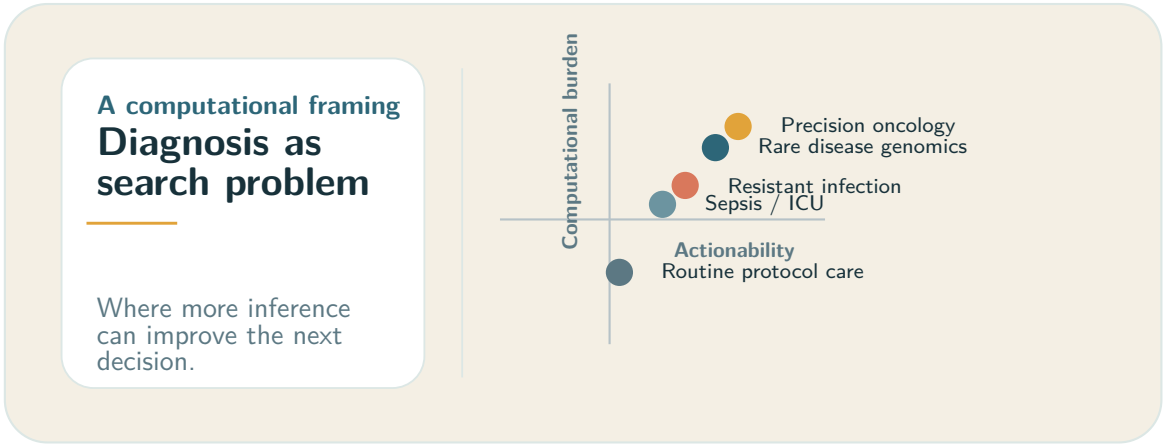
A Computational View of Medicine

Why AI will matter most where disease becomes a search problem

Birger Moëll
Co-Founder, Eir Space

March 17, 2026

Thesis. The most important medical use of advanced AI is not generic automation. It is high-context reasoning over difficult, patient-specific problems where better synthesis can change diagnosis, treatment choice, escalation, or trial matching. In these settings, disease is better understood as a computational burden distributed across time, modalities, and decision branches than as a static label alone.



Abstract

AI will not improve every part of medicine equally. Its highest-value role is emerging in disease classes that generate large patient-specific search problems across genomics, pathology, imaging, longitudinal records, treatment history, and rapidly changing clinical context. This paper argues for a computational view of medicine: one that treats many serious illnesses as reasoning problems whose difficulty is partly measurable and whose outcomes can improve when synthesis improves. Drawing on evidence from precision oncology, genome-scale rare disease diagnosis, antibiotic discovery, sepsis early warning, and longitudinal record modeling, the paper proposes a practical framework for deciding where additional compute should be deployed. It also argues that as patients gain access to their own data, compute itself becomes part of personal health infrastructure rather than something that lives only inside institutions.

1 Introduction

Medicine has long been organized around disease labels, specialty boundaries, and care pathways. Those abstractions are necessary, but they are incomplete. In many of the cases that matter most, the main challenge is not naming the condition. It is navigating a search space: identifying which facts matter, reconciling conflicting signals, selecting among branching interventions, and updating the working model as new evidence arrives.

This paper argues that AI should be evaluated through that lens first. The core question is not whether a model can sound clinical. It is whether more inference can change the answer in settings where the answer is difficult, high-stakes, and patient-specific. That framing is consistent with the broader case for human-plus-machine medicine made by Topol, but it shifts the emphasis from augmentation in general to computational leverage in particular [12].

Definition. A *computational view of medicine* treats disease management as the problem of converting fragmented, multimodal, longitudinal evidence into better patient-specific decisions under uncertainty.

2 A framework for computational leverage

Five properties determine whether a disease class is likely to benefit from larger inference budgets:

1. **Search-space size**: how many plausible hypotheses, targets, interventions, or sequences must be evaluated.
2. **Data burden**: how many modalities and how much longitudinal context are required to reason well.
3. **Personalization**: how much the correct answer depends on the specific patient rather than the average protocol.
4. **Actionability**: whether better synthesis can materially alter diagnosis, escalation, treatment choice, or trial eligibility.
5. **Compute elasticity**: whether additional reasoning depth is likely to improve the result instead of merely restating the same answer more fluently.

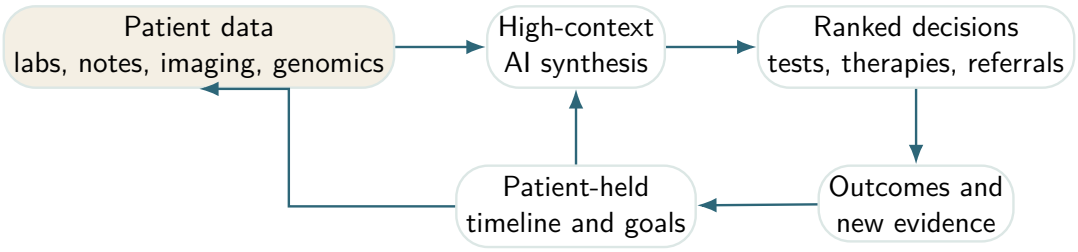


Figure 1: The patient-compute loop. In high-burden disease, value comes from repeated synthesis over the full timeline rather than one-shot question answering.

Under this framework, precision oncology, rare disease genomics, resistant infection, and certain critical care problems rise to the top. They combine large search spaces with heterogeneous data and decisions whose quality directly affects outcome.

Why prevalence is not enough. Common diseases are not automatically the best targets for large inference budgets. The relevant question is whether more synthesis can change a meaningful patient-specific decision. That is why prevalence rankings and compute-opportunity rankings diverge.

3 Evidence from the highest-leverage domains

3.1 Precision oncology

Cancer is already a computational problem. Molecular diagnosis, resistance interpretation, treatment sequencing, and trial matching require synthesis across pathology, imaging, genomics, prior therapy, and time. Personalized cancer vaccines make that explicit: the clinical workflow depends on identifying patient-specific tumor mutations, predicting which neoantigens are immunogenic, and manufacturing a bespoke intervention. In a landmark study, Sahin and colleagues demonstrated that individualized RNA mutanome vaccines could mobilize poly-specific immunity against melanoma [9]. Structure prediction systems such as AlphaFold further expand the search surface by making protein consequence analysis more tractable across vast molecular spaces [5].

The importance of compute here is not cosmetic. It sits inside the decision itself: which mutation matters, which escape mechanisms are plausible, which combination is worth pursuing, and which clinical trial is defensible for this patient now. Oncology is one of the clearest examples of a disease domain where the burden of reasoning has become part of the disease burden itself.

3.2 Rare disease genomics

Rare disease diagnosis is where long-context reasoning becomes immediately clinically useful. The problem is not only finding a variant. It is aligning phenotype, pedigree, prior negative workup, and the growing literature fast enough to affect care. In 2024, Wojcik and colleagues reported in the *New England Journal of Medicine* that genome sequencing materially improved diagnosis for patients with suspected rare disease and supported its use as a first-tier test in appropriate settings [13].

This is a canonical compute-heavy domain: large hypothesis spaces, sparse signal, fragmented records, and a direct payoff when the reasoning succeeds. The same logic applies to critically ill infants, undiagnosed

neurologic disease, and multisystem pediatric cases where small delays compound. When diagnosis arrives late in these settings, the cost is not abstract. It shows up in additional procedures, avoidable admissions, and years spent navigating the wrong branch of care.

3.3 Resistant infection and antibiotic discovery

In antimicrobial resistance, the search problem has two layers. The first is patient-specific treatment selection: organism identity, susceptibility, host state, prior exposure, organ function, and drug interactions. The second is discovery itself: how to search chemical space for compounds with novel activity. Stokes and colleagues used deep learning to identify halicin, demonstrating that machine learning can materially accelerate antibiotic discovery in regions of chemical space that were not obvious to standard pipelines [11].

This is exactly the kind of problem where more compute can change outcomes: better regimen ranking in the clinic and better exploration of candidate space in the lab.

3.4 Sepsis and critical care

Critical care is different from oncology and rare disease, but it is no less computational. Here the challenge is time. Dozens of measurements update continuously, and the relevant signal is often distributed across trends rather than single observations. Adams and colleagues showed that a machine-learning sepsis early warning system deployed prospectively across multiple hospitals was associated with faster antibiotic administration and lower mortality [2]. Rajkomar and colleagues had earlier shown that deep learning on electronic health records could scale across broad clinical prediction tasks using the raw longitudinal record itself [8].

These results matter because they establish two parts of the same claim: longitudinal records contain clinically actionable signal, and systems that can hold more of that context together can improve timing-sensitive decisions.

3.5 Patient access as a force multiplier

The computational view becomes more important, not less, as patients gain access to their own records. OpenNotes showed that patients value direct access to clinicians' notes and that such access can change engagement with care [3]. Follow-on work showed that patients often shared transparent visit notes with family members and caregivers, turning the note into a practical coordination tool rather than a static archive [4]. More recently, Salmi and colleagues showed in a proof-of-concept study that patients can use large language models on open notes in ways that help them interpret and work with their records [10]. Once records become portable and patients can assemble their own longitudinal timelines, compute becomes part of their effective care capacity. The patient is the one actor present for the entire course of illness; what has historically been missing is tooling that lets them reason over that full timeline.

4 Case examples

Table 1 shows why a single mental model fits these otherwise different domains.

Domain	Core data	Why compute matters	Representative evidence
Precision oncology	Genomics, pathology, imaging, treatment history	Patient-specific target prioritization, resistance modeling, trial matching, bespoke vaccine design	Sahin et al. [9]; Jumper et al. [5]
Rare disease	Phenotype timeline, pedigree, genome sequencing, prior workup	Diagnosis depends on long-context synthesis across sparse and distributed clues	Wojcik et al. [13]
Resistant infection	Microbiology, susceptibilities, organ function, prior antibiotic exposure	Better ranking of therapies and faster search across chemical or regimen space	Stokes et al. [11]
Sepsis / ICU	Streaming vitals, labs, notes, medication events	Earlier detection and escalation from weak, time-distributed signals	Adams et al. [2]; Rajkomar et al. [8]
Patient-held records	Notes, meds, labs, imaging reports, symptoms, goals	Converts raw access into reasoning leverage before clinical encounters	Delbanco et al. [3]

Table 1: High-value disease areas under a computational lens.

Illustrative translational case
A widely discussed 2025 media report described how an Australian entrepreneur used sequencing, variant analysis, and AI-assisted reasoning to help design a personalized mRNA treatment strategy for his dog Rosie, who had a mast cell tumour [1]. This is not high-level evidence in the way a randomized trial is. It is useful for a different reason: it makes visible what the computational workflow looks like once motivated people can assemble data, models, and synthesis outside a traditional institutional pipeline.

5 What should be built

The implication is not that every patient should be left alone with a chatbot. The implication is that health systems, startups, and regulators should prioritize tools that make high-burden reasoning portable, inspectable, and patient-aligned.

1. **Patient-assembled case files.** Systems should ingest records from multiple providers and preserve chronology rather than flattening everything into isolated visits.
2. **Explicit uncertainty handling.** Models should separate evidence, inference, and missing information, especially in oncology, genetics, and infection.
3. **Decision-oriented outputs.** The output should not just summarize. It should rank the next tests, consultations, or therapeutic branches.

Study type	What changed	Why it matters for the thesis
Pragmatic randomized clinical trial in cardiovascular screening	AI-enabled ECG alerts were tested as an intervention rather than a retrospective classifier [6]	Supports the claim that computational systems can change downstream decisions and outcomes, not merely generate scores.
Pragmatic cluster-randomized deterioration trial	Real-time surveillance for patient deterioration was tested across live inpatient settings [7]	Strengthens the case that high-context clinical surveillance can function as an operational care system.
Patient-record sharing study	Patients with transparent notes shared them with family and caregivers, extending the practical reach of the record [4]	Shows that once patients have access to records, those records begin acting as coordination infrastructure.
Patient-side LLM proof of concept	Patients used large language models on open notes to better understand and work with their records [10]	Gives direct support to the argument that personal compute can become part of the patient toolkit.

Table 2: Evidence that the computational view extends beyond retrospective prediction.

4. **Re-runnable reasoning.** Every major disease course changes over time. The right unit is a living case model, not a one-time answer.
5. **Patient-clinician collaboration.** The design target is a stronger clinical conversation, not consumer isolation.

6 Risks and constraints

The computational view does not weaken the need for evidence or clinical oversight. It sharpens it. High-burden disease is precisely where hallucinated synthesis, hidden bias, and brittle prompting are most dangerous. Better reasoning systems must therefore be auditable, source-grounded, and explicit about what they do not know.

There is also a distributional risk: if compute meaningfully improves outcomes, unequal access to compute will become a health inequity. Data access without reasoning access will be incomplete empowerment. A future health stack that gives patients records but not useful synthesis will still leave too much leverage with the institution.

7 Operational consequences

If this view is correct, the agenda for health AI changes. Product teams should stop treating every workflow as though it deserves the same level of model complexity. Clinical systems should separate protocol problems from search problems. Regulators should ask not only whether a model is accurate on average, but whether it improves the quality of reasoning in the cases where reasoning is the bottleneck. And patient-facing tools should be built around timeline integrity, evidence provenance, and re-runnable case synthesis rather than one-shot conversational convenience.

The practical payoff is concentration. More compute should go to the places where it buys something real:

fewer missed mutations, faster rare-disease diagnoses, better antibiotic choices, earlier sepsis escalation, clearer preparation for specialist visits, and better continuity when care fragments across institutions. Results from pragmatic intervention studies in ECG alerting and patient-deterioration surveillance suggest that this shift can move from theory into deployed clinical operations when the system is attached to a real decision path [6, 7].

8 Conclusion

The next decade of medical AI should be judged less by how often it automates paperwork and more by how often it helps solve the hard cases. The diseases that matter most here are not simply the most prevalent ones. They are the ones that become search problems when viewed at patient resolution. Precision oncology, rare disease genomics, resistant infection, and certain critical care settings are early proof that this lens is practical, not rhetorical.

The broader shift is cultural as much as technical. Once patients can assemble their own records and run high-context reasoning over them, compute becomes part of care capacity itself. That is why a computational view of medicine matters: it gives us a way to aim AI where additional inference has the best chance of changing outcomes.

References

[1] ABC Health Report. How a dog owner used AI to help design a cancer treatment for rosie. https://www.abc.net.au/listen/programs/healthreport/dog-owner-ai-cancer-treatment/105462166, 2025. URL https://www.abc.net.au/listen/programs/healthreport/dog-owner-ai-cancer-treatment/105462166. Accessed 2025-09-30.

[2] Roy Adams, Katharine E. Henry, Anirudh Sridharan, Hossein Soleimani, Andong Zhan, Nishi Rawat, Lauren Johnson, David N. Hager, Sara E. Cosgrove, Andrew Markowski, et al. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nature Medicine*, 28(7):1455–1460, 2022. doi: 10.1038/s41591-022-01894-0.

[3] Tom Delbanco, Jan Walker, Jonathan D. Darer, Joann G. Elmore, Henry J. Feldman, Suzanne G. Leveille, James D. Ralston, Stephen E. Ross, Elisabeth Vodicka, and Valerie D. Weber. Open notes: Doctors and patients signing on. *Annals of Internal Medicine*, 153(2):121–125, 2010. doi: 10.7326/0003-4819-153-2-201007200-00008.

[4] Sara L. Jackson, Roanne Mejilla, Jonathan D. Darer, Natalia V. Oster, James D. Ralston, Suzanne G. Leveille, Jan Walker, and Tom Delbanco. Patients who share transparent visit notes with others: characteristics, risks, and benefits. *Journal of Medical Internet Research*, 16(11):e247, 2014. doi: 10.2196/jmir.3363.

[5] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021. doi: 10.1038/s41586-021-03819-2.

[6] Chin-Sheng Lin, Wei-Ting Liu, Dung-Jang Tsai, Yu-Sheng Lou, Chiao-Hsiang Chang, Chiao-Chin Lee, Wen-Hui Fang, Chih-Chia Wang, Yen-Yuan Chen, Wei-Shiang Lin, et al. Ai-enabled electrocardiography

alert intervention and all-cause mortality: a pragmatic randomized clinical trial. *Nature Medicine*, 30 (5):1461–1470, 2024. doi: 10.1038/s41591-024-02961-4.

[7] Rebecca Moses, Lu Chen, Irtiza Khan, Zachary Pruitt, Andrew Sharp, Yu Ruan, Yiyun He, Jason Stifter, Hossein Estiri, Stephen A. Collins, et al. Real-time surveillance system for patient deterioration: a pragmatic cluster-randomized controlled trial. *Nature Medicine*, 2025. doi: 10.1038/s41591-025-03609-7.

[8] Alvin Rajkomar, Eyal Oren, Kai Chen, Andrew M. Dai, Nissan Hajaj, Michaela Hardt, Peter J. Liu, Xiaobing Liu, Jake Marcus, Mimi Sun, et al. Scalable and accurate deep learning with electronic health records. *npj Digital Medicine*, 1(1):18, 2018. doi: 10.1038/s41746-018-0029-1.

[9] Ugur Sahin, Evelyna Derhovanessian, Michael Miller, Ben Kloke, Peter Simon, Markus Löwer, Thomas Bukur, Anat Tadmor, Ulrike Luxemburger, Benedikt Schrörs, et al. Personalized rna mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature*, 547(7662):222–226, 2017. doi: 10.1038/nature23003.

[10] Liz Salmi, Dana M. Lewis, Jennifer L. Clarke, Zhiyong Dong, Rudy Fischmann, Emily I. McIntosh, Chethan R. Sarabu, and Catherine M. DesRoches. A proof-of-concept study for patient use of open notes with large language models. *JAMIA Open*, 8(2), 2025. doi: 10.1093/jamiaopen/ooaf021.

[11] Jonathan M. Stokes, Kevin Yang, Kyle Swanson, Wengong Jin, Andrés Cubillos-Ruiz, Nicole M. Donghia, Craig R. MacNair, Shawn French, Lindsey A. Carfrae, Zoe Bloom-Ackermann, et al. A deep learning approach to antibiotic discovery. *Cell*, 180(4):688–702.e13, 2020. doi: 10.1016/j.cell.2020.01.021.

[12] Eric J. Topol. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1):44–56, 2019. doi: 10.1038/s41591-018-0300-7.

[13] Monica H. Wojcik, Gabrielle Lemire, Eva Berger, Maha S. Zaki, Mariel Wissmann, Wathone Win, Susan M. White, Ben Weisburd, Dagmar Wieczorek, Leigh B. Waddell, et al. Genome sequencing for diagnosing rare diseases. *The New England Journal of Medicine*, 390(21):1985–1997, 2024. doi: 10.1056/NEJMoa2314761.