npj Digital Medicine

https://doi.org/10.1038/s41746-026-02717-z

Article in Press

Hubris in the age of intelligent medicine—from clinical hierarchies to algorithmic amplification

Received: 8 February 2026

Accepted: 27 April 2026

Cite this article as: Patel, M.S., Scales, D., Arslan, J. *et al.* Hubris in the age of intelligent medicine—from clinical hierarchies to algorithmic amplification. *npj Digit. Med.* (2026). https://doi.org/10.1038/s41746-026-02717-z

Milit S. Patel, David Scales, Janan Arslan, Paola Pedrelli, Matthew G. Crowson, Rawan Abulibdeh & Leo Anthony Celi

Hubris in the Age of Intelligent Medicine — From Clinical Hierarchies to Algorithmic Amplification

Milit S Patel[1,2], David Scales[3], Janan Arslan[4,5], Paola Pedrelli[6,7], Matthew G Crowson[8,9], Rawan Abulibdeh[10,11,12], Leo Anthony Celi[2,13,14‡]

Affiliations:

1. *Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX, USA.*
2. *Laboratory for Computational Physiology, Massachusetts Institute of Technology, Cambridge, MA, USA.*
3. *Department of Medicine, Weill Cornell Medical College, New York, NY, USA.*
4. *Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, FRA.*
5. *Centre for Eye Research Australia, University of Melbourne, Royal Victorian Eye and Ear Hospital, East Melbourne, Victoria, AUS.*
6. *Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA.*
7. *Department of Psychiatry, Harvard Medical School, Boston, MA, USA.*
8. *Department of Otolaryngology-Head & Neck Surgery, Harvard Medical School, Boston, MA, USA.*
9. *Department of Otolaryngology-Head & Neck Surgery, Mass General Brigham, Boston, MA, USA.*
10. *TRANSFORM HF, Department of Cardiology, University Health Network, Toronto, ON, CAN.*
11. *Peter Munk Cardiac Centre, University Health Network, Toronto, ON, CAN.*
12. *Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, CAN.*
13. *Division of Pulmonary, Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA.*
14. *Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA.*

‡*Corresponding Author*

Corresponding Author:
Leo Anthony Celi, MD, MPH, MSc
Laboratory for Computational Physiology
Massachusetts Institute of Technology
Cambridge, MA 02139
Email: lceli@mit.edu

ORCiD IDs:
Milit S Patel: https://orcid.org/0009-0000-9179-267X
David Scales: https://orcid.org/0000-0001-5727-7148
Janan Arslan: https://orcid.org/0000-0003-2683-3775
Paola Pedrelli: https://orcid.org/0000-0002-2865-2583
Matthew G Crowson: https://orcid.org/0000-0001-9950-0985
Rawan Abulibdeh: https://orcid.org/0009-0008-1179-4201
Leo Anthony Celi: https://orcid.org/0000-0001-6712-6626

Abstract: Hubris in contemporary medicine is less an individual failing than a structural condition produced by professional privilege, institutional incentives, and global asymmetries of power. This Perspective argues that modern health systems mistake technical competence for epistemic

completeness, fueling iatrogenic harm and algorithmic inequity. We propose disciplined humility and structural reflexivity as operational principles, translated into concrete practices across clinical, global health, and AI contexts.

In ancient Greek tragedy, hubris precipitated downfall not through malice, but through an inflated sense of capability that blinded protagonists to their limitations[1]. Consider Daedalus, whose ingenuity in crafting wings of wax embodied both technical brilliance and fatal overreach; the tool that promised liberation became the instrument of catastrophe when its limits were ignored. Today, analogous overestimation of competence permeates healthcare systems, technological innovation, and global health initiatives, creating ripple effects that disproportionately harm the most vulnerable[2,3].

Hubris manifests not as deliberate harm, but as an inability to acknowledge what we do not, and cannot, know[4]. In leadership theory, hubris can be distinguished from simple overconfidence as a complex combination of over-ambition and arrogance, typically catalyzed by the possession of significant, unfettered power[5]. It manifests as an 'untenable faith' in one's own abilities, often leading to a contemptuous disregard for advice or criticism. In contemporary health systems, this is less a personal flaw than a structural condition produced by professional privilege, institutional incentives, and global asymmetries of power.

It is essential to distinguish between individual hubris—which can be addressed through personal 'deliberate practice' and humility—and institutional hubris. In the medical system, power is structurally 'baked in' through legal authorities like physician licensure. Institutional hubris occurs when organizational cultures, such as the 'move fast and break things' ethos of tech, incentivize innovation over reflection, effectively silencing those who might highlight uncertainty.

Throughout this Perspective, we distinguish among several related but distinct expressions of this phenomenon. Overconfidence refers to miscalibrated certainty about specific judgments; arrogance denotes a dispositional dismissiveness toward others' knowledge. Hubris, by contrast, is the systemic pattern in which institutional structures, professional incentives, and power asymmetries together produce

sustained overestimation of epistemic reach. We use "epistemic hubris" to describe overconfidence in the completeness of current knowledge, "technological hubris" to describe unwarranted faith that technical sophistication compensates for contextual gaps, and "institutional hubris" to describe organizational cultures that structurally suppress uncertainty. These categories are not mutually exclusive; they interlock across the three domains we examine below: clinical practice, global health, and artificial intelligence.

Structural and individual hubris are interconnected and mutually reinforcing: institutional cultures that reward certainty provide the cognitive scaffolding for the psychological manifestation of hubris, often termed the "God complex," in which excessive pride blinds practitioners to the limits of their competence[6]. Research suggests hubris may manifest as an acquired disorder known as 'Hubris Syndrome,' characterized by a behavioral transformation following the attainment of power[6]. Unlike stable personality disorders that emerge in youth, this syndrome is an 'illness of position,' recognizable through specific symptoms such as a messianic manner of speaking and a loss of contact with reality[7]. This refusal to acknowledge uncertainty, while far from universal, recurs across many domains of medicine. It appears in the rigid hierarchies of the operating room and in the 'technological hubris' of AI, where algorithmic outputs are mistaken for objective truths immune to bias or epistemic limits. In both cases, the result is the same: catastrophic failure.

These examples are instructive not merely as instances of medical error, which may arise from many causes, but specifically as cases in which institutional cultures actively suppressed corrective feedback. In the Bristol heart scandal, pediatric cardiac mortality exceeded national benchmarks for years, yet whistleblower reports were disregarded and institutional prestige shielded practitioners from scrutiny[8,9]. The "oculostenotic reflex" in cardiology, a term Topol and Nissen coined to warn that the impulse to intervene can override clinical evidence, illustrates how procedural confidence can become self-reinforcing even when data suggest restraint[10]. The failure of IBM Watson, whose executives publicly promoted AI-driven cancer care while internal documents revealed frequently inaccurate recommendations, demonstrates the same dynamic in a technological register: organizational investment in a narrative of competence overrode honest appraisal of limitations[11]. In each case, the defining feature was not ignorance per se, but the structural refusal to reckon with available evidence of failure, the hallmark of hubris as we define it here.

When these dynamics enter the domain of artificial intelligence, the consequences become tangible at scale. Dermatology algorithms trained predominantly on light-skinned populations have demonstrated significantly lower sensitivity for melanoma detection in darker skin tones[12]. Sepsis prediction models validated in single-center academic settings have shown degraded performance when deployed in community hospitals with different patient populations and documentation practices[3,13]. In each instance, the failure was not purely technical but reflected an unexamined assumption that training conditions would generalize universally, an assumption that institutional confidence left unchallenged.

The same hubris appears in reverse: resistance to AI adoption can reflect institutional defensiveness, as established power structures reject tools not on evidence of harm, but to preserve existing hierarchies. These failures do not arise from isolated misjudgments, but from entrenched patterns that shape how expertise is produced, trusted, and exported.

Having established hubris as a structural condition rather than individual failing, we now examine how it manifests across three interlocking domains: clinical practice, global health interventions, and algorithmic systems. Each arena reveals distinct expressions of the same underlying pattern—the mistaking of technical competence for epistemic completeness.

The Myth of Context-Free Control

In surgical morbidity and mortality conferences, cases have traditionally been dissected for individual error, as though execution alone determines outcomes[14]. To be clear, the act of taking individual responsibility in such conferences reflects professional integrity, and these forums have evolved considerably in many institutions toward systems-oriented quality improvement[14]. Yet to the extent that M&M culture historically overemphasized individual mastery while underemphasizing systemic contributors to adverse events, it instantiated a particular form of hubris: the belief that human control extends further than it does. The surgeon who believes they alone determine patient destiny commits the same error as the data scientist who assumes their model captures all relevant variables: both risk mistaking technical competence for epistemic completeness. When the moral and epistemic scope of medicine expands beyond its traditionally narrow goals, as Hofmann argues[15], assumptions forged in one context may prove insufficient in another, and the failure to anticipate this gap constitutes a form of hubris.

This overconfidence manifests differently across contexts but shares common roots. In technology sectors, developers frequently deploy artificial intelligence tools without adequate consideration of how socioeconomic, cultural, and infrastructural differences will alter performance[3,16]. A diagnostic algorithm achieving high accuracy in Boston may perform less well in rural Ghana—not because the technology is inherently flawed, but because its design presumes universality where none exists[12,17]. When context is treated as noise rather than the ground on which care is delivered, errors are not anomalies, but the expected result[18].

The medical profession's historical privilege creates conditions in which hubris can flourish, though it is neither universal nor inevitable[19]. Elevated social status can widen the distance between perceived and actual capability[1]. In elite academic institutions and major technology companies, success may breed confidence that, in some cases, metastasizes into the conviction that good intentions guarantee beneficial outcomes[2,20]. Such environments can also be sites of rigorous self-critique and epistemic humility; the risk we identify is not that prestige necessarily produces hubris, but that it creates structural conditions, including cultural rewards for confidence and insufficient mechanisms for surfacing doubt, in which hubris can take root more readily.

The Geography of Overconfidence

These patterns of overconfidence extend beyond institutional walls to shape global health interventions, where geographic and economic asymmetries amplify the consequences of epistemic overreach. Global health has made significant strides toward recognizing these dynamics, yet legacy patterns of epistemic asymmetry persist. The risk of hubris emerges when scientific proficiency is conflated with situational omniscience: when a policy drafted in Geneva, Washington, D.C., or London, based on aggregate data and economic modeling, is treated as inherently superior to local knowledge without mechanisms for integrating the two[2,20]. This is not to suggest that all global health actors hold such beliefs; many have moved well beyond colonial-era paternalism. Rather, the structural incentive architecture of global health funding, publication, and agenda-setting can reproduce these asymmetries even when individual practitioners are well-intentioned. Similarly, cost-effectiveness analysis (CEA) is a necessary budgetary tool, but when it is deployed as the sole arbiter of resource allocation without attention to local values and priorities, it risks functioning as a proxy for epistemic hierarchies that determine whose suffering "counts."

The pattern repeats in the deployment of artificial intelligence across low- and middle-income countries, where solutions designed for high-resource settings arrive with assumptions about digital infrastructure, literacy, and healthcare system capacity that may not hold[3]. When algorithms require stable internet connectivity in regions with intermittent power, or assume patients have private spaces for telehealth

consultations in multigenerational households, the problem may initially reflect a knowledge deficit rather than hubris. It becomes hubris when such assumptions persist in the face of available evidence to the contrary, or when deployment proceeds without mechanisms to surface and respond to contextual mismatches[16,21]. Yet, hubris flows in multiple directions. High-income countries assuming their regulatory frameworks for AI ethics will transfer seamlessly to different legal and cultural contexts demonstrate a similar blindness[16]. These dynamics intersect with broader debates about evidence hierarchies: when "evidence-based medicine" is narrowly equated with randomized controlled trials regardless of contextual feasibility, the resulting framework can inadvertently dismiss forms of knowledge that do not conform to a single methodological standard[2]. The result is a global epistemic gradient in which some actors define rigor, while others are expected to conform to it, regardless of context or consequences. This gradient ensures that the hubris of high-income settings is not merely an export, but a foundational logic that redefines what constitutes valid evidence globally.

Problems with Hubris and AI: The Algorithmic Amplification

Among the mechanisms that propagate this geographic gradient, artificial intelligence warrants particular scrutiny—not as a separate problem, but as a technology that inherits and amplifies existing patterns of hubris at unprecedented scale. Artificial intelligence systems possess a unique capacity to scale the consequences of hubris[3,12]. Individual biases can cause significant harm, but algorithmic systems replicate those biases with a reach and velocity that far exceed any single human decision[13,22]. An AI model trained predominantly on data from light-skinned individuals fails to detect melanoma in darker skin tones at rates that would be unconscionable if openly stated as policy—yet the outcome is identical[12]. Algorithmic decisions lend a veneer of neutrality to patterns that would otherwise be recognized as inequitable.

The hubris here is multilayered: first, in assuming that training data represent universal human biology rather than historically privileged populations; second, in believing that technical sophistication compensates for representational gaps; and third, in deploying systems without adequate validation across the populations they will serve[16,23]. Each assumption reflects a failure to interrogate the limits of what has been built[16]. They create an epistemic blind spot: the conviction that computational refinement outweighs the biases embedded in the data.

Certain patterns within academic research can contribute a distinct strain of algorithmic hubris. Studies on AI in healthcare, published overwhelmingly in high-income countries and using data from high-income populations, nonetheless claim generalizability without cross-cultural validation[12]. Implicit in these claims is a belief that generalizability is a property earned through model performance rather than through alignment with diverse lived realities[3,17]. What appears as “rigor” in controlled research environments often dissolves when confronted with the variability of real clinical practice.

Ultimately, this manifests as two distinct but related failures: *epistemic hubris*, an overconfidence in the completeness of current knowledge[24], and *deployment hubris*, an overconfidence in the safety of releasing unvalidated tools into complex environments.

Disciplines of Deflection

Different fields can cultivate distinct expressions of hubris, though these tendencies are neither universal nor uncontested within any discipline. In academic medicine, publication pressure can incentivize novel claims over rigorous replication, contributing to a literature that may overstate certainty[15]. In technology entrepreneurship, the ethos of rapid iteration, while valuable in some contexts, carries particular risks in healthcare where failures have direct human consequences[21]. Some global health initiatives have been critiqued for measuring success through intervention deployment rather than sustained community-

determined outcomes[2,20]. In each case, the concern is not that these orientations are uniformly harmful, but that structural incentives can cause performance metrics to eclipse context and confidence to substitute for accountability.

The medical education system may further entrench these tendencies, though the picture is not monolithic. Hidden curricula can teach trainees to project confidence regardless of their actual knowledge, and research suggests that tolerance for ambiguity is unevenly cultivated across training programs[19,25]. At the same time, growing recognition of these dynamics has prompted educational interventions that explicitly prepare trainees for clinical uncertainty[26]. The risk is not that medical education uniformly punishes uncertainty, but that, absent deliberate structural countermeasures, the default incentive structure may normalize intellectual overreach and discourage honest appraisal of limits. The same pattern appears in grant writing, where investigators must promise definitive results; in corporate projections, which demand linear growth; and in policy briefs, which must appear unequivocal to remain politically viable[1415]. Complexity becomes a liability, and hubris becomes an adaptive strategy.

Across disciplines, a recurrent risk is what might be termed selective epistemic closure: defining rigor in ways that can inadvertently exclude challenges to foundational assumptions[2]. Clinical trials that exclude patients with multiple comorbidities generate internally valid data but may produce guidelines whose applicability to the broader patient population requires careful, context-specific judgment[15]. AI researchers who optimize for accuracy on benchmark datasets create models that perform poorly in clinical practice[13]. Global health programs that define success by donor metrics rather than community priorities build unsustainable interventions[20]. In each case, the boundaries of legitimate evidence are drawn narrowly, ensuring that what falls outside those boundaries is dismissed as irrelevant rather than as a signal of design failure.

Reflexivity as Resistance

The diagnostic phase complete, we turn to intervention—recognizing that identifying hubris can be insufficient without viable alternatives. The antidote to hubris is not self-flagellation, but structured reflexivity—a disciplined practice of examining how positionality, privilege, and institutional context shape what we take to be true[20]. For individual practitioners, this means cultivating awareness of how training, institutional culture, and personal circumstances shape perception[19]. For institutions, it requires creating systems where acknowledging uncertainty is rewarded rather than penalized[14]. Humility becomes a collective competency, not an individual trait.

To mitigate institutional hubris, healthcare can look to High-Reliability Organizations (HROs), such as commercial aviation[5]. HROs have successfully reduced 'command-and-control' cultures in which a captain's positional power previously overrode the crew's expertise[5]. By implementing procedural safeguards—like the surgical checklists now used in many theaters—institutions can reduce the 'power distance' that typically prevents the communication of risks.

Concrete examples of such competencies are emerging. In medical education, simulation-based training that deliberately exposes trainees to diagnostic ambiguity has shown promise in building tolerance for uncertainty[26] Structured peer reflection and narrative medicine exercises can help practitioners recognize how training and institutional culture shape their perception of what constitutes adequate evidence[25]. In AI development, "model cards" and datasheets for datasets operationalize reflexivity by requiring developers to document intended use cases, known limitations, and demographic performance gaps before deployment[13]. In global health, participatory action research frameworks that position community members as co-investigators rather than subjects represent an institutional mechanism for redistributing epistemic authority[2,20].

In AI development, reflexivity might manifest as mandatory algorithmic audits across demographic groups before deployment, with transparent reporting of differential performance[13]. In global health, it could mean genuine power-sharing in agenda-setting, where communities define problems and external partners provide requested support rather than predetermined solutions[2,20]. In clinical practice, it involves recognizing that patient expertise about their own lives constitutes essential knowledge that no amount of medical training can replace[15]. In each arena, reflexivity functions as a counterweight to the inertia of established authority.

Dialogical reflexivity—reflection through engagement with those holding different perspectives and power positions—offers particular promise[20]. When researchers from high-income institutions collaborate with colleagues in low- and middle-income settings, whose analysis is centered? When algorithms are developed for diverse populations, who defines "fairness"? When clinical guidelines are written, whose values determine acceptable risk-benefit tradeoffs? Such questions reveal that neutrality is often an illusion; every design choice reflects a hierarchy of perspectives and priorities[16,20].

The practice of reflexivity does not require elaborate frameworks. Even simple questions - Who is not represented in the development process? What assumptions about context, capacity, or values remain unexamined? Under what conditions might we be wrong? - can create friction against the momentum of unchecked confidence[2,13]. Small acts of interrogation accumulate into structural resistance, making it harder for hubris to pass as competence.

Collective Humility

Reflexivity, however, cannot remain an individual practice. Hubris flourishes in isolation—individual certainty, disciplinary silos, institutional defensiveness[1]. Its remedy requires collective commitment to uncertainty-embracing practices[2]. This is not relativism that abandons standards, but precision about the boundaries of knowledge and the contexts in which claims hold and where they do not[20].

Medical education that integrates "anti-hubris competencies" alongside clinical skills could produce practitioners better equipped to navigate uncertainty[19]. Technology companies that reward teams for identifying limitations of their tools before deployment might build systems that are safer and more robust. Research funding that values rigorous characterization of where interventions fail as highly as demonstration of where they succeed would generate more honest evidence bases[2]. In each instance, humility becomes structurally encoded rather than left to individual discretion. What unites these proposals across clinical practice, global health, and AI is a shared mechanism: each converts the recognition of epistemic limits from an individual aspiration into an institutional requirement, ensuring that the structural conditions that reproduce hubris **(see Figure 1)** are met with structural countermeasures rather than individual good intentions alone.

The goal is not eliminating confidence—effective action requires conviction—but right-sizing it to match actual capability[1]. A surgeon's confidence in their technical skill serves patients; confidence that this skill alone determines outcomes becomes dangerous[14]. An algorithm's performance on validation data merits appropriate trust; assuming this translates universally without verification courts disaster[12,13]. Humility, in this sense, is a calibration: an ongoing adjustment of certainty to context, evidence, and impact.

Perhaps the most profound humility lies in accepting that intelligence—human or artificial—is always partial, situated, and dependent on relationships[16,20]. The most sophisticated diagnostic algorithm remains useless without healthcare workers who trust it, infrastructure to run it, and patients willing to act on its outputs[3,21]. The most elegant research design cannot compensate for questions that matter to researchers but not communities[2]. The most skilled clinician cannot heal without the patient's participation in their

own care[15]. These interdependencies are not weaknesses but constitutive features of systems that involve human lives **(Figure 1)**.

These principles require concrete form. For AI developers specifically, humility must become a discipline of attention—a set of repeatable questions that interrupt momentum and surface the limits of what a model can claim **(Box 1)**.

Box 1.

1. **What assumptions about people, context, and clinical practice have we built into this model, and which of them might be wrong?** This shifts focus from performance to epistemic boundaries.
2. **Who is missing – from our data, from our design process, from the imagined user population – and what harms follow from that absence?** Absence becomes a diagnostic signal.
3. **Under what plausible conditions might this system fail, and who will be affected first or most severely?** Failure is assumed, not exceptional.
4. **What mechanisms exist for noticing when the model begins to drift, degrade, or cause harm, and who, including affected communities and end users, is empowered to intervene?** Accountability becomes operational, not symbolic.
5. **How do we clearly communicate the limits of generalizability so clinicians and patients can understand where not to trust the system?** Limitations are treated as essential information, not as an obligatory paragraph.
6. **What would restraint look like here?** Humility sometimes requires choosing not to deploy.

Such questions establish the intellectual posture required before deployment. They convert humility from aspiration to practice, ensuring that acknowledging limits becomes as routine as optimizing performance.

The Practice of Not Knowing

We return, finally, to where we began: the question of how knowledge producers confront their own limitations. Every system we build, every intervention we deploy, every algorithm we train represents a choice about which uncertainties to acknowledge[2,13]. Hubris risks becoming the refusal of that acknowledgment—the insistence that technical sophistication, good intentions, or institutional prestige renders us exempt from limitations that constrain others and manifests as a "universalizing" tendency that treats local infrastructural and cultural frictions not as critical design constraints to be respected, but as temporary deviations from a standard model that need to be corrected[1,19].

The practitioners, researchers, and technologists reading this occupy positions of relative privilege in global knowledge production[20]. This creates both responsibility and risk: responsibility to use that position toward equity, risk of perpetuating the very hubris we critique[20]. Self-reflection is not sufficient; it must connect to action that redistributes power, resources, and legitimacy[20].

What would it mean to build AI systems that foregrounded their limitations as prominently as their capabilities? To conduct global health initiatives that prioritized community governance over external metrics? To practice medicine that treated uncertainty as valuable information rather than professional failure? These questions have no algorithmic solutions—they require human choices about values[16,19].

We address this Perspective to the overlapping communities of clinicians, AI developers, global health practitioners, and policymakers who shape how medical knowledge is produced, validated, and deployed.

The structural nature of hubris means that no single audience can address it alone; the interventions we propose are most effective when adopted in concert across these groups.

The ancient Greeks understood that hubris invited nemesis: not downfall per se, but just consequence, the natural reckoning that follows overreach[1]. Our modern nemesis arrives as health inequities amplified by biased algorithms, communities harmed by well-intentioned interventions, and patients endangered by overconfident care[3,21].

Acknowledgements
LAC is funded by the National Institute of Health through DS-I Africa U54 TW012043-01 and Bridge2AI OT2OD032701, the National Science Foundation through ITEST #2148451, a grant of the Boston-Korea Innovative Research Project (RS-2024-00403047) and a grant of the Korea Health Technology R&D Project (RS-2024-00439677) through the Korea Health Industry Development Institute (KHIDI) as funded by the Ministry of Health & Welfare, Republic of Korea. The funders had no role in the conceptualization, writing, or decision to submit this manuscript. No additional funding was received for this research.

Author Contributions
Leo Anthony Celi conceptualized the study and supervised the overall direction of the work. Milit S Patel drafted the initial manuscript. Milit S Patel, David Scales, Janan Arslan, Paola Pedrelli, Matthew G Crowson, Rawan Abulibdeh, and Leo Anthony Celi critically reviewed, revised, and edited the manuscript for important intellectual content. All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work in accordance with ICMJE authorship criteria.

Competing Interests
MGC is a Chief Medical Officer at Drive Health, a digital health company. The other authors declare no competing financial or non-financial interests.

References

1. Selten, J.-P. Consider the hubris syndrome for inclusion in our classification systems. *Psychol Med* **53**, 5889–5891 (2023).

2. George, A. S. *et al.* Hubris, humility and humanity: expanding evidence approaches for improving and sustaining community health programmes. *BMJ Glob Health* **3**, e000811 (2018).
3. Fletcher, R. R., Nakeshimana, A. & Olubeko, O. Addressing Fairness, Bias, and Appropriate Use of Artificial Intelligence and Machine Learning in Global Health. *Front Artif Intell* **3**, 561802 (2020).
4. Donnelly, M. Best Interests in the Mental Capacity Act: Time to say Goodbye? *Med Law Rev* **24**, 318–332 (2016).
5. Sadler-Smith, E., Akstinaite, V., Robinson, G. & Wray, T. Hubristic leadership: A review. *Leadership* **13**, 525–548 (2017).
6. Owen, D. Hubris syndrome. *Clin Med (Lond)* **8**, 428–432 (2008).
7. Owen, D. & Davidson, J. Hubris syndrome: An acquired personality disorder? A study of US Presidents and UK Prime Ministers over the last 100 years. *Brain* **132**, 1396–1406 (2009).
8. Dyer, C. Bristol inquiry. *BMJ* **323**, 181–181 (2001).
9. Keogh, B. *et al.* The legacy of Bristol: public disclosure of individual surgeons' results. *BMJ* **329**, 450–454 (2004).
10. Topol, E. J. & Nissen, S. E. Our Preoccupation With Coronary Luminology: The Dissociation Between Clinical and Angiographic Findings in Ischemic Heart Disease. *Circulation* **92**, 2333–2342 (1995).
11. Strickland, E. How IBM Watson Overpromised and Underdelivered on AI Health Care. *IEEE Spectrum* https://spectrum.ieee.org/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care (2019).
12. Zou, J. & Schiebinger, L. Ensuring that biomedical AI benefits diverse populations. *EBioMedicine* **67**, 103358 (2021).
13. Nazer, L. H. *et al.* Bias in artificial intelligence algorithms and recommendations for mitigation. *PLOS Digit Health* **2**, e0000278 (2023).
14. Humbyrd, C. J. Virtue Ethics in a Value-driven World: Medical Errors and the Albatross. *Clin Orthop Relat Res* **481**, 661–663 (2023).

15. Hofmann, B. Managing the moral expansion of medicine. *BMC Med Ethics* **23**, 97 (2022).
16. Kerasidou, A. Ethics of artificial intelligence in global health: Explainability, algorithmic bias and trust. *J Oral Biol Craniofac Res* **11**, 612–614 (2021).
17. Pasipamire, N. & Muroyiwa, A. Navigating algorithm bias in AI: ensuring fairness and trust in Africa. *Front Res Metr Anal* **9**, 1486600 (2024).
18. Greenhalgh, T. Will COVID-19 be evidence-based medicine's nemesis? *PLoS Med* **17**, e1003266 (2020).
19. Abbasi, K. Foundation of wisdom: "I don't know". *BMJ* l6053 (2019) doi:10.1136/bmj.l6053.
20. Liwanag, H. J. & Rhule, E. Dialogical reflexivity towards collective action to transform global health. *BMJ Glob Health* **6**, e006825 (2021).
21. d'Elia, A. *et al.* Artificial intelligence and health inequities in primary care: a systematic scoping review and framework. *Fam Med Community Health* **10**, e001670 (2022).
22. Chinta, S. V. *et al.* AI-driven healthcare: Fairness in AI healthcare: A survey. *PLOS Digit Health* **4**, e0000864 (2025).
23. Ferryman, K., Mackintosh, M. & Ghassemi, M. Considering Biased Data as Informative Artifacts in AI-Assisted Health Care. *N Engl J Med* **389**, 833–838 (2023).
24. Pongiglione, F. Epistemic Hubris. *Social Epistemology* **39**, 91–105 (2025).
25. Sarkis, S. & Raphael, C. Understanding uncertainty and ambiguity in medicine and medical education: a narrative review with implications for training. *Postgraduate Medical Journal* qgaf170 (2025) doi:10.1093/postmj/qgaf170.
26. Ilgen, J. S. & Dhaliwal, G. Educational Strategies to Prepare Trainees for Clinical Uncertainty. *N Engl J Med* **393**, 1624–1632 (2025).

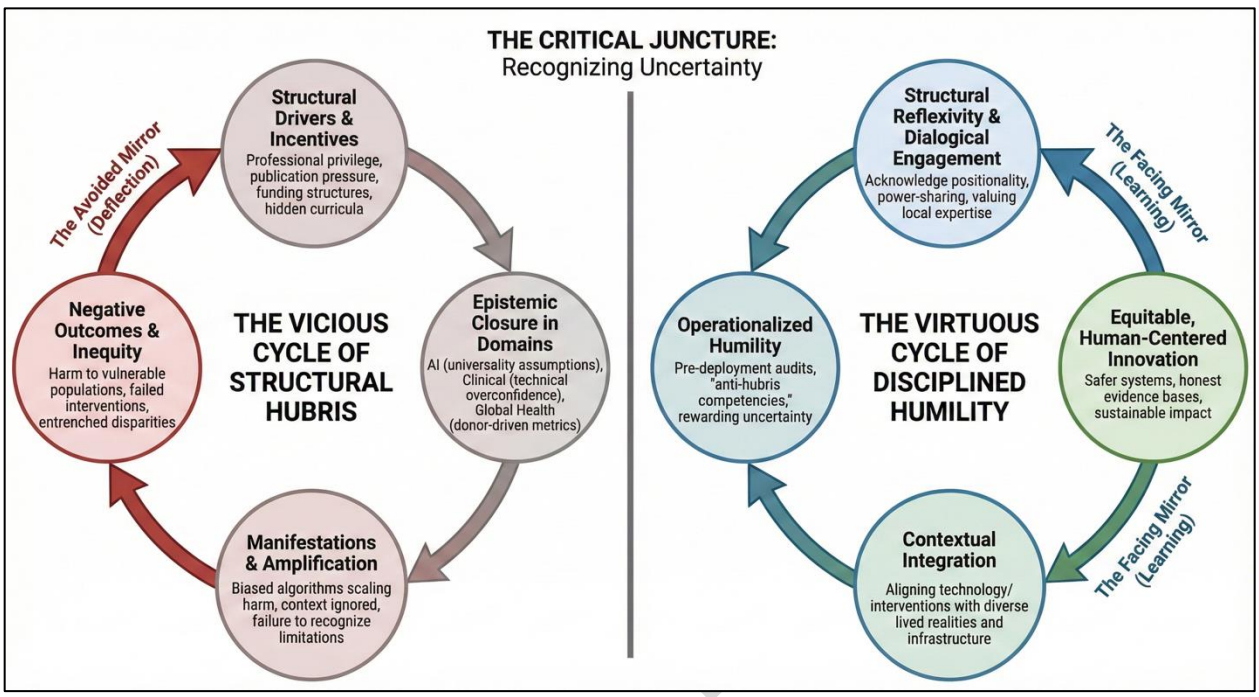

Figure 1. Divergent Pathways in Healthcare Innovation—From Structural Hubris to Disciplined Humility. The diagram contrasts the self-reinforcing mechanisms of the current paradigm (left) with the proposed framework for reflexive practice (right). The Vicious Cycle (red) illustrates how professional privilege and institutional incentives foster epistemic closure, leading to algorithmic and clinical failures that are deflected rather than examined. The Virtuous Cycle (blue/green) demonstrates how structural reflexivity interrupts this pattern by operationalizing humility and integrating context, converting potential failures into learning opportunities ("The Facing Mirror") and resulting in more equitable, robust systems.