcomes research (HEOR) and real-world evidence workflows, from evidence synthesis to modeling and digital endpoint analytics.

Guidance for reporting AI use in HEOR is an important foundation, but health technology assessment agencies and payers still need task-specific criteria for validation, human review, auditability, and decision fitness.

Organizations that apply proportionate controls early will be better positioned to preserve scientific credibility and stakeholder trust.

Introduction

When a health technology assessment (HTA) agency receives a dossier in which artificial intelligence (AI) assisted with literature screening, evidence synthesis, statistical programming, and economic modeling, what documentation should the assessors expect before trusting the evidence?

This question is already relevant across health economics and outcomes research (HEOR) and real-world evidence (RWE). AI earned first place in [ISPOR's 2026-2027 Top 10 HEOR Trends](#), rising from the third position in the 2024-2025 list. Automation of evidence synthesis was anticipated more than a decade ago; generative AI has brought that ambition closer to practical use at scale. Reports from the ISPOR Working Group on Generative AI [describe applications](#) in systematic reviews, RWE analysis, economic modeling, and dossier development, and assess opportunities, challenges, and policy considerations [specific to HTA](#). Evaluations show potential efficiency gains for evidence synthesis but also meaningful, variable errors, [reinforcing the need for human verification](#). This article focuses on generative AI, particularly large language models (LLMs), whose probabilistic outputs raise questions about reproducibility and traceability distinct from traditional machine learning.

Healthcare decision makers depend on evidence that is transparent, reproducible, and scientifically defensible. As AI enters evidence-generation workflows, stakeholders are asking: How was it used? What human oversight was applied? Can the workflow be reproduced and validated? Who remains accountable for scientific integrity?

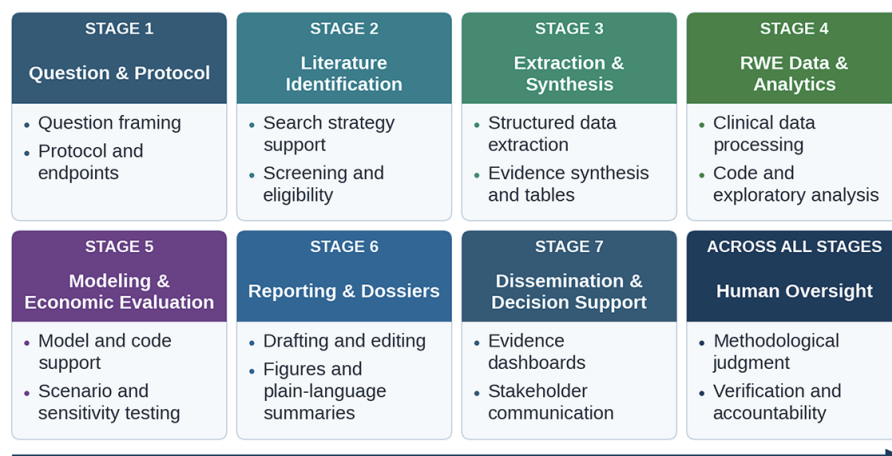
The central challenge is no longer whether AI will be used in evidence generation, but whether stakeholders will trust the resulting evidence.

Some [evaluated tools](#) already perform well in [bounded tasks](#) such as screening or structured extraction, but performance remains task- and implementation-dependent. Without safeguards, AI-assisted workflows can introduce hallucinations, undocumented decisions, automation bias, inconsistent outputs, and limited auditability, all of which [must be weighed against the potential benefits](#) of AI adoption in HEOR. Productivity alone will not establish the scientific credibility that decision makers need.

Where AI Is Already Entering Evidence Generation

AI-assisted technologies are being integrated across the evidence life cycle ([Figure 1](#)). In systematic reviews, they can support searching, screening, data

Figure 1. AI Integration Across the HEOR and RWE Evidence Life Cycle



Examples are illustrative, not exhaustive.

AI indicates artificial intelligence; HEOR, health economics and outcomes research; RWE, real-world evidence.

extraction, summarization, and evidence-table development. In RWE study planning and analysis, generative AI is being explored to process structured and unstructured clinical data and support code generation, exploratory analysis, and interpretation. It can also assist with protocols, statistical programming, economic models, dossiers, and scientific writing.

In digital health and patient-centered evidence generation, AI-enabled analytics [are being integrated with wearable and biosensor data](#) for continuous monitoring, pattern detection, and prediction, generating digital endpoints that may feed RWE studies and HTA submissions.

Most organizations are not deploying fully autonomous evidence-generation systems. AI is more commonly implemented as a copilot, with researchers, statisticians, and HEOR scientists retaining responsibility for methods, interpretation, and scientific review. Even in hybrid workflows, AI outputs can influence downstream decisions.

The Emerging Scientific Credibility Gap

Reproducibility is one of the clearest challenges. LLM outputs can change with the model version, prompt wording, configuration, and stochastic settings. HEOR guidance, including the [ELEVATE-GenAI](#) framework, therefore recommends documenting model characteristics, prompts, data sources, and validation procedures. Without this information, an analysis may be difficult to reconstruct or independently verify.

Transparency presents a related challenge. A [2025 review of 159 clinical LLM studies](#) found that exact prompt wording was reported in 49%, query dates in 51%, and stochasticity management in only 15%. Within HEOR, the editors [introducing *Value in Health's* 2025 generative AI section](#) observed that none of the articles to which ELEVATE-GenAI could have been applied would have fully met its criteria. Missing documentation limits reconstruction and obscures the boundary between automated output and human judgment.

Figure 2. Scientific Risks and Principal Safeguards for AI-Assisted Evidence

Controls should be proportionate to the task and consequences of error.

RISK AREA	PRINCIPAL SAFEGUARD
F Hallucinations & factual inaccuracies	Verify material claims against primary sources; require human fact-checking.
R Reproducibility challenges	Record model/version, prompts, settings, and source set; test repeatability.
A Automation bias & over-reliance	Use independent expert review, predefined checks, and escalation criteria.
T Transparency & traceability gaps	Maintain workflow and decision logs; disclose AI and human contributions.
P Data, privacy & security risks	Use approved secure environments; minimize and de-identify data; restrict access.
G Accountability & governance gaps	Define ownership, SOPs, monitoring, audit rights, and accountable decision makers.

TRACEABLE CONTROLS PROTECT SCIENTIFIC CREDIBILITY

AI indicates artificial intelligence; SOPs, standard operating procedures.

Hallucinations and factual inaccuracies remain important concerns, particularly in evidence synthesis. A recent systematic review [found substantial and highly variable error rates](#) across searching, screening, extraction, and risk-of-bias assessment. Automation bias can compound these risks when users accept fluent outputs without adequate source verification. These risks are summarized in **Figure 2**.

Human scientific oversight remains central. Qualified experts must verify source fidelity, critically appraise outputs, resolve discrepancies, and remain accountable for methods, interpretation, and scientific validity.

Policy and governance expectations are already evolving. The [European Union AI Act](#) establishes risk-based obligations. The [US Food and Drug Administration's 2025 draft guidance](#) proposes a context-of-use, risk-based credibility framework for AI models supporting regulatory decisions about drugs and biological products. The [US National Institute of Standards and Technology \(NIST\) Generative AI Profile](#) similarly emphasizes life cycle risk management.

HTA agencies have also begun to state their expectations. A [2024 position statement](#) from the National Institute for Health and Care Excellence in the United Kingdom asks submitters to declare and justify AI use, follow existing methods guidance, engage early when planning its use, and retain human oversight and accountability; [Canada's Drug Agency issued](#) a closely aligned statement in 2025. The [2025 HTAi Global Policy Forum](#) discussed the themes of trust, human agency, and risk-based approaches when prioritizing next steps for the integration of GenAI into HTA. Subsequent [commentary](#) has focused on translating those principles into validation and accountability in HTA practice.

What Standards Are Needed?

HEOR is not starting from zero. The ISPOR Working Group on Generative AI has published [a taxonomy, an HTA-focused report](#), and [ELEVATE-GenAI](#), which specifies what LLM-assisted studies should report about models, prompts, accuracy, and reproducibility. The remaining gap lies less in what to disclose than in how assessors will judge fitness for a decision, which validation benchmarks and shared test sets each task requires, and how organizations govern AI use across projects.

From the evidence-generator and submitter perspective, 4 principles frame the controls needed for decision-ready evidence (**Figure 3**).

First, transparency requires disclosure of where and how AI was used. Documentation should identify the purpose, model and version, prompts or prompt logic, workflow steps, data sources, validation procedures, and known limitations.

Second, reproducibility requires audit trails, version control, and workflow records that allow analyses to be reviewed, replicated, and validated. Criteria and performance thresholds should be defined before use whenever AI influences evidence selection, analytic code, or model inputs.

Third, governance requires standard operating procedures, risk classification, tool validation, monitoring, escalation pathways, and clear accountability. Controls should be proportionate to the task, uncertainty, and potential consequence of error.

Finally, human scientific oversight remains central. Qualified experts must verify source fidelity, critically appraise outputs, resolve discrepancies, and remain accountable for methods, interpretation, and scientific validity.

Lessons Learned

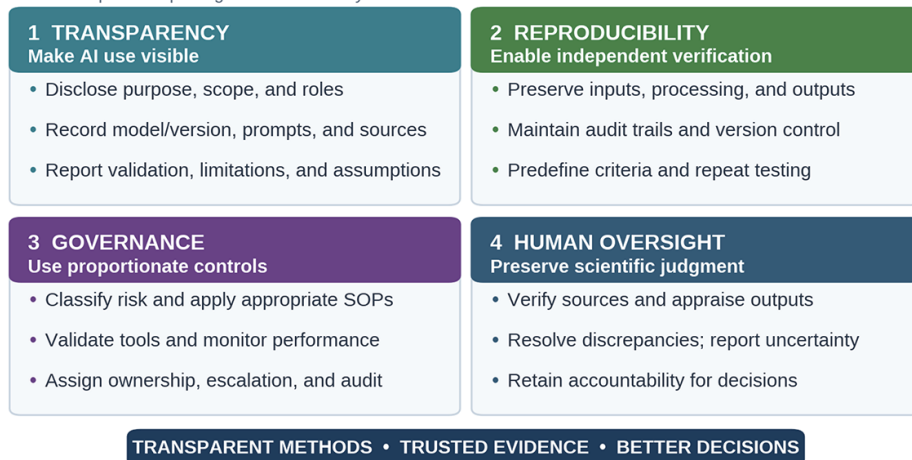
Recent evidence suggests a practical lesson: Efficiency gains are credible only when the workflow can be reconstructed. Generative AI can assist evidence synthesis, and some tools perform well in bounded tasks, but performance varies and material errors remain possible.

HTA agencies and payers need evidence-lineage records, reproducibility safeguards, and transparent validation before accepting AI-assisted evidence.

An AI-assisted literature-screening workflow may appear faster than traditional methods yet still be unsuitable for decision making if the model version, prompts, source set, inclusion criteria, and validation results are not preserved. Without that record, teams may be unable to explain or reproduce why individual studies were included,

Figure 3. Four Pillars of Trustworthy AI-Assisted Evidence Generation

From transparent reporting to decision-ready evidence



AI indicates artificial intelligence; SOP, standard operating procedure.

excluded, or summarized in a particular way. The same principle applies to code generation, RWE analysis, economic modeling, and dossier development.

A tool used only to improve wording does not require the same validation as a tool used to select evidence, generate analytic code, or influence model inputs. This distinction is central to proportionate, risk-based governance.

HTA agencies and payers need evidence-lineage records, reproducibility safeguards, and transparent validation before accepting AI-assisted evidence. HEOR and RWE teams need documented workflows, version control, and oversight calibrated to the task and consequences of error. Journals and regulators need to set clear disclosure expectations. Patient organizations may also expect transparency about whether AI contributed to evidence supporting treatment, access, or reimbursement decisions.

ISPOR has already shaped this agenda through the Working Group on Generative AI, ELEVATE-GenAI, and its [AI resources and policy work](#). The next opportunity is to translate that foundation into HTA-side acceptance criteria; task-specific validation benchmarks and shared test cases; proportionate human-review requirements; and auditability standards across systematic reviews, RWE analyses, economic models, and HTA submissions.

Consensus on decision-ready documentation would help organizations implement these principles consistently in practice.

Future progress will depend not only on what AI systems can do, but on whether the HEOR community can make their use transparent, reproducible, accountable, and worthy of trust.

Author disclosure and disclaimer:

Mallik Greene is an employee of Abbott. This article was prepared independently in the author's personal capacity. The views expressed are those of the author and do not necessarily reflect the views or positions of Abbott.

AI Use Disclosure Statement:

OpenAI ChatGPT (GPT-5.6 Pro) and Anthropic Claude were used to support editorial refinement, manuscript review, drafting of revision text, and figure-content and layout development. ChatGPT (GPT-6 Astra Pro) also supported this copyediting round. Python 3 with Matplotlib was used to render the figures; Pillow was used to replace the Figure 1 title band. All AI-assisted output was reviewed, verified against the cited sources, and edited by the author, who retains full responsibility for the scientific integrity, accuracy, and interpretation of the article.

Artificial Intelligence in HEOR: From Potential to Practice in Standards, Oversight, and Real-World Implementation

Kathy Lang, PhD; Ross Maclean, MD, Precision AQ, New York, NY, USA

KEY TAKEAWAYS

Artificial intelligence can substantially improve the efficiency of health economics and outcomes research by accelerating evidence reviews, data extraction, real-world data analyses, modeling, and reporting.

Trustworthy adoption depends on transparency, validation, governance, and accountability, with human experts responsible for study design, interpretation, methodological decisions, and final conclusions.

The future is likely a hybrid model where artificial intelligence augments but does not replace expert judgment, supported by standards, oversight frameworks, and workforce readiness.

Introduction

Health economics and outcomes research (HEOR) has long served as a cornerstone for evidence-based decision making—guiding policy, reimbursement, and clinical practice through rigorous evaluations of patient outcomes, healthcare costs, comparative effectiveness, disease burden, cost-effectiveness, and budget impact. Traditionally, HEOR has relied on structured methodologies—including systematic literature reviews, network meta-analyses, economic modeling, and observational research—to assess the value of medical interventions. Today, artificial intelligence (AI), machine learning (ML), and large language models (LLMs) are reshaping the HEOR workflow across literature review, data extraction, simulation, and reporting.

Trained human experts must remain the drivers and decision makers to ensure that AI augments decision making and doesn't replace it.

The potential for AI in HEOR is clear. Less clear is what that potential looks like in terms of responsible adoption of AI and the conditions that must be met for the technology to be trusted. Trust is essential, given that HEOR directly informs reimbursement, pricing, and access decisions that affect millions of patients and billions of healthcare dollars.

Successful adoption of AI in HEOR will depend on internal and external governance, transparency, accountability, and organizational readiness, as opposed to technological capability alone, to establish trust among health technology assessment (HTA) agencies and other decision makers.

What Constitutes Credible AI-Enabled HEOR

AI tools are being tested to automate and augment core tasks across HEOR, including the analysis of evidence bases,

identification of gaps in care, and sizing of unmet need.

- Targeted and systematic literature reviews can be expedited via natural language processing (NLP) and interactive searches.
- LLMs can apply predefined inclusion/exclusion criteria, extract key data points, and produce concise summaries.
- In [health economic evaluations](#), AI can be used to support cost-effectiveness and budget-impact models by driving country adaptations and parameter documentation.
- ML can be used in [real-world evidence \(RWE\) studies](#) to process large claims and electronic health record datasets and evaluate patient treatment patterns, healthcare resource utilization and costs, and predictors of outcomes.
- NLP can be used to [extract outcomes](#) from trials and publications to populate models.
- GenAI can be used to automate [sensitivity analyses](#), [synthetic patient scenarios](#), and [model adaptation and documentation](#), subject to fit-for-purpose validation and expert review.

While these enhancements can enable substantial time savings, in a field that relies on and prioritizes expert peer review, they must be used in a way that is transparent and verifiable if they are to be trusted and accepted. This is particularly important given the challenges associated with their use. Currently, AI cannot apply scientific judgment or anticipatory strategic thinking in drawing conclusions and making recommendations, and it cannot make methodologic judgments required in building models or interpreting evidence. It also cannot foresee and account for evolving patient, provider, payer, and HTA preferences and priorities. Furthermore, AI can infuse bias if the data on which it is trained are incomplete or nonrepresentative. As a result, unchecked use of AI creates the risk of introducing factual inaccuracies, hidden biases, inconsistent reasoning, and erroneous conclusions that

undermine the credibility of generated evidence.

The likely trajectory is a hybrid intelligence model with workflows that intentionally combine the computational power of AI with expert human judgment and guidance to ensure reproducibility and validity. Trained human experts must remain the drivers and decision makers to ensure that AI augments decision making and doesn't replace it.

Published guidance for AI use in HEOR consistently references the need for transparency, validation, accountability, and auditability, all of which necessitate human expert oversight. Early HTA position statements from the [National Institute for Health and Care Excellence](#) in the United Kingdom and [Canada's Drug Agency](#), as well as [Joint Clinical Assessment dossier preparation guidance](#) from the European Union's member state HTA Coordination Group, stress the need for human oversight, explicit justification of AI use, and full disclosure of methods and data—positioning AI to augment, not replace, expert judgment in HTA. Standards such as CHEERS-2022 and [CHEERS-AI](#) and evolving LLM reporting guidance (eg, [ELEVATE-GenAI](#)) also define consistent expectations for transparency, accuracy, and reproducibility, including specifying what to disclose when LLMs contribute to synthesis, drafting, or coding.

The ISPOR Generative AI Working Group offers a [shared taxonomy](#) (eg, retrieval-augmented generation, prompt engineering, domain fine-tuning) to improve accuracy, traceability, and reproducibility across systematic literature reviews, RWE, and modeling. The International Committee of Medical Journal Editors have outlined [similar priorities](#) related to transparency, accountability, disclosure, and responsibility in using AI for the drafting of peer-reviewed publications.

Together, these frameworks establish a shared floor of accountability that makes AI-assisted HEOR legible to HTA agencies, journals, and payers alike.

The Role of Human Accountability in Driving Credibility

As drivers of AI activity, human experts must remain responsible and

Table 1.

HEOR Activity	AI Role	Human Role
Drafting written content	Initial drafts	Oversight and accountability
Literature searching (SLR)	Initial screening for identification	Oversight, verification, and supplementation
Data extraction	Draft extraction	Oversight and verification
Economic models	Drafting with human input	Methods and parameter estimation oversight and final validated version
Value dossiers	Initial drafting	Verification and final accountability for content
Conclusions	Support only	Responsible and accountable

SLR indicates systematic literature review.

Table 2.

Governance Requirement	How to Operationalize	Why It Matters	Implementation Considerations
Prespecification of AI role in HEOR studies (planned uses)	Determine and articulate, at study outset, the methods and uses of AI that will be leveraged	Transparency, accountability	Organizations will need governance policies that define which AI uses are permitted and how to document intended use before study initiation
Documentation of use	Explicit noting of where AI has been used (eg, drafting of written sections, literature screening, data extraction, parameter estimation for HEOR models)	Transparency, credibility, accountability	SOPs should require disclosure of specific AI use and who was responsible for verification
Prompt and parameter documentation (log prompts, model assumptions, parameter sources)	Document specific prompts, assumptions, sources	Transparency, reproducibility, accountability	Version-control repository should retain source documentation for auditability and validation
Quality checking	Feed results through AI application to quality/validity check	Credibility, validity, accountability	Independent human review should verify AI outputs against predefined acceptance criteria
Map evaluation of AI intervention to existing standards (eg, CHEERS-AI)	Ensure established standards are leveraged	Credibility, reproducibility, validity	Existing reporting templates should be updated such that AI-specific reporting is routine vs optional
Align with NICE, CDA, other positions and guidance for generating AI-supported evidence	Ensure consistency with published guidance	Transparency, credibility, validity, accountability	Organizations should monitor evolving guidance and update SOPs as expectations change

AI indicates artificial intelligence; CDA, Canada's Drug Agency; CHEERS, Consolidated Health Economic Evaluation Reporting; HEOR, health economics and outcomes research; NICE, National Institute for Health and Care Excellence; SOP, standard operating procedure.

accountable for all aspects of an HEOR workflow; it is not sufficient for a human to be simply “in the loop”. The emerging distinction between *human-in-the-loop*, where experts approve each AI output at defined checkpoints, and *human-over-the-loop*, where experts set parameters and review final outputs, has direct

implications for how accountability is assigned in HTA submissions. **Table 1** illustrates how this accountability should be distributed across core HEOR activities.

Operationalizing credible AI use in HEOR requires governance policies that evolve

alongside the technology. **Table 2** outlines key controls, their rationale, and implementation considerations for HEOR.

Workforce and Organizational Readiness

To prepare for the future, HEOR professionals at all career stages will need training alongside emerging AI competencies to maintain methodologic standards in an AI-augmented environment. Core AI competencies for HEOR professionals include:

- Rigorous prompt/query design for conducting and validating research
- Methodologic validation of AI-generated outputs
- Identification of potential bias or inaccurate conclusions
- Documentation skills and awareness of internal and external AI governance rules and requirements

To guide responsible use of AI, organizations will need to build out governance policies, validation

procedures for all AI applications, version control systems, audit trails, and staff training programs. Given the rapid pace of AI advancement, these processes and procedures should be designed to be living documents that are regularly reviewed and updated by human experts as AI technologies evolve.

Conclusion

AI is not a replacement for human expertise, although it is a powerful partner. Teams that understand where AI adds value and where human judgment must prevail will lead. With appropriate standards and oversight, AI-assisted HEOR can deliver more efficient and credible evidence to support health technology decisions.

The defining challenge for the next decade of HEOR will not be whether AI can generate evidence. It will be whether organizations can establish appropriate governance, accountability, and rigor sufficient for HTA, payers, clinicians, and patients to trust AI-assisted evidence.

Disclosures

Authors Lang and Maclean are employees of Precision AQ, a consultancy offering AI-assisted HEOR services where appropriate.

ChatGPT (OpenAI) was used to assist with drafting selected sections of this paper, improving language and organization, and editing for clarity and conciseness. The conceptual framework, interpretation of the evidence, opinions expressed, and final content were developed and critically reviewed by the authors. The authors verified all factual statements, references, and citations and accept full responsibility for the accuracy, originality, and integrity of the paper.

The Verification Era: Agentic AI and the Future of Evidence Generation in HEOR

Neil Warnock, MD, MS, MadeAi, Inc, Boston, MA, USA

KEY TAKEAWAYS

Demand for evidence proving a medicine's value is outpacing the resources available to most health economics and outcomes research teams. This is a gap that greater effort alone cannot close.

Artificial intelligence can compress months of evidence extraction and synthesis into weeks.

Meanwhile, the scientist's value shifts from producing evidence to ensuring that AI-assisted evidence work is accurate, credible, and decision-ready; human judgment is more important than ever.

Introduction

The health economics and outcomes research (HEOR) discipline has reached a paradoxical inflection point. While the global demand for robust, real-world evidence (RWE) and value demonstration has never been higher—driven by the implementation of the Inflation Reduction Act (IRA) in the United States and the Joint Clinical Assessment (JCA) in the European Union (EU)—the organizational structures supporting these efforts are being dismantled.

In the 2026 Tufts Center for the Evaluation of Value and Risk in Health [HEOR Leaders Survey](#), nearly all of the 59 senior professionals from US-based pharmaceutical companies who were surveyed said their companies' senior leaders believed HEOR was critical to the firm's success and would become increasingly important in RWE-based strategic decisions. However, only 34% of the respondents (and fewer than 20% from large companies) expected the size of their team or their budget to increase as a result.

In the Verification Era, AI is not merely an efficiency tool but a structural solution to a workforce capacity problem that has no purely human answer at current resourcing levels.

To survive this capacity problem, the field must embrace agentic artificial intelligence (AI) frameworks. In the current Verification Era, AI's ability to generate large amounts of information increases the importance of verifying whether that information can be trusted. This transition redefines the HEOR scientist not as a manual data handler, but as a scientific governor of automated evidence pipelines.

Convergence of Crisis: Why the Old Model is Failing

For decades, HEOR operated as the scientific engine of value demonstration. However, recent evidence marks a

decisive inflection point: Traditional, manual workflows are no longer compatible with the pace of global market demands.

- **The IRA and the QALY vacuum.** The [Inflation Reduction Act](#) restricts the use of cost-per-quality-adjusted life year (QALY) and similar measures in Medicare price negotiations, creating a methodological vacuum for traditional value assessment frameworks. HEOR teams must now develop novel, product-specific value frameworks on accelerated timelines that they were not originally structured to meet. [More than two thirds of HEOR leaders agree](#) the IRA has driven a greater need for HEOR and RWE evidence generation.
- **The end of sequential submissions.** The EU Health Technology Assessment (HTA) Regulation's [JCA](#), operational for oncology since January 2025, requires a single coordinated submission across member states. Teams must simultaneously satisfy the diverse methodological standards of Germany's [IQWiG](#) (Institute for Quality and Efficiency in Health Care), France's [Haute Autorite de Sante](#), and the United Kingdom's [National Institute for Health and Care Excellence](#) (NICE). Evidence-generation plans designed around sequential national submissions are now structurally inadequate.
- **The organizational paradox.** Despite rising stakes, [HEOR associates have been direct targets](#) in recent restructuring cycles. Teams are being dispersed into Medical Affairs or Market Access structures, losing centralized strategic influence at the exact moment their expertise is most required.

The Solution: AI as Evidence Infrastructure

In the Verification Era, AI is not merely an efficiency tool but a structural solution to a workforce capacity problem that has no purely human answer at current resourcing levels. The focus has shifted toward agentic frameworks.

Agentic AI refers to coordinated systems of specialized AI agents that autonomously perform discrete research

tasks—such as literature screening, data extraction, and evidence synthesis—under human oversight. These AI architectures have been validated for [systematic literature reviews](#) (SLRs) and RWE generation.

Transforming SLRs

The SLR is the cornerstone of HEOR, yet it traditionally takes 4 months to 1 year or longer.

- The [A4SLR framework](#), published in *Value in Health* in 2025, coordinates 7 specialized agents (search, inclusion/exclusion criteria deployment, abstract/full-text screening, text/table preprocessing, data extraction, risk of bias analysis, report) to handle the entire life cycle of an SLR.
- The framework achieves F1 screening scores of 0.917–0.977. F1 score combines precision (how often the AI's selections are correct) and recall (how often it finds what it should) into a single 0-1 measure, with A4SLR's scores indicating very high screening accuracy.
- Data extraction F-scores exceed 0.99.
- The process compresses the timeline from months to a 2- to 4-week window.
- Each agent produces traceable reasoning chains, providing the auditability required by HTA bodies like NICE and the [European Medicines Agency](#).

Unlocking Unstructured Text

The importance of RWE is increasing, but extracting relevant information from health system databases is not always straightforward.

- Approximately 80% of clinically relevant outcome data reside in unstructured text, such as physician notes and pathology reports.
- Clinical [large language models](#) can convert these narratives into structured datasets, capturing side-effect burdens and quality-of-life proxies that traditional RWE pipelines miss.
- This capability represents a genuine expansion of the evidentiary universe without requiring additional study investment.

Defining the AI Orchestrator Role

The emergence of high-speed AI does not replace scientists; it elevates them. The HEOR professional's role is shifting from primary evidence producer to AI orchestrator and quality assurance architect. For example, instead of manually building models, screening literature, and extracting evidence, the HEOR professional directs AI agents to perform these tasks at scale, then applies scientific judgment to validate the methods, outputs, and conclusions.

The role of the health economics and outcomes research professional is shifting from primary evidence producer to AI orchestrator and quality assurance architect.

- **Governance as a scientific competency.** The primary risk of AI remains hallucination—the generation of plausible but fabricated data. HEOR experts must design the deterministic validation checks that ensure agentic extraction is accurate. A false-positive rate tolerable in early stage research is not acceptable in a final NICE submission; this judgment cannot be automated.
- **Human-in-the-loop validation.** The [ISPOR ELEVATE-GenAI](#) framework and [Academy of Managed Care Pharmacy RWE standards](#) emphasize that human-in-the-loop validation is a methodological requirement, not a best practice. The HEOR professional provides the irreplaceable quality assurance layer that makes AI-generated evidence credible to regulators and payers.

The Business Case: The Translation Imperative

The most existential challenge for HEOR is organizational. HEOR is no longer evaluated primarily on its ability to produce data, but on its ability to influence decisions under constraint.

- **Closing the translation gap.** Historically, HEOR teams have been poor at translating scientific output

into the business case language that C-suite leaders use for resource decisions.

My own experience bears this out. Running an 8-member field team, I found our models were rarely the problem. The gap was between a defensible analysis and an argument an access lead could carry into a resource meeting.

AI changes this dynamic by allowing teams to generate executive-facing summaries that translate complex cost-effectiveness analyses into budget impact and revenue protection narratives.

- **Demonstrating value.** The reported efficiency gains are substantial. A [time and motion analysis](#) of HTA-standard SLRs found hybrid AI screening was 51.4% more efficient than manual methods, fully automated screening was 81.1% more efficient, and AI-assisted extraction was 93.1% more efficient for AI-assisted data extraction. An [earlier case study](#) completed a full systematic review in 2 weeks using automation tools (with experienced human reviewers in the loop at every decision point) versus a typical timeline of approximately 1 year.

Compression of this order can advance market access decisions by 1 or more review cycles—a concrete and communicable return on investment in AI-enabled evidence infrastructure.

The larger return is upstream in the development process. Evidence synthesis that takes a year is late by default—commissioned after the value story is set, defending decisions instead of shaping them. With a time frame of 2 weeks, evidence synthesis can start early and run continuously: competitor assumptions can be tested while trial design is still open and evidence gaps flagged while there is time to close them. Most of the evidence problems I have dealt with involved analysis that arrived too late to change anything.

Implications for HEOR Practice

The transition to AI-enabled evidence generation requires structural adjustments within HEOR organizations.

- Redefining the HEOR skill set. HEOR professionals will increasingly need competencies in AI governance, validation workflows, and computational research design in addition to traditional health economic modeling and epidemiology.
- Integrating AI into evidence-generation pipelines. Rather than treating AI tools as isolated productivity applications, organizations should integrate agentic AI systems directly into systematic literature review, evidence synthesis, and real-world data pipelines.
- Establishing methodological governance frameworks. Human-in-the-loop validation processes must be formally defined, including audit checkpoints for AI-assisted literature screening, data extraction, and outcome classification.
- Strengthening HEOR's strategic role. AI-enabled efficiency allows HEOR teams to shift focus from manual evidence generation toward earlier strategic engagement in clinical development, pricing strategy, and payer evidence planning.
- Building the training pipeline. These competencies are not yet taught systematically, and most current practitioners will acquire them mid-career. Professional societies and employers will need to establish structured curricula and credentialing in areas such as AI validation and governance, rather than relying on self-directed learning.

Conclusion

HEOR functions are operating in an environment that demands more evidence, earlier, and at higher methodological standards—demands that are outpacing available resources. AI, deployed through validated agentic frameworks and governed by the

methodological expertise of HEOR professionals, offers the only credible path forward. The Verification Era does not diminish the scientist; it redefines the role around the use of rigorous scientific judgment to govern AI-generated evidence.

The transition is not without risk. Reported screening gains come with low precision, and sufficiency measured without verification will be overstated.

The remedy is largely specified already: [Canada's Drug Agency](#) and [NICE](#) require that AI use be declared, justified, and validated, the capable, informed human in the loop.

Teams that build in these controls from the outset will not have to retrofit them under submission pressures.

Treating Multiple Sclerosis Early and Aggressively in Mexico: What the Numbers Say

Karina Carrillo, MSc, National Institute of Neurology and Neurosurgery, Mexico City, Mexico; Víctor Gasca, MD, Mexican Social Security Institute, Mexico City, Mexico; Margarita Fuentes, MBA; Iliana González, MD, Merck S.A. de C.V., State México, Mexico; Pedro Oliva, MD, MSc; Nadia López, MSc, Mexico City, Mexico

KEY TAKEAWAYS

In Mexico, patients with relapsing-remitting multiple sclerosis (MS) who start high-efficacy therapy early spend 3.3 more years able to walk without aid and maintain activities compared with those prescribed an escalation approach.

Despite higher drug costs, early high-efficacy treatment saves \$13,161 per patient over 15 years compared to escalation, mainly by avoiding caregiver costs, disability aids, and productivity losses.

Delaying high-efficacy therapy is more expensive and riskier clinically. These findings support updating formulary and access policies in Mexico toward earlier, more effective MS treatment.

Four Realities of MS Treatment in Mexico

Imagine being diagnosed with multiple sclerosis (MS) in Mexico. Depending on where you live and which healthcare system you have access to, the treatment you receive or don't receive could be vastly different from the patient in the next city. MS is a chronic immune-mediated disease that attacks the nervous system, and its most common form, relapsing-remitting MS (RRMS), causes episodic flare-ups that [progressively accumulate neurological damage](#). Left uncontrolled, it can lead to severe, permanent disability.

Disease-modifying therapies (DMTs) represent the most significant therapeutic innovation in MS of the past 3 decades. Unlike symptomatic treatments, DMTs act on the underlying immune mechanisms of the disease, reducing the frequency of relapses and slowing the accumulation of disability. Moderate-efficacy agents include injectable interferons, glatiramer acetate, and oral options like teriflunomide and dimethyl fumarate; high-efficacy therapies include natalizumab, ocrelizumab, alemtuzumab, and cladribine tablets. The distinction matters: In clinical trials, high-efficacy DMTs [reduce annualized relapse rates](#) by 50–70% relative to moderate-efficacy agents, and observational data suggest they are associated with substantially lower long-term disability accrual.

Disease modifying therapies represent the most significant therapeutic innovation in multiple sclerosis of the past 3 decades.

Today in Mexico, 4 treatment realities coexist: some patients receive no DMTs at all; others get moderate-efficacy agents; some follow an escalation path starting with moderate-efficacy drugs and switching to high-efficacy ones only

when the disease breaks through; and a growing number are now starting high-efficacy therapy early, right from diagnosis. This heterogeneity is driven not by clinical preference but by structural barriers: delayed diagnoses, [unequal access across public and private systems](#), and limited availability of innovative therapies in the public sector.

The central question [our research](#) sets out to answer is a practical one: over 15 years, which approach delivers better outcomes for patients and for the health system's budget?

What We Did

We built a Markov model (a simulation tool widely used in health economics) to follow a cohort of adult patients with active RRMS over a 15-year horizon. The model tracked patients moving between 4 disability states, defined by the Expanded Disability Status Scale (EDSS):

- Mild disability (EDSS <4, patients ambulatory and largely independent)
- Moderate disability (EDSS 4–6, patients with limited walking ability or requiring some assistance)
- Severe disability (EDSS ≥6, patients needing a wheelchair or bedridden, highly dependent on caregivers)
- Death

Transition probabilities between these states were drawn from [published clinical literature](#), calibrated to reflect how each treatment strategy actually slows disease progression.

We modeled all 4 strategies head-to-head. Costs were assigned to each disability state and included everything relevant to a real patient's burden: drug acquisition (with wastage factored in), hospitalizations, relapse management, neurology consultations, rehabilitation, out-of-pocket expenses, caregiver time, assistive devices, absenteeism, presenteeism, and early disability pensions. All costs were expressed in

US dollars, applying a standard 5% annual discount rate. All cost inputs are original model estimates derived from Mexican institutional sources and national labor data; systematic reviews by [Samjoo et al](#) and [Schauf et al](#) informed the selection of cost categories included in the model.

What We Found

The results tell a clear story and one that should matter to clinicians, payers, and policy makers alike.

Clinically, early high-efficacy treatment changes the trajectory of the disease.

As shown in **Table 1**, patients on early high-efficacy DMTs spent 12.8 of their 15 modeled years in the mild disability state, compared with just 9.3–9.5 years in all other strategies. Just as striking: under early high-efficacy treatment, patients accumulated zero years in the severe disability state, compared with 1.2–1.6 years under the other approaches. In practical terms, that means delaying years of dependence on caregivers and wheelchairs, along with lost productivity.

Economically, paying more upfront for better drugs costs less overall.

This is the counterintuitive finding that deserves the most attention (**Table 2**). Early high-efficacy DMTs carry the highest direct drug costs (\$155,186 over 15 years), yet their total discounted cost is \$172,931, which is \$13,161 less than the escalation strategy (\$186,092). Why? Because early disease control dramatically reduces what gets spent on caregivers, home support, disability devices, and lost wages. Nonmedical and indirect costs under early high-efficacy treatment are 57% lower than under no-DMT scenarios. Those savings more than offset the higher drug bill.

The escalation strategy, by contrast, combines the drug costs of 2 sequential therapy lines with the disability costs accumulated while waiting for a treatment switch, making it both more expensive and less clinically efficient than early high-efficacy DMT initiation.

Lessons Learned

What worked

- Adopting a societal cost perspective was a key methodological choice and one that changed the conclusions. From a narrow healthcare budget

perspective, early high-efficacy treatment appears to be the most expensive option due to higher drug costs. It is only when caregiver burden, productivity losses, and disability aids are included that the full picture emerges: early high-efficacy treatment becomes the more cost-efficient option compared to escalation. Decision makers who operate under a strict healthcare budget perspective may still weigh these trade-offs differently depending on their mandate.

- Modeling all 4 coexisting strategies rather than just comparing 2 gave decision makers a realistic picture of the Mexican landscape. It surfaced the counterproductive economics of the escalation paradigm in a way a simpler comparison would not.
- Applying EDSS-based health states made clinical and economic outcomes directly linkable: Each year spent in the severe disability state has a calculable cost in caregiving, devices, and lost productivity.

Limitations

- Not including quality-adjusted life years (QALYs) limits the study's formal cost-effectiveness claims. A future version should incorporate utility weights so that the results can feed directly into health technology assessment (HTA) decisions, which typically require an incremental cost-effectiveness ratio.
- Transition probabilities came from published literature, not Mexican real-world data. Local registry data, when available, would make this model significantly more credible to national regulators.
- The 15-year horizon may still understate the full benefit. MS is a lifelong disease; extending the simulation to 20–25 years is likely to widen the cost advantage of early high-efficacy treatment as disability divergence compounds over time.

Implications for stakeholders

- Payers and HTA bodies: Consider incorporating indirect and nonmedical

Table 1. Clinical Outcomes by Treatment Strategy (15-year horizon)

Outcome	No DMTs	Moderate-efficacy DMTs	Escalation	Early High-efficacy DMTs
Years of life gained	13.9	14.7	15.0	15.0
Years with mild disability (EDSS <4)	9.3	9.5	9.5	12.8
Years with moderate disability (EDSS 4–6)	3.0	3.9	4.3	2.1
Years with high disability (EDSS ≥6)	1.6	1.3	1.2	0.0
Years of life lost	1.1	0.3	0.0*	0.0*

*Both escalation and early high-efficacy strategies achieved near-zero years of life lost (0.02).

Green = most favorable; red = least favorable per row.

DMT indicates disease-modifying therapy; EDSS, Expanded Disability Status Scale.

Table 2. Total Discounted 15-year Costs per Patient by Treatment Strategy (USD)

Cost category	No DMTs	Moderate-efficacy DMTs	Escalation	Early High-efficacy DMTs
Direct medical (USD)	\$117,073	\$107,710	\$148,767	\$155,186
Direct nonmedical (USD)	\$25,626	\$25,697	\$26,114	\$11,187
Indirect costs (USD)	\$11,489	\$11,251	\$11,211	\$6558
TOTAL (discounted, USD)	\$154,188	\$144,657*	\$186,092	\$172,931

*Lowest total cost strategy. Exchange rate: MXN 18.9244/USD (June 25, 2025). 5% annual discount rate applied.

Green = lowest cost; red = highest cost. DMT indicates disease-modifying therapy.

costs into assessment frameworks, as a budget-only perspective may underestimate the full societal burden of MS disability beyond drug costs.

- Formulary committees in public institutions ([IMSS](#), [ISSSTE](#), [IMSS Bienestar](#)): This analysis provides cost and outcomes data that can inform access decisions. The findings suggest that restricting high-efficacy agents to patients who have already failed moderate-efficacy therapy may generate higher downstream disability-related costs; this evidence can be one input in formulary deliberations.
- Clinical guideline developers: The escalation paradigm is not just clinically suboptimal. This analysis shows it is also economically inefficient in the Mexican healthcare context. Guidelines should reflect that evidence.

Why This Matters Beyond Mexico

The story here is not unique to Mexico. In many middle-income countries, treatment access for MS is shaped by the same structural constraints: inadequate diagnosis infrastructure, fragmented payer systems, and risk aversion toward high-cost drugs. The lesson from this analysis is transferable: When you account for the full downstream costs of uncontrolled disability, early intensive treatment often costs less, not more than sequential escalation.

The HEOR community has a role to play in making that argument visible to the right audiences. Models like this one are decision tools. Their value lies not just in the numbers they generate, but in placing those numbers in front of the people who set formularies, write guidelines, and allocate health budgets

and helping them see that spending on better treatment today is, in many cases, an investment that pays for itself.

Funding & Disclosures

This study was sponsored by Merck Biopharma Distribution S.A. de C.V., Mexico, an affiliate of Merck KGaA, Darmstadt, Germany (CrossRef Funder ID: 10.13039/100009945). Margarita Fuentes and Iliana González are employees of Merck Biopharma Distribution S.A. de C.V., Mexico, an affiliate of Merck KGaA, Darmstadt, Germany. All other authors declare no additional conflicts beyond the sponsorship relationship. The work was presented as poster EE130 at ISPOR Europe 2025, Glasgow, UK, November 9–12, 2025.

Smartphone-Based Screening Could Change the Economics of Obstructive Sleep Apnea Care

Marcel Braun, PhD, MBA, University of Duisburg-Essen, Essen, Germany

KEY TAKEAWAYS

Obstructive sleep apnea (OSA) remains substantially underdiagnosed despite being linked to cardiovascular disease, accidents, and premature mortality.

Smartphone-based screening, using a built-in microphone to analyze overnight breathing sounds, may help close this gap at scale by identifying individuals at elevated risk for OSA who would benefit from confirmatory sleep testing.

A German health economic model suggests smartphone-based OSA screening would reduce costs and improve outcomes over 10 years, while remaining cost-effective over a lifetime horizon.

A Silent Epidemic Hidden in Plain Sight

Obstructive sleep apnea (OSA) is one of the most common chronic disorders that health systems still struggle to detect efficiently. Although prevalence estimates suggest that [up to 50% of adults](#) may be affected in some countries, a large proportion of moderate-to-severe cases remain undiagnosed for years. This diagnostic gap matters because untreated OSA is not simply a sleep disorder. It is [associated with](#) increased risks of cardiovascular disease, stroke, road traffic accidents, reduced workplace productivity, and premature mortality.

The downstream burden extends far beyond sleep clinics and affects employers, insurers, and healthcare systems broadly, and furthermore, the condition is associated with substantial economic impact through healthcare utilization, accidents, and productivity losses. A 2023 systematic review of 31 studies reported annual cost estimates ranging [from €1,669 to €5,186 per patient](#) and found that OSA-related costs can increase years before diagnosis.

Yet the conventional diagnostic pathway is difficult to scale. Traditional instrumental sleep assessments require specialist infrastructure, trained personnel, and waiting lists that often delay diagnosis and treatment initiation. The first-line treatment for OSA is [positive airway pressure \(PAP\) therapy](#), which aims to

maintain airway patency during sleep via a face mask and has been associated with reduced risk of all-cause mortality and cardiovascular mortality in patients with OSA. Because PAP requires regular use to prevent airway obstructions, the extent to which modeled benefits are realized in practice depends on successful initiation as well as sustained and consistent use.

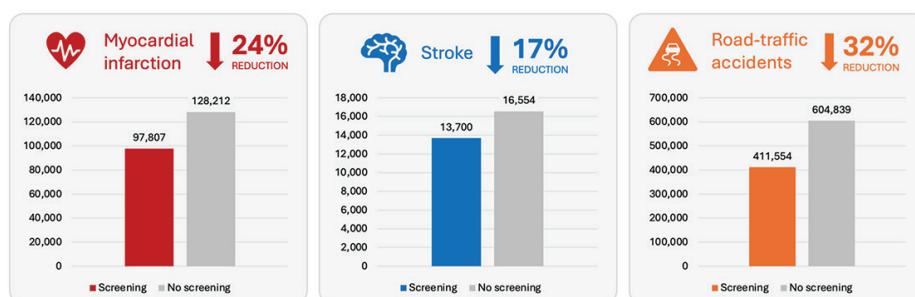
This raises an important question for stakeholders: What if diagnosis could be shifted earlier and made substantially more accessible? Digital health technologies, particularly smartphone-based tools, may offer an answer.

Why Smartphone-Based Screening Matters

Smartphone-based applications that record nocturnal breathing sounds through the phone's built-in microphone and analyze acoustic patterns to estimate the risk of OSA represent an emerging approach to accessible, inexpensive screening in which app-based identification of an elevated risk triggers a referral for confirmatory sleep testing. The smartphone technology should not be considered a replacement for gold-standard diagnostics, but it can change who enters the diagnostic pipeline in the first place. In [our study](#), presented at ISPOR Europe 2025, we evaluated the cost-effectiveness of a smartphone-based screening application for identifying individuals at risk for moderate-to-severe OSA.

Figure 1. Earlier detection and treatment of OSA through smartphone-based screening can prevent serious health events while reducing burden on the healthcare system (population-based modeling with ~5.6 m. women and men aged 50 years and time-horizon of 10 years).

Modeled clinical and economic outcomes over 10 years: smartphone screening versus no screening.



The concept is straightforward:

- Large populations can be screened at low cost and low friction.
- High-risk individuals can then be referred for confirmatory testing.
- Earlier diagnosis may enable earlier treatment initiation.
- Earlier treatment could reduce downstream complications and costs.

This “triage-first” approach resembles digital transformation already occurring in other disease areas, such as [atrial fibrillation screening](#), where scalable front-end screening tools help prioritize limited specialist resources. Importantly, the economic value is not generated by the app alone. It emerges from changing the timing of diagnosis and intervention.

What We Modeled

We developed a Markov model from the perspective of German statutory health insurance (SHI) comparing smartphone-based screening versus no systematic screening.

The model simulated a cohort entering at age 50 years and followed it over a 10-year horizon (approximately to age 60 years) and over a lifetime horizon. The model incorporated:

- screening uptake
- confirmatory sleep investigation and treatment initiation
- long-term treatment adherence
- cardiovascular complications, road traffic accidents, mortality
- quality-adjusted life-years (QALYs)
- healthcare costs and the incremental cost-effectiveness ratio (ICER) calculated as the difference in costs divided by the difference in QALYs between screening and no screening

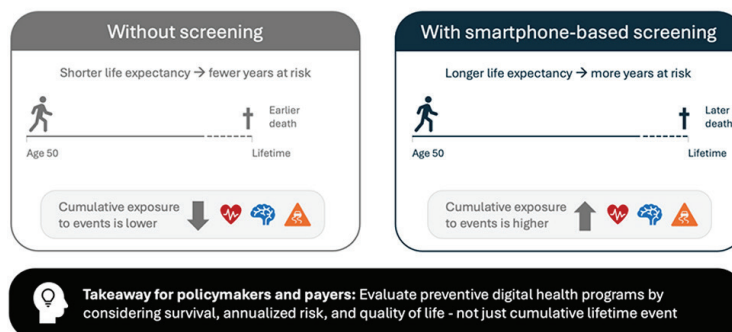
The model also reflected an important real-world challenge: diagnosis alone is insufficient. Clinical benefits depend on patients successfully progressing through the pathway from screening to confirmatory testing to sustained treatment adherence.

The Modeled Economic Signal Was Strong

Over a 10-year horizon, smartphone-based screening generated more QALYs at lower cost than no screening, and was hence economically dominant in both

Figure 2. Longer lifetime changes the interpretation of lifetime event counts in economic assessment of preventive screening technologies. Over a lifetime horizon, longer modeled survival generates additional person-time at risk and changes the age distribution of the surviving population. Absolute event counts and crude rates that are not age-standardized should therefore be interpreted alongside survival and QALY outcomes.

Modeled clinical and economic outcomes over the lifetime horizon: smartphone screening versus no screening.



male and female populations. In practical terms, this means the strategy both improved outcomes and reduced costs.

Over the 10-year horizon, compared with no screening, the model projected:

- approximately 18% lower healthcare costs
- 20% higher QALYs
- fewer clinical events (**Fig. 1**) and
- lower mortality

Over the lifetime horizon, the model projected higher total costs with screening because longer modeled survival generated additional years of healthcare use. Nevertheless, the resulting lifetime ICERs remained below the illustrative cost-effectiveness benchmark of 50,000 EUR applied in the analysis.

Survival and Outcomes

One of the more interesting findings from the analysis was how strongly survival influenced long-term outcome interpretation. Over a lifetime horizon, individuals in the screening arm lived substantially longer, resulting in more cumulative years at risk for age-related health events (**Figure 2**).

This highlights an important principle in preventive health economics: effective interventions can sometimes generate higher cumulative lifetime numbers of events simply because patients survive

longer. This leads to more events occurring because more life-years are gained, and not because outcomes worsen.

In our model, when outcomes were interpreted relative to survival time, screening was consistently associated with lower annualized risk exposure, particularly for road traffic accidents.

What if diagnosis could be shifted earlier and made substantially more accessible? Digital health technologies, particularly smartphone-based tools, may offer an answer.

The broader message is that preventive digital health interventions should be evaluated not only by cumulative event counts, but also by their effects on survival, quality of life, and time at risk. For policy makers and payers, this distinction becomes increasingly relevant as digital screening programs are evaluated over long time horizons and across large populations.

A New Bottleneck

Digital screening technologies are advancing rapidly. This means that, in many situations, organizational

challenges may be greater than technological ones. Our analysis suggests that the economic value of smartphone-based OSA screening depends heavily on implementation factors, such as screening uptake, confirmatory diagnostic capacity, treatment initiation, and long-term therapy adherence.

This means digital screening tools only create value if the corresponding downstream pathway can handle increased volume. A scalable front-end screening strategy could overwhelm sleep laboratories if diagnostic capacity is not expanded simultaneously. Likewise, poor treatment adherence could erode much of the projected long-term benefit.

This has important implications for payers and health systems considering digital OSA programs. Reimbursement discussions should not focus exclusively on the digital tool itself, but on the performance of the entire care pathway.

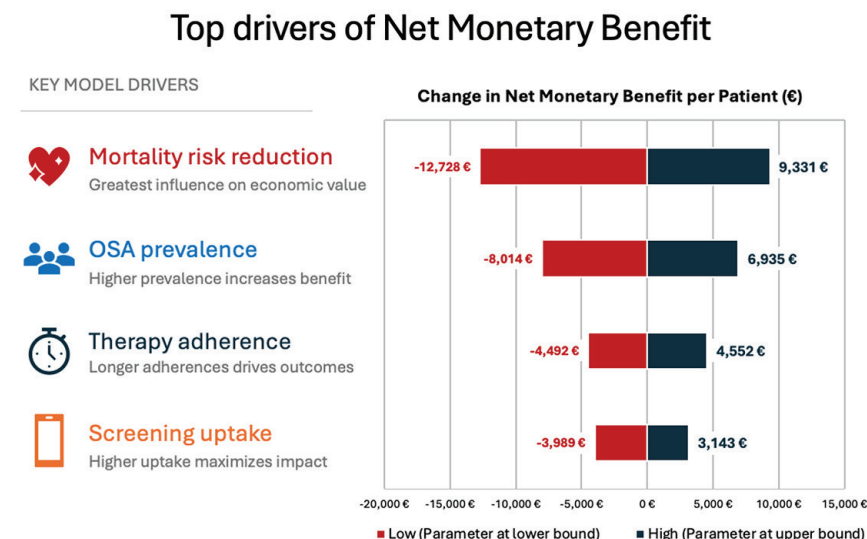
Lessons Learned

This work also highlights broader lessons for HEOR evaluations of digital health technologies.

Digital interventions should not be assessed solely as standalone products. Their value often depends on pathway redesign and operational integration. The economics of preventive digital interventions may also depend on the time horizon selected. Short-term analyses may show cost savings, whereas lifetime analyses often capture the economic consequences of improved survival.

Implementation assumptions matter. In our model, parameters such as treatment adherence and disease prevalence had greater impact on results than the screening cost itself (**Figure 3**). It may look paradoxical that reduced mortality is the strongest driver of economic value while also raising lifetime costs. Net monetary benefit reconciles the 2 findings by weighing the monetized value of health gained against the additional spending. Preventing premature death adds many QALYs, and at conventional willingness-to-pay thresholds, the value of those years far outweighs the extra age-related care that typically accompanies a longer life. Mortality reduction lifts both costs and

Figure 3. Top drivers of Net Monetary Benefit: One-way sensitivity analysis showing how the economic value of smartphone-based OSA screening responds when each key model parameter is varied across its plausible range.



health, but the health gain is by far the larger movement.

Published economic evaluations of digital health interventions are [heterogeneous in their methods and analytical perspectives](#), which limits comparison across interventions. Our model necessarily simplifies real-world processes, such as spontaneous diagnosis, switching of treatment modalities, or digital literacy barriers.

Next Steps for Decision Makers

For healthcare systems facing growing demand and capacity constraints, smartphone-based OSA screening may represent an opportunity to modernize access to care. But implementation should proceed pragmatically.

Three priorities appear especially important:

- Integrate screening into structured pathways: without confirmatory testing of patients identified as at-risk, followed by treatment of those who are confirmed, screening is unlikely to deliver meaningful value.
- Measure real-world adherence to treatment and retention: long-term effectiveness in chronic conditions—which affects the value of screening—depends heavily on sustained treatment use.

- Evaluate equity and digital access: smartphone-based pathways must not widen disparities among older adults or digitally underserved populations.

The broader implication extends beyond sleep medicine. OSA may become an important example of how health systems can evaluate and operationalize scalable digital screening technologies more generally.

AI disclosure

The author used an AI-based language tool to improve the English-language phrasing and readability of this article. The tool was not used to generate content, conduct analysis, or interpret results. The author reviewed and edited all text and is solely responsible for the accuracy and integrity of the work.

AI-Estimated Depression Scores as Real-World Evidence of Treatment-Resistant Depression

Pedro Alves; Carl D. Marci, MD; Costas Boussios, PhD, OM1, Inc., Boston, MA, USA

KEY TAKEAWAYS

The PHQ-9 is a validated patient-reported measure of depression severity but is inconsistently documented in electronic medical records, creating a significant evidence gap.

The ePHQ-9—an AI model that estimates PHQ-9 scores from clinical notes—had a higher correlation with physician-attested treatment-resistant depression than the directly observed PHQ-9 across a real-world cohort of more than 76,000 patients.

The ePHQ-9 may serve as a valuable proxy when patient reports are unavailable in real-world datasets.

Major depressive disorder (MDD) is one of the most prevalent and disabling conditions worldwide [with widespread economic impact](#), yet measuring its severity in routine clinical practice remains surprisingly difficult. The [Patient Health Questionnaire 9-item](#) (PHQ-9) is the standard tool for quantifying depressive symptoms, but clinicians often omit it from the medical record, leaving large gaps in real-world data (RWD) and no objective markers of disease severity. For health economists, outcomes researchers, and healthcare decision makers who rely on RWD to track treatment effectiveness, this gap is more than an inconvenience—it is a fundamental limitation to advancing therapies and knowledge.

Our team of researchers at OM1, Inc. set out to fill this gap. In a study [presented at ISPOR Europe 2025 in Glasgow](#), the team tested whether an artificial intelligence model called the ePHQ-9—which infers PHQ-9 depression severity scores directly from clinical notes in electronic medical records (EMRs)—could serve as a reliable proxy when a conventional PHQ-9 score is absent. We aimed to determine the usefulness of ePHQ-9 for identifying one of the most clinically meaningful presentations of severe MDD, namely treatment-resistant depression (TRD).

The PHQ-9 Problem in Real-World Data

The PHQ-9 is a 9-item questionnaire, typically completed by the patient, that scores depressive symptoms on a 0–27 scale across 5 severity categories:

- None-Minimal (PHQ-9 = 0–4)
- Mild (PHQ-9 = 5–9)
- Moderate (PHQ-9 = 10–14)
- Moderate-Severe (PHQ-9 = 15–19)
- Severe (PHQ-9 = 20–27)

Clinicians and researchers alike regard PHQ-9 as a gold standard for evaluating depression severity and treatment response. The problem is that it only appears in the medical record when someone remembers—or chooses—to administer it. In large, population-level real-world datasets, this is far from guaranteed, with many patients who have

only a few observed scores over time and [others with no severity scores at all](#).

The ePHQ-9 was developed to help fill the gaps in real-world measurement. Using a previously published machine learning model trained on clinical notes from a large condition specific dataset, it estimates what a patient's PHQ-9 score would be based on the language clinicians use to describe their recent history, symptoms, and course of illness. A published [validation study](#) demonstrated strong analytic performance of the model in estimating PHQ-9 scores from both structured (ie, templated) and unstructured (ie, free text) note elements. The current study takes the next step: Does the ePHQ-9 reflect clinically meaningful variation in disease burden?

The ePHQ-9 estimates what a patient's depressive symptom score would be based on the language clinicians use in their notes to describe the patient's recent history, symptoms, and course of illness.

Treatment-Resistant Depression: A Meaningful Benchmark

TRD—broadly defined as MDD that fails to respond adequately to at least 2 antidepressant treatments of adequate dose and duration—is [a well-recognized clinical entity](#) affecting a substantial proportion of patients with MDD.² Critically, TRD has no ICD diagnostic code despite the fact that they tend to suffer with [more severe and more debilitating symptoms](#) compared with depressed patients without TRD. Identifying patients with TRD from real-world datasets requires natural language processing (NLP) of clinical notes to find physician attestations of the condition—exactly the kind of rich clinical signal that the ePHQ-9 was built to harness. This makes TRD an ideal real-world benchmark for testing a severity measure derived from clinician narratives.

Study Design

The study drew on a large, curated US real-world dataset of patients with MDD (OM1, Inc., Boston, MA). The dataset combines linked claims and EMR data with both structured and unstructured clinical notes. The final cohort comprised 76,873 patients, of whom 28,641 had an observed PHQ-9 score and 61,602 had an ePHQ-9 score recorded within 30 days of their index date.

Physician-attested TRD was identified via NLP of clinical notes (for example, “Continue Abilify to 7.5 mg PO QHS for treatment-resistant depression symptoms and irritability/agitation”), with human coders reviewing a subset as validation. We found that 1699 patients were labeled as TRD, including 1082 with an observed PHQ-9 and 1232 with an ePHQ-9 score. To evaluate the relationship between depression severity and TRD, the team examined the proportion of TRD patients within each of the 5 PHQ-9 severity categories.

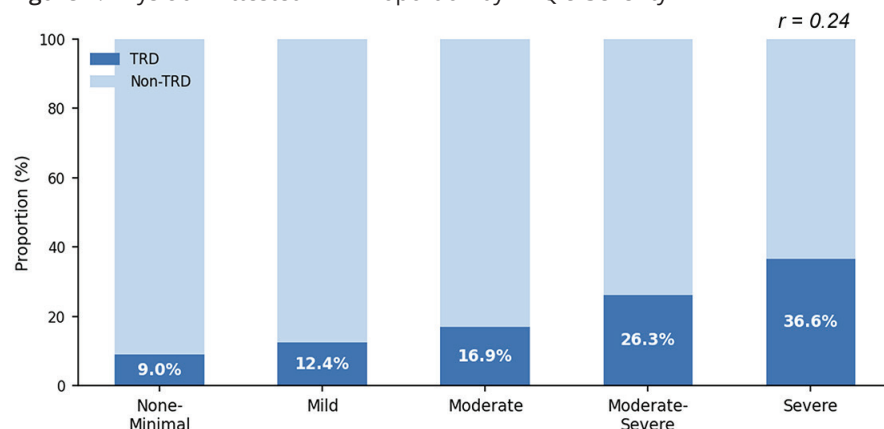
A Clear Signal: Stronger Relationship in the AI Model

Both the observed PHQ-9 and the ePHQ-9 revealed a modest, monotonic pattern: As depression severity increased, so did the proportion of patients with physician-attested TRD (Figures 1 and 2). This is exactly what one would expect if both measures are genuinely capturing disease burden.

Any tools that extend the reach and fidelity of real-world evidence are not merely methodological conveniences—they are essential to informed decision making.

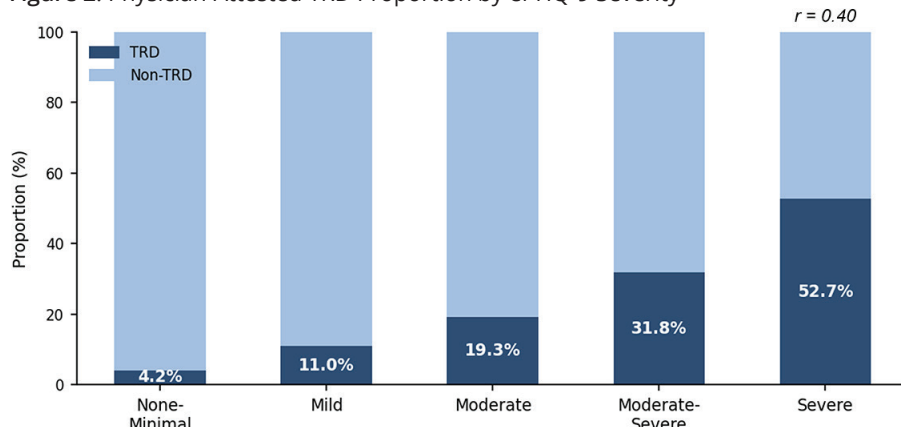
The difference lies in the strength of that association. For the observed PHQ-9, TRD prevalence rose from 9% in the None-Minimal category to 36.6% in the Severe category, yielding a Pearson correlation of $r = 0.24$. The ePHQ-9 told a more vivid story: TRD prevalence ranged from 4.2% in the None-Minimal group to 52.7% in the Severe category, with a stronger correlation of $r = 0.40$. Put differently, more than half of all

Figure 1. Physician-Attested TRD Proportion by PHQ-9 Severity



Pearson correlation: $r = 0.24$ | $n_{\text{TRD}} = 1082$; $n_{\text{non-TRD}} = 4328$ (fixed 1:4 subsample). TRD indicates treatment-resistant depression; PHQ, Patient Health Questionnaire.

Figure 2. Physician-Attested TRD Proportion by ePHQ-9 Severity



Pearson correlation: $r = 0.40$ | $n_{\text{TRD}} = 1232$; $n_{\text{non-TRD}} = 4928$ (fixed 1:4 subsample). TRD indicates treatment-resistant depression; ePHQ, machine-learning based Patient Health Questionnaire.

patients classified as severely depressed by the ePHQ-9 also carried a physician attestation of TRD—a convergence that supports the clinical relevance of the AI-derived measure.

Why Does the AI Model Outperform?

The finding that the ePHQ-9 correlates more strongly with TRD than the observed PHQ-9 may seem counterintuitive given the ubiquitous use of the PHQ-9. The explanation likely lies in the source of the data. The PHQ-9 is a patient-reported outcome completed at a single point in time, often in a waiting room or asynchronously online prior to a clinical encounter. The score reflects their subjective assessment of symptoms over the prior 2 weeks. As such, it reflects only the patient perspective and may not capture the full clinical picture.

The ePHQ-9, by contrast, is derived from clinician narratives: the language physicians use when documenting a patient's history. This includes clinical judgment combined with patient reports of symptom burden, treatment course, and response over time. Because TRD is also identified through physician attestation in clinical notes, the 2 measures share a common clinical “vocabulary,” producing a strong alignment. Both the ePHQ-9 and TRD attestation, in effect, speak the language of the treating physician, which may account for the results of this study.

Implications for HEOR and RWE

These findings carry practical implications for health economics and outcomes research. TRD is expensive, burdensome, and increasingly the target of novel therapeutic interventions.

Yet studying it in real-world settings has been hampered by the absence of a diagnostic code and the limited availability of validated severity measures interrogated into EMR data.

The ePHQ-9 offers a potential solution on both counts. With its ability to generate depression severity estimates across a far larger patient population than the observed PHQ-9 (61,602 versus 28,641 patients in this study), and its stronger alignment with clinically attested treatment resistance, it may enable more robust and comprehensive RWD research into MDD and its most refractory forms.

While the RWD in this study were limited by documentation quality and consistency and the potential for selection bias, the findings demonstrate the potential value of AI-enabled innovations like the ePHQ-9. For healthcare decision makers evaluating comparative effectiveness, health

technology assessments, and coverage policies, any tools that extend the reach and fidelity of real-world evidence are not merely methodological conveniences—they are essential to informed decision making.

Lessons Learned

This study represents a meaningful step in the validation of AI-derived clinical measures for use in RWD research. While more research is needed involving additional datasets from additional health networks, the ePHQ-9's stronger correlation with physician-attested TRD than the conventional patient-reported PHQ-9 underscores the potential value of mining clinician narratives for outcomes research.

As EMR data grow richer and NLP methods continue to mature, studies like this help support a new generation of AI-enabled measures that can fill the gaps where traditional instruments fall short.



Translating World-Class Research Into Innovations With Real Impact in Germany

Interview with **Maïke Sander, MD**, Scientific Director & Chair of the Board, Max Delbrück Center, and Vice President, Helmholtz Health

Photo credit: Pablo Castanogla / Max Delbrück Center

Q&A

After almost 3 decades in the United States, Maïke Sander, MD, returned to Germany with firsthand experience of an ecosystem where scientific discovery, entrepreneurship, investment, and industry intersect. Now leading the [Max Delbrück Center](#), she is working to strengthen those connections in Germany while building on the distinct advantages of Europe's research model. From cell therapy and AI-enabled discovery to venture building and the culture of scientific entrepreneurship, Sander sets out what it will take to turn world-class research into innovations with real impact.

“AI predictions still need to be experimentally validated. We are certainly not at a point where AI alone will give us the next therapy.”

— Maïke Sander

PharmaBoardroom: What has shaped your scientific career, and what motivated your return to Germany after almost 30 years in the United States?

Maïke Sander: I trained as an MD in Germany and then spent almost 30 years pursuing my research career in the United States, particularly in California, where I joined the University of California system immediately after medical school. My focus has been stem cell biology, particularly in diabetes. Career-wise, I became director of the Pediatric Diabetes Research Center at the University of California San Diego, where I was also co-director of the Center on Diabetes within the Institute of Engineering in Medicine.

Within the West Coast ecosystem, I saw firsthand how the handshake between academia, fundamental research, investors, large pharmaceutical companies, and small biotech can turn discoveries into something that really matters to society. I witnessed this impact as my foundational research enabled a biotech startup to initiate clinical trials, a pipeline that was later acquired and advanced by a major pharmaceutical company.

Over those same 3 decades, I saw fundamental and discovery science in Germany become stronger, supported by substantial government investment, and watched its quality and breadth evolve considerably. What I also saw, however, was a critical gap in the translational pipeline connecting laboratory discoveries to commercial markets. That became a major motivation for returning to Europe.

Having experienced an ecosystem in the United States where the different components come together very effectively, I felt I could bring some of that experience back and help shape the ecosystem in Germany and Europe.

PB: How does the Max Delbrück Center fit within Germany's biomedical research ecosystem, and what sets its approach apart?

MS: The Max Delbrück Center is part of the Helmholtz Association, Germany's largest scientific research organization. Being part of Helmholtz gives us access to an interdisciplinary scientific ecosystem of a scale that an individual institution could not achieve on its own. We collaboratively advance infrastructure and scientific expertise across universal health research domains, with an emphasis on AI, data science, and biomedical engineering.

I firmly believe that scientists conducting creative, discovery research should not be expected to become business entrepreneurs to see their work commercialized.

Our research is disease-agnostic, so rather than focusing on one particular disease, we pioneer technologies and approaches that can enable new ways to diagnose, prevent, and treat disease, and then work with clinical partners to identify where they can have the greatest impact. Cell therapy is a good example: CAR-T cell approaches developed at the Max Delbrück Center were initially directed towards cancer, but we are now seeing the potential of related approaches in autoimmune and other diseases.

PB: What is needed to translate more of Germany's scientific excellence into medical innovation and patient impact?

MS: For us, it is really about enabling scientists to take those next steps by creating the right structures around them. We have reconceptualized what was previously our Technology Transfer Office and created an Innovation & Entrepreneurship Department with a much broader remit that goes beyond patenting to support business development, industry partnerships, venture building, and spin-offs around promising scientific discoveries. What we are trying to do is bring that expertise much closer to the labs and the science.

I firmly believe that scientists conducting creative, discovery research should not be expected to become business entrepreneurs to see their work commercialized. Our approach at the Max Delbrück Center is to build an umbrella of expertise around them, particularly in areas such as venture building and regulatory affairs. Investigators can then remain focused on what they do best, while having access to the capabilities needed to connect their discoveries with industry and move them toward the market.

When we look at why ecosystems such as Boston or the US West Coast have been so successful, however, the differences are not

purely structural. Culture and individual attitudes towards risk emerge as important factors. At the Max Delbrück Center, we are bringing in people who have successfully built companies to share their experiences with our trainees, because it is difficult to envision a career path for yourself if you have never met someone who has taken it.

Alongside that cultural shift, Germany also needs to consider whether it wants to create a more venture-friendly environment, from its regulatory frameworks and the time it takes to move from an idea to a venture to taxation and the conditions for high-risk investment. Investors will always consider where and how they can generate a return, including the prospects for a product once it reaches the market. Germany remains relatively slow in this respect, and the overall conditions for high-risk investment are still less favorable than in some of the world's leading innovation ecosystems. Building a stronger translational ecosystem will require both cultural and structural progress.

PB: What potential does cell therapy hold for type 1 diabetes, and how could artificial intelligence accelerate the development of new treatments?

MS: Refined insulin therapy regimens have made huge strides in improving both quality of life and life expectancy, while helping to prevent many long-term complications. Yet that only tells part of the story. For many people with type 1 diabetes, there is hardly a waking moment when they can stop thinking about the disease. They constantly have to consider when and what they eat, how they time it, and what that means for their insulin. Even with the technologies and apps that have tremendously improved diabetes management, the psychological burden remains enormous.

At the same time, we need to communicate the timelines for developing new therapies more clearly and avoid overpromising how quickly they can reach patients. Safety is paramount, particularly when treatments can carry serious or even life-threatening risks.

I think we will increasingly see teams working on data science-enabled predictions alongside scalable validation platforms that allow us to establish whether those predictions are mechanistically true.

Bringing a fundamentally new therapy from discovery to market can take decades, even within a well-functioning ecosystem. CAR T-cell therapy is one example, and cell therapy for type 1 diabetes is another. We are also getting close in Parkinson's disease, where efforts began around the same time as those in diabetes and followed a remarkably similar path. The timelines have progressed largely in parallel, which suggests there may be limits to how much we can shorten the overall development process without compromising the validation needed to ensure safety.

Where I do expect us to gain time is on the discovery science side, particularly as we enter a new era of AI-enabled research.

AI can help us predict new drug targets and, in the case of molecular interventions, identify where a compound might alter the properties of a protein, while automation and scalable screening capabilities can accelerate laboratory processes that previously relied on individual scientists working manually.

However, AI predictions still need to be experimentally validated. We are certainly not at a point where AI alone will give us the next therapy. I think we will increasingly see teams working on data science-enabled predictions alongside scalable validation platforms that allow us to establish whether those predictions are mechanistically true.

PB: Where would you like to see the Max Delbrück Center in the next 5 years, and what can Europe learn from the United States while building on its own strengths?

MS: If the end goal is to bring a new medicine to market, the Max Delbrück Center cannot achieve that alone. We need interconnected ecosystems in which institutions and different players work together, including through stronger partnerships with industry. For us, success is therefore about becoming even better at what we do best: producing innovative, validated discovery science that can then be taken forward by the next players in that ecosystem.

At the same time, we need to become better at thinking through what a discovery could ultimately enable, whether that is a new diagnostic, a new therapeutic, or perhaps a new way of preventing disease, and then carry out the early validation needed to establish that potential. An important part of this is creating patents that have real value, because discoveries are unlikely to reach the market without strong intellectual property. For me, an important measure of the Max Delbrück Center's future success will therefore be our ability to combine scientific breakthroughs with validated, patentable inventions that can be taken forward by the next entity in the ecosystem.

Europe often looks to the United States as the exemplary innovation ecosystem, and there is certainly much we can learn from it. But Europe has important strengths of its own. Germany, for example, has a relatively coordinated approach to funding science and provides long-term support for discovery research. The US system is incredibly dynamic, which is one of its great strengths, but the short-term nature of much of its funding can hamper the longer-term thinking required to create innovation. For me, the opportunity is therefore not simply to replicate the US model, but to understand what works in each system and build on the strengths we already have in Europe.

Membership Has Its Privileges!

Unlock exclusive benefits by joining ISPOR—
The Professional Society for Health Economics
and Outcomes Research.



Member Benefits Include:

- 1-year subscription to the HEOR Learning Lab™
- 1-year online subscription to *Value in Health* (including online access to all past issues)
- 25% discount for members on open access publishing fees for *Value in Health*
- Access to member groups (eg, special interest groups, task force review groups, chapters)
- Networking through online member communities and volunteer opportunities
- Discounts on conference, Summit, short course, and ISPOR Education Center course registrations

Go to ISPOR.org today to join or renew your membership.



ISPOR

Improving healthcare decisions

ISPOR-The professional society for health economics and outcomes research
505 Lawrence Square Blvd, South
Lawrenceville, NJ 08648 USA
www.ispor.org