RB Weekly AI Brief - Issue 15 - 15.07.2026

Covering the week of 15.07.2026 · Issue 15 of the RB Weekly AI Brief

Recurring themes: Regulatory & HTA Signals (3 of last 4 issues) · Regulation & Policy (3 of last 4 issues) · Healthcare & Life Sciences (3 of last 4 issues) · Models & Research (3 of last 4 issues)

AI News Roundup

Regulatory & HTA Signals

No qualifying HTA news items identified this week. This section requires stories from official HTA body sources or specialist health policy outlets — general AI regulation stories are excluded.

Regulation & Policy

EU Extends High-Risk AI Classification Consultation to July 23 (ONGOING)

[ONGOING] — The European Commission extended its targeted consultation on draft guidelines for classifying high-risk AI systems under the AI Act by four weeks to 23 July 2026, following stakeholder requests. The draft guidelines, published 19 May 2026, interpret the high-risk conformity assessment test broadly, potentially capturing more AI systems in regulated products — including medical devices and health insurance pricing tools — than companies may have anticipated. Final guidelines are due by end of 2026, with high-risk Annex III obligations now deferred to December 2027 under the AI Omnibus agreement.

So what? *Pharma, medtech, and digital health companies have a narrowing window — closing 23 July 2026 — to submit formal feedback that could materially shape how the Commission defines 'high-risk' AI in the medicines lifecycle, clinical decision support, and health insurance pricing, with the final text directly determining compliance scope and timelines.*

European Commission

Healthcare & Life Sciences

Takeda Signs $600M AI Drug Discovery Deal With Insilico Medicine

[IN CASE YOU MISSED IT] — On July 2, 2026, Takeda and Insilico Medicine announced a strategic collaboration in which Insilico's Pharma.AI platform — covering target identification, de novo molecule generation, and clinical trial probability forecasting — will lead early-stage discovery across Takeda's therapeutic areas. Takeda receives exclusive worldwide rights to develop and commercialise any selected candidates. Insilico earns approximately $60 million in near-term payments, with the total deal potentially reaching $600 million through preclinical, clinical, commercial, and sales milestones plus tiered royalties.

So what? *As AI-generated drug candidates approach pivotal Phase III trials in 2026, pharma and HTA professionals should expect pipeline disclosures to increasingly cite AI-derived evidence packages — including AI-predicted clinical transition probabilities from platforms like InClinico — which will require new frameworks for evaluating the evidentiary quality and reproducibility of computationally designed molecules in submissions and value dossiers.*

Insilico Medicine

Models & Research

OpenAI Releases GPT-5.6 Family with Science-Tuned Flagship Sol

On July 9, 2026, OpenAI launched the GPT-5.6 model family — Sol, Terra, and Luna — to general availability following a two-week government-gated safety review tied to the models' advanced cybersecurity and scientific reasoning capabilities. Sol, the flagship, is specifically tuned for biology, chemistry, and cybersecurity, sets a new state of the art on the Artificial Analysis Coding Agent Index at 80 points, and scored 60.5 on HealthBench Professional — up 8.7 points over GPT-5.5. OpenAI has implemented "Trusted Access for Biology Research" and "Trusted Access for Cyber" programs to gate the most sensitive capabilities to vetted organisations.

So what? *GPT-5.6 Sol's domain-specific tuning for biology and chemistry — combined with its gated Trusted Access programmes — signals that frontier AI models are moving toward regulated, credentialed deployment models that pharma R&D, HEOR, and medical affairs teams will need to qualify and validate before integrating into evidence generation or regulatory submission workflows.*

OpenAI

FLI Safety Index: No AI Lab Scores Above C+, Safety Pledges Eroding

The Future of Life Institute published its Summer 2026 AI Safety Index on July 7, evaluating nine leading AI developers — including Anthropic, OpenAI, Google DeepMind, Meta, xAI, DeepSeek, and Mistral — across 37 indicators in six domains using a GPA grading system. Anthropic led with a C+, followed by OpenAI and Google DeepMind at C, while xAI, DeepSeek, and Mistral all received failing grades. The independent expert panel found that multiple top labs had weakened or eliminated earlier commitments to pause development if systems approached specified danger thresholds, with the report concluding that voluntary self-regulation is insufficient as capabilities accelerate.

So what? *For pharma, market access, and HTA professionals deploying AI in regulatory submissions, HEOR analyses, or clinical decision support, the finding that even top-ranked AI developers score no better than C+ on independent safety governance metrics underscores the need for robust internal AI validation, audit trails, and vendor due-diligence processes when selecting AI partners for high-stakes evidence generation.*

Future of Life Institute

Academic Paper Summaries

Selected from PubMed · Published within the last 12 months · New selections each week

Domain Paper — HEOR / Health Economics / Market Access

Breaking barriers in women's pelvic health: claims-based economic analysis and healthcare utilization of an AI care program compared to usual care.

Pereira AP, M Seet A, Domingues B, et al. · Journal of medical economics · 2026

#HTA · #MarketAccess · #RealWorldEvidence

This study compared healthcare spending for women using an AI-guided pelvic care programme against those receiving usual in-person care, using a matched cohort of 602 women drawn from a US nationwide insurance claims database. Women in the AI programme generated mean per-person savings of around $3,082 over 12 months, driven mainly by fewer surgical procedures, alongside improvements in symptom burden, work productivity, and mental health. For pharmaceutical and healthcare executives, this is a rare example of an AI intervention evidenced on the endpoints payors actually care about — utilisation and spend — rather than diagnostic accuracy alone.

PMID: 41733548	PubMed →	DOI →

AI Research Paper 1

Large Language Model Performance and Clinical Reasoning Tasks.

Rao AS, Esmail KP, Lee RS, et al. · JAMA network open · 2026

#ClinicalAI · #Diagnostics · #Regulation

This study rigorously tested 21 leading AI language models, including GPT-5 and Claude 4.5, on their ability to perform real-world clinical reasoning across the full patient care workflow—not just multiple-choice questions. Top models showed meaningful but variable performance, with notable differences across diagnostic and management tasks. For healthcare executives, this provides a more credible benchmark for evaluating which AI tools are genuinely ready for clinical deployment versus those that only perform well on simplified tests.

PMID: 41973425	PubMed →	DOI →

AI Research Paper 2

Large Language Model Analysis of Reporting Quality of Randomized Clinical Trial Articles: A Systematic Review.

Srinivasan A, Berkowitz J, Friedrich NA, et al. · JAMA network open · 2025

#ClinicalAI · #DrugDevelopment · #Regulation

This study built and validated an automated AI pipeline using GPT-4o-mini to check whether published clinical trial reports follow established CONSORT reporting standards, then applied it to over 21,000 trials spanning decades. The system performed comparably to human expert reviewers and revealed widespread, persistent gaps in reporting quality across disciplines and funding sources. For pharmaceutical and healthcare executives, this highlights both a scalability solution for research quality oversight and the systemic reporting deficiencies that can undermine evidence-based decision-making and regulatory submissions.

PMID: 40875232	PubMed →	DOI →

This brief is produced using an AI-assisted workflow with expert human editorial review. Stories are curated and sources verified before publication. Readers are encouraged to consult source links directly.