Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Beyond Transhumanist Perfectionism: A Prohumanist Framework for AI, Digital Health, and Dignified Longevity

Israel Huerta Castillo

israel.huerta@rai.usc.es

University of Santiago de Compostela

open-forum

Keywords: prohumanism, transhumanist perfectionism, optimisation drift, digital health, AI governance, dignified longevity

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8457186/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

AI-enabled medicine does not express a single anthropology, and digital health is not intrinsically transhumanist. Some platformised health arrangements nevertheless stabilise optimisation drift: a transition in which bounded aims of support, prevention, or capability preservation are progressively recoded as an open-ended duty of measurable correction. The drift becomes structurally closer to transhumanist perfectionism than to bounded enhancement when three properties converge: principled stopping rules lose practical force, embodiment is governed as an indefinitely corrigible substrate, and selective interventions escalate into a recursively expanding package of optimisation. Platform incentives, preventive rationalisation, and organisational pressures explain the emergence and diffusion of this pattern more directly than transhumanist doctrine; structural affinity identifies the normative grammar enacted once those properties converge. A reconstructive prohumanist framework evaluates that grammar through dignified longevity, understood as technologically supported continuation of life that preserves embodied agency, plural goods, recognition, justice, and meaningful exit. Four constraints follow: proportionality, contestability, non-totalisation, and epistemic autonomy. Applied to remote patient monitoring and episodic diagnostic AI, they distinguish bounded support from optimisation-by-default across design, deployment, review, and withdrawal.

1. Introduction

Artificial intelligence and digital infrastructures are changing the temporal and normative organisation of health care. Clinical encounters increasingly sit within longer chains of data capture, risk prediction, remote monitoring, behavioural recommendation, and automated triage. These arrangements can improve access, detect deterioration earlier, reduce avoidable work, and support people who wish to remain at home. They also change what counts as a relevant health signal and how quickly a signal becomes a reason for intervention. Once bodies and routines are continuously rendered as scores, alerts, and trajectories, the possibility of improvement can acquire the force of a standing expectation (Lupton 2016; Ruckenstein and Schüll 2017; Sharon 2017).

The resulting problem is narrower than a general objection to medical AI. Diagnostic assistance, rehabilitation technologies, remote monitoring, and preventive tools differ in purpose, duration, institutional setting, and effect on agency. An episodic system that evaluates a retinal image for a specified referral decision does not organise a life in the same manner as an adaptive platform that gathers data across daily activities, repeatedly adjusts thresholds, and couples access to care with continued participation. A defensible analysis therefore requires criteria for distinguishing bounded technological support from arrangements in which intervention becomes recursively self-extending.

The term optimisation drift designates that institutional transition. A bounded aim of care - for example, detecting a defined complication, supporting recovery after discharge, or preserving a particular capability - is progressively re-specified as a continuing opportunity to measure and correct. Success does not close the intervention; it reveals another variable, domain, or threshold to be acted upon. The

relevant shift is visible when purposes expand, stopping rules weaken, and the burden of justification moves from those who extend intervention to those who seek to pause or refuse it. Datafication makes this movement possible, while organisational incentives and platform economics can make it attractive (van Dijck 2014; van Dijck et al. 2018; Zuboff 2019).

Optimisation drift is not equivalent to transhumanism. Transhumanist thought is internally heterogeneous, and health technologies need not arise from transhumanist commitments. The comparison becomes philosophically useful at a different level. Some institutional arrangements enact rules of improvement that are structurally closer to a perfectionist configuration within transhumanism than to selective human enhancement. The relevant relation is one of structural affinity: a correspondence between operational rules and a normative grammar, without claims of historical influence, ideological membership, or shared intention. Three properties make the comparison discriminating. Intervention lacks a credible stopping point; embodiment is treated as an indefinitely corrigible substrate; and selective improvements escalate towards an open package in which refusal appears increasingly anomalous.

Political economy and bureaucratic rationalisation remain the stronger causal explanation for why such arrangements spread. Health systems seek efficiency, earlier detection, population management, and workload reduction; firms seek data, market extension, and continued engagement. Those causes can produce optimisation drift without any transhumanist doctrine. The affinity claim addresses the normative orientation stabilised by the drift: whether correction remains subordinate to a specified human good or becomes the default direction of responsible conduct. The comparison has value only if it helps to distinguish cases and generates governance implications unavailable from a generic critique of commercialisation.

A positive normative account is required because rejecting optimisation drift does not entail rejecting enhancement. Prohumanism organises resources from embodied cognition, the capability approach, vulnerability ethics, epistemic justice, contextual integrity, and relational accounts of autonomy around a specific question: how can technological improvement remain answerable to the embodied and social conditions that make improvement worthwhile? Its evaluative centre is dignified longevity. Longer or technologically supported lives are valuable insofar as people can remain capable, recognised, situated, and practically authoritative in relation to their own trajectories. This standard does not supply a universal metric of a good life. It identifies conditions that optimisation must not erode while pursuing health-related benefits.

Four governance constraints translate that standard into institutional requirements. Proportionality governs the relation between patient-relevant benefit and the scope, burden, reversibility, and duration of intervention. Contestability requires systems and organisations to remain responsive to reasoned challenge throughout their lifecycle. Non-totalisation protects boundaries around domains, times, and purposes that should not automatically become objects of health governance. Epistemic autonomy preserves the standing of patients and professionals to interpret, question, and negotiate the

significance of algorithmic outputs. Together, these constraints differentiate bounded enhancement from optimisation-by-default without treating legal compliance, technical performance, or individual consent as sufficient on their own.

2. From digital health to optimisation drift

2.1. Platformised health and recursive evaluation

Digital health is often described through a catalogue of devices: applications, wearables, telemedicine systems, clinical decision support, remote sensors, and predictive models. The catalogue obscures the organisational change produced when those devices are connected through platforms. Platformisation links measurement to prediction, prediction to recommendation, recommendation to behavioural response, and response to further measurement. Clinical attention can consequently extend beyond episodic encounters as daily activity is converted into an environment of continuously renewable signals (Ruckenstein and Schüll 2017; van Dijck et al. 2018).

This environment has genuine clinical uses. Continuous or frequent data can identify patterns missed during a consultation, support self-management, and help allocate scarce attention. The ethical issue begins with the relation between the signal and the action it authorises. A threshold can be a provisional aid whose meaning is negotiated within a clinical context. It can also become an institutional trigger that shifts the burden of proof. Once an alert, risk category, or adherence score enters workflow, departing from it may require more explanation than following it. Kempt et al. (2023) show how disagreement with a diagnostic decision-support system can place a new justificatory burden on the physician even when neither human nor machine is straightforwardly at fault. The same dynamic can reach patients when absence of data, refusal of a recommendation, or deviation from a target is interpreted as non-adherence.

Datafication further changes the ontology of the problem presented to care. Variables that are measurable and comparable are easier to incorporate into targets, dashboards, and procurement criteria. Their availability can make them appear more actionable than dimensions that remain narrative, relational, or context-dependent. Obermeyer et al. (2019) demonstrated the practical consequences of a proxy choice when an algorithm used health costs as a surrogate for health need and thereby reproduced racial inequality. The calculation operated on a target that encoded a socially patterned relation between expenditure and illness. Optimisation begins before the model produces an output, in the selection of what counts as the object of improvement.

The quantified-self literature shows how this logic can move from institutions into subjectivity. Self-tracking may support reflection and agency, while metrics can establish normative expectations about discipline, productivity, and responsible self-management (Lupton 2016; Sharon 2017). The empowerment-surveillance binary leaves this ambiguity unresolved. The decisive questions concern who defines the categories, what consequences follow from them, how long they remain active, and whether

affected persons can change their practical meaning. Autonomy is weakened when participation consists mainly in complying with categories and thresholds that remain externally fixed.

Platform economics and administrative pressures intensify these tendencies. Data-rich infrastructures benefit from scale, interoperability, retention, and repeated use. Health organisations face incentives to reduce admissions, standardise decisions, demonstrate measurable outcomes, and direct attention towards cases classified as most urgent. Firms may benefit from adding modules, integrating new domains, and maintaining engagement. These incentives are sufficient to explain why a circumscribed intervention can expand. The language of personalisation may coexist with a standardising infrastructure in which the same kinds of signals, targets, and responses are reproduced across populations (van Dijck et al. 2018; Zuboff 2019).

Optimisation drift occurs when these mechanisms change the telos of intervention. Three indicators are especially important. First, the original purpose broadens: post-discharge monitoring becomes chronic surveillance, or a single risk score becomes a gateway to multiple services. Second, actionability expands: additional deviations are treated as reasons for correction even when their connection to a patient-relevant benefit is weak. Third, stopping becomes exceptional: the system continues by default, while withdrawal requires explanation or entails loss of support. No single indicator proves a perfectionist orientation. Their combination reveals a transition from using measurement to pursue a bounded good towards treating measurable difference as a standing invitation to intervene.

2.2. Remote patient monitoring as a situated test

AI-enabled remote patient monitoring provides a concrete setting in which benefits and drift can be examined together. Behdadi et al. (2026) analyse MonitAir, a system developed for chronic obstructive pulmonary disease care in Sweden. The intended benefits include earlier detection of exacerbations, support for self-management, and reduction of clinical workload. Their multidimensional evaluation also identifies risks involving data bias, intelligibility, uneven access, responsibility, role shifts, and patient burden. Its analytical value lies in tracing ethical significance across the patient, the patient-clinician relationship, the organisation, and the health sector without reducing the technology to a promise-or-harm binary.

MonitAir supplies the documented starting point for the analysis. The later sequence in which monitoring acquires additional modules, cross-domain scoring, or an indefinite duration is an analytical composite used to test how a bounded arrangement could drift. It is not presented as a feature of MonitAir or as a description of any particular institution.

A bounded deployment of remote monitoring can remain oriented towards a defined clinical purpose. Data collection may be limited to variables necessary for detecting deterioration; monitoring may be reviewed at specified intervals; the clinical team may contextualise alerts; and participation may be discontinued without punitive consequences. Under those conditions, technical mediation can support

capability and enable a person to remain at home without converting the rest of life into a health optimisation programme.

The same infrastructure can move in another direction. Monitoring designed for exacerbation detection may acquire modules for activity, sleep, adherence, nutrition, or general risk. Scores initially used to trigger clinical attention may influence eligibility, priority, or resource allocation. Data absence may become evidence of disengagement. As the platform accumulates functions, the initial stopping point becomes difficult to recover because each new stream offers another prospective benefit. Expansion can be motivated by workload, prevention, research, reimbursement, or commercial strategy. None of these motives is transhumanist in itself. Their convergence can nevertheless establish a regime in which more monitoring and correction are presumed beneficial until someone demonstrates otherwise.

The shift also redistributes epistemic authority. Remote data may reveal clinically important changes, while the dashboard can marginalise information that is difficult to standardise. Fatigue related to grief, reluctance shaped by domestic conditions, or a patient's account of what constitutes a tolerable daily routine may enter as unstructured context beside highly visible numerical indicators. Van Leersum and Peine (2026) show that explainability needs in long-term care vary across technical, managerial, and care settings; practical understanding and situated knowledge cannot be replaced by a single technical account of the algorithm. Maccaro (2026) similarly treats healthcare algorithms as infrastructures that stabilise assumptions about care, authority, and social worth, while warning that context-insensitive transfer can produce epistemic displacement and dependency.

These observations locate optimisation drift at the intersection of infrastructure and interpretation. The relevant question concerns whether the institution preserves conditions under which prediction remains one contribution to care instead of becoming the authoritative representation of the person. A system may be technically explainable and still organise decisions around a proxy that affected persons cannot meaningfully challenge. Conversely, a system whose internal model is difficult to interpret may remain governable if its purpose is narrow, its consequences are limited, and its outputs are subject to competent review. Explanation is one component of governance; justification and contestability remain independently necessary (Babic et al. 2026).

2.3. The normative significance of drift

Optimisation drift differs from ordinary quality improvement because the object and endpoint of improvement are unstable. A hospital can seek to reduce avoidable readmissions while maintaining a defined population, timeframe, evidential threshold, and review procedure. Drift begins when that aim authorises a continuing enlargement of what is measured and corrected. The institution no longer asks only whether a particular intervention advances a specified good; it treats the availability of another signal as presumptive reason to extend intervention.

The resulting normativity is practical rather than doctrinal. Repeated classifications shape what professionals notice, what patients are expected to report, and what organisations regard as defensible

action. An alert creates a reason to act; an accumulating record creates expectations of compliance; a risk score renders future intervention administratively legible. These habits can stabilise even when no participant endorses an explicit philosophy of perfection. Digital-health systems then participate in defining responsible conduct by making correction the easiest institutional response to detectable variation.

A causal account centred on markets and administration explains much of this development. It identifies who benefits from data accumulation, why organisations prefer standardised categories, and how preventive rationalities extend action upstream. The evaluative question remains because two arrangements may emerge from similar incentives while enacting different rules of improvement. One remains answerable to a bounded purpose and treats withdrawal as legitimate; the other converts technical possibility into a continuing mandate. Structural affinity discriminates between those normative configurations while leaving political economy's causal role intact.

3. Structural affinity without causal reduction

3.1. Transhumanism, human enhancement, and a delimited ideal type

Transhumanism comprises a heterogeneous family of positions concerned with using science and technology to overcome limitations and expand human capacities. Its historical and philosophical range includes commitments to autonomy, welfare, longevity, morphological freedom, risk reduction, and possible posthuman futures (Bostrom 2005). Treating this family as a unified doctrine would erase disagreements about the value of embodiment, the pace and scope of enhancement, distributive institutions, and acceptable limits.

Human enhancement is equally broader than a single ideology. Selective interventions can seek to repair function, prevent harm, increase opportunity, or extend capabilities while remaining within a recognisably human range. Agar's defence of limits rejects the inference that every technically available increase is desirable and distinguishes moderate from radical transformation (Agar 2013). García-Barranquero and Alonso (2025) sharpen the distinction through two axes: whether improvement is gradual or directed towards exponential transformation, and whether the biological body remains ethically significant or is treated mainly as a constraint to be surpassed. These axes permit criticism of maximalist trajectories without construing every enhancement as a step towards the posthuman.

Recent debate also complicates a simple moderate–radical division. Brennan’s “weak transhumanism” retains radical enhancement as an ultimate aim while proposing an incremental route through repeated moderate interventions. Lyreskog and McKeown argue that some transformative enhancements may create epistemic gaps that undermine fully informed rational choice. Together, these positions separate technological magnitude, temporal sequence, and the agent’s capacity to anticipate the evaluative consequences of enhancement. They also sharpen the present ideal type. Perfectionist affinity does not

follow from radicality alone, nor is it absent whenever each step appears moderate. A sequence of individually moderate interventions may acquire a perfectionist structure when each stage authorises the next, stopping rules lose force, and refusal becomes progressively costly. Conversely, a radical intervention can remain outside the ideal type when its purpose, endpoint, and conditions of refusal are bounded. The comparison therefore tracks the governance structure of improvement rather than a simple scale of technological intensity (Brennan 2023; Lyreskog and McKeown 2022).

Transhumanist perfectionism names an ideal type within this diverse terrain. It combines three tendencies: improvement is oriented towards an open horizon rather than a specified sufficiency; inherited embodiment has weak authority as a source of limits; and selective refusal becomes difficult to sustain once enhancement is understood as an integrated project. Hauskeller's analysis is relevant because it reconstructs transhumanist arguments through narratives of human self-design, immortality, the elimination of suffering, and expanded autonomy, showing how technological possibilities acquire direction from broader accounts of human nature (Hauskeller 2016). The ideal type is legitimate only as an analytical device. It does not classify every transhumanist, nor does it attribute motives to developers or clinicians.

3.2. The strongest causal rival

Platform incentives, bureaucratic rationalisation, preventive medicine, and population management explain optimisation drift more directly than transhumanist doctrine. Platforms benefit from data continuity and functional expansion. Health systems seek earlier detection, standardised workflows, and allocation of attention under scarcity. Preventive strategies move the temporal boundary of intervention from manifest illness towards risk, predisposition, and behaviour. These forces can produce continuous monitoring and behavioural steering without any commitment to transcending the human condition.

This rival explanation has priority. It accounts for adoption, diffusion, and institutional durability. It also identifies power asymmetries that a purely philosophical comparison could obscure. Sharon's analysis of digital health and capitalism shows how commercial platforms can acquire roles in health policy and reshape common goods, while the platform-society framework explains how infrastructures organise public values through connective ecosystems rather than isolated devices (Sharon 2018; van Dijck et al. 2018). Maccaro's comparative account adds that algorithmic governance is embedded in unequal infrastructures and may reproduce dependency, digital paternalism, and locally inappropriate forms of authority (Maccaro 2026).

The affinity thesis survives only by accepting this causal division of labour. Its task begins after adoption and expansion: it identifies the grammar of improvement stabilised by those processes. Political economy explains why expansion is attractive; structural affinity asks what makes that expansion appear normatively self-justifying. The comparison adds value only when it reveals a shift from bounded support to open-ended corrigibility that efficiency or profit alone does not classify.

3.3. Three conditions of structural affinity

Structural affinity is a correspondence between operational rules and a normative orientation. It requires more than shared vocabulary or a general desire for better health. Three properties must converge.

A first condition concerns the weakening or displacement of stopping rules. A stopping rule identifies when a purpose has been achieved, when further intervention requires renewed justification, or when withdrawal must be available. Long duration alone is not decisive; chronic care may require continuing support. The difference lies in the allocation of justificatory burden. A bounded arrangement periodically asks whether continued monitoring remains necessary and proportionate. A drifting arrangement presumes continuation and treats interruption as a deviation. Once every successful intervention supplies data for another potential intervention, improvement loses a stable point of sufficiency.

Embodiment supplies the second condition when it is governed as indefinitely corrigible. Medicine necessarily treats many bodily states as open to change. The perfectionist tendency appears when technical corrigibility becomes a general normative presumption: measurable variation is interpreted primarily as shortfall, and the body matters chiefly as material for further adjustment. Embodiment then ceases to function as the lived condition through which agency, dependency, pleasure, fatigue, identity, and care acquire meaning. A person's refusal to optimise a variable can look irrational because the system recognises the possibility of correction more readily than the reasons for leaving a condition, routine, or risk unaltered.

Package-style escalation completes the configuration. Selective enhancement permits differentiated judgements: one intervention may be accepted while another is refused. A package orientation links domains so that each improvement increases the apparent need for complementary monitoring or adjustment. Remote monitoring for exacerbation risk expands towards activity, sleep, nutrition, adherence, and general behavioural prediction; a score designed for clinical attention becomes relevant to resource allocation or eligibility. Integration can be clinically justified, and multidimensional care is not inherently perfectionist. Escalation becomes significant when added functions inherit legitimacy from the original purpose without independent justification and when selective refusal carries accumulating costs.

Taken separately, the properties do not sustain the comparison. Weak stopping rules may reflect poor administration; treating a particular pathology as corrigible is ordinary medicine; coordinated interventions may be necessary for complex illness. The comparison begins only when these features interact so that intervention remains open-ended, the body is continually actionable, and improvement advances as an expanding package. Under those conditions, optimisation drift is structurally closer to transhumanist perfectionism than to bounded enhancement because both privilege continuing transformation over context-specific sufficiency.

Their interaction changes the normative role of each property. Weak stopping rules can produce administrative overreach while the body still retains authority as a source of practical limits and individual modules remain independently rejectable. Treating a pathology as corrigible remains ordinary medicine when the intervention has a defined endpoint. Complex, multimodal care also stays bounded

when each component has its own justification and refusal does not jeopardise access to the rest of care. Affinity emerges when weakened stopping rules make functional expansion presumptive, expansion renders further bodily variation actionable, and indefinite corrigibility turns that actionability into a continuing expectation of correction. The resulting feedback relation authorises additional modules, widens the domain of actionable variation, and makes sufficiency increasingly difficult to defend. Political economy explains why organisations seek more data, functions, and engagement; the affinity diagnosis identifies why the resulting arrangement treats further transformation as the normal direction of responsible conduct. Where this feedback relation is absent, commercial or bureaucratic overreach may remain, but the perfectionist comparison should be withdrawn.

Autonomous diabetic-retinopathy screening clarifies the boundary. The system evaluated by Abràmoff et al. (2018) performs a defined diagnostic task on retinal images and produces a referral outcome within an episodic encounter. The contrastive configuration considered here can remain bounded by a specified target, data type, clinical consequence, and endpoint, while still requiring validation, equitable access, quality control, and routes for handling error. Its status is conditional rather than representative: a deployment that reuses images across purposes, links screening to broader scoring, or lacks workable remedy would require a different evaluation. Autonomous AI alone does not determine the affinity; the operational organisation of improvement does.

3.4. Defeat conditions

Structural affinity fails when the three properties do not discriminate between bounded and open-ended arrangements. It also fails if the comparison adds no evaluative consequence beyond an account of commercial expansion or administrative overreach. Evidence that a platform maintains enforceable purpose limits, periodic necessity review, non-punitive withdrawal, and independent justification for each added function would weaken the diagnosis even if monitoring remained extensive. A health system that treats patient narratives and professional disagreement as reasons capable of changing outputs and consequences would likewise resist the corrigible-substrate orientation.

As a conceptual and comparative claim, structural affinity does not establish the prevalence of perfectionist configurations. Empirical work would be required to determine how often purposes expand, how stopping rules operate in practice, and how patients experience refusal or exit. Its immediate value lies in identifying what a positive normative framework must protect. Bounded improvement requires governance to specify the goods being served, retain limits that technical possibility cannot erase, and preserve the authority of affected persons within the interpretation of their own lives.

4. Prohumanism and dignified longevity

4.1. Embodied agency and constitutive vulnerability

Health technologies act upon persons whose agency is embodied, relational, and institutionally mediated. Bodies are lived conditions through which perception, effort, dependence, pleasure, pain, and

practical identity are organised, while remaining open to alteration and support. Embodied and enactive approaches describe agency as situated sense-making (Varela et al. 1991; Thompson 2007). This account gives the body normative relevance without treating every inherited condition as desirable.

Vulnerability follows from exposure, dependency, and finite capacities, yet it should not be romanticised. Disease, avoidable suffering, and exclusion can rightly motivate treatment and technological support. The normative point is that vulnerability cannot be exhausted by a catalogue of technical deficits. It also describes relations of reliance, unequal power, and susceptibility to institutional decisions (Goodin 1985). A system that reduces vulnerability to correctable risk may overlook the conditions under which assistance becomes burdensome, withdrawal becomes costly, or a person's account loses authority.

The remote-monitoring case illustrates the ambivalence. Earlier detection of COPD exacerbations can protect capability and support ageing at home. The same arrangement may transfer work to patients, require sustained digital competence, or make access dependent on compliance with monitoring. Behdadi et al. (2026) identify uneven access, role shifts, responsibility gaps, and patient burden alongside potential benefit. These conditions determine whether the technology supports agency or reorganises dependency in ways that narrow it.

4.2. Prohumanism as a reconstructive framework

Prohumanism is a prudential framework for evaluating technological improvement under conditions of optimisation drift. Its ingredients are not new in isolation. The capability approach asks what people are effectively able to do and be, rather than treating resources or biological performance as sufficient measures of advantage (Nussbaum 2011). Relational accounts of autonomy show that choice depends on social conditions, recognitional standing, and available alternatives (Mackenzie and Stoljar 2000). Epistemic injustice identifies harms that arise when persons are discredited or lack the interpretive resources needed to make their experiences intelligible (Fricker 2007). Contextual integrity explains why information flows must be assessed against the norms and purposes of particular settings rather than through consent or secrecy alone (Nissenbaum 2010).

The framework's distinct function is to organise these resources around the boundary between bounded enhancement and open-ended optimisation. It operates as a second-order grammar, coordinating considerations that comprehensive moral theories may weigh differently. A technological intervention is prohumanly justified when its aims remain answerable to embodied agency and vulnerability, plural human goods, justice, and meaningful possibilities of refusal and exit. These considerations identify what must enter the justification and what aggregate performance cannot displace.

Prohumanism's claim to originality is correspondingly limited. The term would be redundant if the capability approach combined with relational autonomy, epistemic justice, and contextual integrity generated exactly the same distinctions and governance requirements. Its use is defensible insofar as the reconstruction makes a specific problem visible: health technologies can advance a capability while eroding the conditions under which that capability remains practically one's own. A remote-monitoring

system may reduce exacerbations and still weaken agency if it expands without limits, penalises withdrawal, or turns a proxy into the authoritative account of responsible conduct.

Its second-order function becomes visible when the source frameworks identify different dimensions of one arrangement. A capability assessment may register fewer exacerbations and greater opportunity to remain at home. Relational autonomy asks whether participation and refusal rest on viable social conditions; epistemic justice whether the patient's account can affect interpretation; contextual integrity whether information flows remain appropriate to clinical purpose. Prohumanism assesses these judgements together at the boundary of enhancement. Suppose core respiratory support depends on accepting an optional module monitoring sleep and activity. Capability may improve, while effective exit narrows and domestic life acquires a clinical purpose. This changes governance: core care should be decoupled from ancillary optimisation, and each module should require renewed justification instead of inheriting legitimacy. If the source frameworks already require the same classification and remedy, the additional label should be reduced.

The framework also differs from bioconservatism. Embodiment supplies a normative anchor, not a prohibition on alteration. Repair, prevention, diagnosis, assistive technologies, and some enhancements can widen capabilities and reduce unjust constraints. The diabetic-retinopathy system provides a straightforward example: a bounded diagnostic function can expand access to screening and enable timely referral without governing the patient's life as a continuing improvement project (Abràmoff et al. 2018). The relevant distinction lies between support directed towards a specified good and optimisation whose endpoint continually recedes.

4.3. Dignified longevity as an evaluative *telos*

Dignified longevity names the evaluative commitment that gives prohumanism direction. Lifespan extension counts years, healthspan seeks years lived in better health, and the World Health Organization's account of healthy ageing focuses on functional ability produced through the interaction between intrinsic capacity and environment (World Health Organization 2015). Together, these concepts clarify that longevity's value depends on how life is lived. They leave unresolved how data-intensive systems should be limited when the pursuit of function becomes a continuing regime of measurement and behavioural correction.

Dignified longevity evaluates extended or supported life through the conditions that make continued life practically inhabitable and normatively one's own. These conditions include capabilities, recognition, relationships of care, meaningful participation, access to support, and standing in decisions that shape one's trajectory. Dignity does not describe independence from others or freedom from vulnerability. Dependence can coexist with agency when relations and institutions preserve voice, respect, and practical authority. Nor does dignity require a single account of successful ageing. People may reasonably differ about risk, treatment burden, privacy, enhancement, and the value of continued monitoring.

Dignified longevity cannot be represented by a composite score. Metrics may inform specific judgements, but no dashboard can aggregate agency, recognition, care, and meaning into a universally valid target. The concept functions as a test of institutional orientation. A technology supports it when the benefits are connected to goods the person can recognise, when burdens and dependencies remain open to review, and when refusal does not erase access to reasonable alternatives. A system undermines it when longer or safer life is purchased through continuous subordination to an evaluative regime that the person cannot meaningfully question or leave.

The criterion is counterfactual and distributive: would the longevity benefit remain justified if the person could understand its practical meaning, supply context, refuse ancillary optimisation, and retain reasonable care? Choice does not settle every conflict; urgent risk, professional duties, and fair allocation remain. Healthy ageing identifies capabilities and supportive environments; dignified longevity asks whether the institutional means of producing them preserve practical authorship of the trajectory. Measured function may improve while this condition deteriorates when assistance requires pervasive monitoring or disagreement cannot change consequences. Dependence and intensive support can satisfy the criterion when the person remains recognised, heard, and able to influence care. If this test never changes a judgement beyond healthy ageing, the concept is redundant.

Remote monitoring makes the distinction concrete. Remaining at home and avoiding exacerbations can be central goods, yet their value does not automatically justify indefinite surveillance, expansion into unrelated domains, or an adherence score that determines access. The practical meaning of ageing at home depends on whether the home remains a lived environment or becomes an extension of the clinic. In long-term care, different actors require different forms of explanation and practical knowledge, and older adults themselves must be included in determining what forms of intelligibility and assistance matter (van Leersum and Peine 2026). Dignified longevity treats that situated participation as constitutive of success.

4.4. From telos to constraints

Once improvement is judged against dignified longevity, four governance problems become unavoidable. Intervention can impose burdens that exceed the patient-relevant benefit; authority can become insulated from challenge; health-related monitoring can spread across the whole of life; and algorithmic representations can displace the interpretive standing of patients and professionals. Proportionality, contestability, non-totalisation, and epistemic autonomy respond respectively to those risks.

The constraints are analytically distinct but interdependent. Proportionality is difficult to assess when targets and assumptions cannot be contested. Contestability is hollow when withdrawal carries unacceptable penalties. Non-totalisation requires an account of which human goods justify protected boundaries. Epistemic autonomy depends on institutions that treat contextual reasons as capable of changing decisions. Justice runs through all four because costs, digital competence, data quality, and access to alternatives are unequally distributed. The constraints acquire substance only when they

shape design choices, organisational responsibilities, and the consequences attached to system outputs.

5. Deriving and operationalising prohumanist governance

5.1. Proportionality: benefit, burden, and reversibility

Proportionality asks whether the scope and burden of an AI-mediated intervention are justified by a concrete, patient-relevant good. Predictive accuracy contributes to that judgement but does not settle it. Data collection, monitoring frequency, behavioural direction, false alerts, workload transfer, and the consequences of classification all carry burdens. The justificatory threshold should rise as an arrangement becomes more intrusive, persistent, difficult to reverse, or consequential for access to services.

The proportionality judgement begins with a precise statement of purpose. The institution should identify the health-related good being pursued, the population and setting for which the intervention is intended, the evidence connecting each data stream to that good, and the conditions under which monitoring or recommendation will cease. Additional functions require independent justification instead of inheriting legitimacy from the original purpose. The principle also favours less intrusive alternatives when they can provide comparable benefit and requires periodic review of whether the original balance still holds.

In MonitAir-like remote monitoring, earlier exacerbation detection can justify collecting specified clinical and patient-reported variables. It does not automatically justify indefinite expansion to sleep, movement, nutrition, or broad behavioural scoring. A proportional deployment would review whether each added stream improves outcomes important to patients, whether it creates avoidable burden, and whether a lower-intensity mode is available. Reversibility is part of the same analysis: withdrawal should lead to a safe fallback pathway rather than exclusion from ordinary care.

In episodic retinal screening, the purpose, data type, and consequence can be more tightly delimited. Images are collected for a specified diagnostic target and the output guides referral. Proportionality still requires validation across relevant populations, quality controls, and attention to access, but the intervention does not need continuing access to unrelated aspects of conduct. The contrast shows that the constraint is sensitive to architecture and context rather than hostile to automation.

5.2. Contestability: responsive institutions, not explanation alone

Contestability concerns the capacity to challenge how a system represents the world and what institutions do with that representation. It extends beyond receiving an explanation. A person may understand that a score is based on particular variables yet remain unable to correct the data, question the proxy, request competent review, or alter the consequence. Contestable AI must be responsive to

intervention throughout the lifecycle, including the definition of targets, design of interfaces, deployment rules, and resolution of disputes (Alfrink et al. 2023).

Practical requirements vary by role. Clinicians need authority to register and act upon reasoned disagreement without being converted into rubber stamps. Patients need accessible routes to correct information, add context, challenge consequences, and obtain human consideration. Organisations need records of overrides, appeals, error patterns, and recurring disputes so that contestation can reshape the system as well as repair individual cases. Kempt et al. (2023) call attention to meaningful disagreement in medical contexts, where the burden of proof should not shift automatically towards the human professional. A nominal override that is penalised or ignored does not preserve responsible judgement.

The object of challenge includes more than the final output. Obermeyer et al. (2019) show why proxy choice must be contestable: a model can operate as designed and still institutionalise an unjust target. In remote monitoring, a patient's low engagement score may reflect disability, connectivity, caring responsibilities, or disagreement with the programme. A process that permits correction only of sensor error leaves the normative meaning of engagement untouched. Effective contestation allows reasons from lived context to question the category itself and the consequences attached to it.

Explanation remains valuable when it is fitted to the decision and the audience. Long-term-care professionals may need practical knowledge about how to act, while managers and developers require different forms of system information (van Leersum and Peine 2026). Babic et al. (2026) argue that justifiability should not be collapsed into explainability: in high-impact settings, the central question is whether a decision can be supported by reasons that affected persons and responsible institutions can evaluate. Contestability operationalises that difference by linking reasons to possible correction and remedy.

5.3. Non-totalisation: preserving limits around health governance

Non-totalisation protects the distinction between supporting health and governing life through health. Digital systems can cross that boundary gradually because new variables appear relevant to prediction and prevention. Sleep, movement, location, consumption, social interaction, and emotional state may each bear some relation to health. Their relevance does not grant an unlimited institutional mandate to collect, combine, and act upon them.

Its requirements appear at three levels. Rights and expectations protect domains and periods in which people are not continuously assessed. System architecture needs purpose limitation, data minimisation, retention limits, and barriers against cross-domain scoring. Organisational deployment requires sunset clauses, periodic necessity review, restrictions on secondary use, and non-punitive options for reduced monitoring or exit. Contextual integrity is useful here because the ethical status of a data flow depends on actors, purposes, information types, and transmission principles, while consent alone cannot establish that integrity (Nissenbaum 2010).

Non-totalisation does not impose a fixed maximum duration for care. Chronic conditions can justify long-term monitoring, and integrated care can require data from several domains. The constraint asks whether each extension remains tied to a specified purpose and whether parts of life retain protection from automatic actionability. A home-monitoring platform becomes totalising when domestic routines are progressively absorbed into a clinical-administrative regime and when opting out carries the loss of ordinary support. The boundary is institutional, not merely spatial: the home can remain a lived environment while containing medical technology, or it can be reorganised as a permanently supervised site of compliance.

Function creep provides an observable failure mode. Data gathered for exacerbation detection may later train commercial models, determine eligibility, or support behavioural nudging. A governance process should map actual data flows, record new uses, require renewed justification, and provide deletion or de-linking where appropriate. The EU AI Act establishes lifecycle and risk-management duties for relevant systems, while WHO guidance emphasises autonomy, accountability, inclusion, and sustainability in AI for health (European Parliament and Council of the European Union 2024; World Health Organization 2021). Non-totalisation supplies a normative reason to use those mechanisms to preserve bounded purposes rather than treating compliance as an end in itself.

5.4. Epistemic autonomy: interpretive standing in mediated care

Epistemic autonomy is the practical capacity to understand, contribute to, question, and negotiate the meaning of information used to govern one's health. It does not require complete technical mastery. Patients and professionals occupy different epistemic roles, and responsible care depends on combining clinical knowledge, technical expertise, and lived experience. The constraint protects the standing of each contribution against an infrastructure that can render one representation authoritative by default.

Fricker's account of epistemic injustice illuminates two risks. Testimonial injustice occurs when a person's account receives less credibility than a metric or model output without adequate reason. Hermeneutical injustice arises when available categories cannot adequately express an experience or concern. Digital-health systems can intensify both when dashboards privilege standardised variables and leave contextual information as an unstructured residue (Fricker 2007). The harm is not only informational error. It is the loss of practical authority in decisions about one's own trajectory.

Institutional protections include interfaces that communicate uncertainty and purpose, opportunities to add contextual information, documentation of professional disagreement, and decision processes in which reasons can alter outcomes. Organisations should examine whether patients understand the consequences of participation and withdrawal, whether alternative pathways remain available, and whether explanations are adapted to different users. Subías-Beltrán et al. (2026) connect autonomy to diagnostic governance questions distributed across organisational roles, while Herzog et al. (2025)

locate trustworthiness within medical AI ecosystems. These approaches shift attention from isolated consent towards ongoing institutional responsibility.

Meaningful exit is part of epistemic autonomy because a person cannot negotiate a system freely when rejection eliminates reasonable care. Exit is not absolute; scarce resources and clinical safety may constrain alternatives. It must nevertheless be more than a formal button. The institution should disclose consequences, provide proportionate fallback, and avoid interpreting withdrawal as evidence of irresponsibility. Voice without exit can become managed participation, while exit without voice leaves the underlying classifications unchanged. Prohumanist governance therefore links both to contestability and non-totalisation.

5.5. Cross-case evaluation and institutional responsibility

The four constraints produce different evaluations of the two cases. In AI-enabled COPD monitoring, proportionality asks whether data and workload are connected to benefits that matter to patients; contestability reaches alerts, proxies, and organisational consequences; non-totalisation limits extension into unrelated domains; and epistemic autonomy requires the patient's account and clinical judgement to remain capable of changing interpretation. The system may satisfy these requirements, but its ethical status depends on implementation rather than the label "remote monitoring".

In the contrastive retinal-screening configuration considered here, the architecture is narrower. Proportionality focuses on diagnostic performance, access, and the burden of referral. Contestability requires routes for handling poor-quality images, error, and disputed outcomes. Non-totalisation limits secondary use and prevents the screening result from becoming a general behavioural score. Epistemic autonomy requires clear communication and a pathway for clinical discussion. The task remains bounded only while these controls preserve its specified purpose and endpoint.

Responsibility must be allocated across the ecosystem. Providers determine intended functions and technical limitations; deploying organisations set workflows, thresholds, data practices, and fallback routes; clinicians interpret or act on outputs; procurement and governance bodies decide which benefits and risks count; regulators establish minimum duties. Patients should not carry the burden of coordinating failures among these actors. The ecosystem perspective is especially important where adaptive systems, vendors, and institutional routines evolve after initial approval (Herzog et al. 2025).

Taken together, the constraints structure the reasons that institutions must be able to give. Their value is defeated if they merely rename familiar principles without changing design or deployment. They become substantive when they identify bounded purposes, preserve avenues of challenge, maintain protected limits, and keep affected persons within the production of practical meaning. Legal duties remain a floor for that work, while the constraints specify the normative questions that implementation must continue to answer.

6. Scope, limits, and implications

The argument is conceptual and normative. It does not establish how prevalent optimisation drift is across digital health, and the two cases do not support statistical generalisation. MonitAir provides a situated implementation case whose risks and benefits are context-specific; diabetic-retinopathy screening illustrates a more bounded architecture but can still fail through bias, access barriers, poor integration, or inadequate remedy. Empirical research is needed to examine how purposes expand, how stopping rules operate, what costs attach to withdrawal, and how patients and professionals experience disagreement.

The remote-monitoring analysis contains two evidential layers that must remain separate. MonitAir is a documented implementation examined by Behdadi et al. (2026); the later expansion into sleep, nutrition, adherence, cross-domain scoring, and eligibility is an analytical extension used to test the concept of drift. No claim is made that MonitAir has undergone that expansion. Generalisation would require longitudinal evidence about changes of purpose, procurement and reimbursement incentives, module addition, withdrawal practices, override records, and patient and professional experience across multiple institutions. The retinal contrast is equally conditional: its bounded status depends on a specified task, endpoint, data domain, route of remedy, and absence of functional expansion. Actual deployments that lack those properties may fall outside the contrastive class.

Transhumanism also remains more diverse than the perfectionist ideal type used here. Some transhumanist positions endorse autonomy, social justice, risk reduction, and limits that may resist optimisation-by-default. The comparison should be abandoned wherever the three conditions fail to discriminate or where it obscures more direct explanations of power and political economy. Structural affinity has a modest role: it identifies an operative grammar of open-ended corrigibility, not a hidden ideology or a universal trajectory.

Prohumanism is similarly defeasible. Its normative resources derive from established traditions, and the term is unnecessary if their combination already yields the same distinctions and obligations without loss. The framework also faces conflicts that cannot be resolved by stipulation. A capability valued by one person may impose burdens on another; meaningful exit may be constrained by scarce alternatives; patient preference can conflict with professional duties or distributive justice. Prohumanism supplies a field of reasons and limits, leaving those conflicts to practical and institutional judgement.

The framework is most developed for platformised, data-intensive, and behaviourally directive systems in relatively formal health-care institutions. Maccaro's (2026) analysis cautions against generalising governance assumptions across unequal infrastructures and culturally plural practices of care. Technological subsidiarity, solidarity, and local knowledge may require different institutional forms. Dignified longevity cannot be detached from material conditions such as housing, income, digital access, care labour, and the availability of non-digital alternatives. Otherwise, appeals to autonomy and choice can conceal structural abandonment.

These limits also indicate an empirical programme. Comparative studies could test whether the three conditions reliably distinguish bounded from drifting deployments; participatory work could examine

how patients define tolerable monitoring and meaningful exit; organisational research could track overrides, appeals, purpose expansion, and decommissioning. Governance evaluation should address model accuracy after deployment and the institution's continuing capacity to limit what the model is allowed to make actionable.

The concepts also have distinct failure tests. Structural affinity is over-specified if one condition, or any loose pair of conditions, classifies cases as effectively as the full configuration. Prohumanism is redundant if removing the term leaves every priority, classification, and governance decision unchanged. Dignified longevity is indeterminate if it cannot distinguish acceptable dependence from safety secured through coercive monitoring. Findings of that kind would require reformulation or abandonment rather than merely further illustration.

7. Conclusion

AI-enabled health systems alter more than the technical capacity to detect and predict. They organise the practical relation between a detectable difference and a justified response. When measurement becomes recursive, purposes expand, and stopping requires defence, a bounded intervention can drift towards optimisation-by-default. That transition is driven most directly by platform incentives, preventive rationalisation, and organisational pressures. It acquires a structural affinity with transhumanist perfectionism only where weak stopping rules, indefinitely corrigible embodiment, and package-style escalation converge.

The comparison matters because it distinguishes a problem of governance from a general objection to technology. A bounded diagnostic system and a continuous monitoring platform can both use advanced AI while creating different normative relations to the person. The relevant question is whether correction remains subordinate to a specified human good or becomes an open horizon that progressively absorbs more of life.

Prohumanism provides a reconstructive standard for making that distinction. Dignified longevity evaluates longer and technologically supported lives through embodied agency, plural goods, recognition, justice, and meaningful exit. Proportionality, contestability, non-totalisation, and epistemic autonomy translate those commitments into institutional requirements. They ask whether benefits justify burdens, whether authority remains responsive to challenge, whether health governance retains boundaries, and whether patients and professionals remain participants in the interpretation of their own situations.

The central limit is also the central test. Prohumanist governance has value only when it changes what institutions design, justify, review, and discontinue. It should permit effective technologies whose purposes remain bounded and contestable, while resisting arrangements in which technical possibility silently becomes a duty of continuous self-correction. Dignified longevity depends on preserving that distinction: improvement should enlarge the conditions of a life one can inhabit and author, rather than turning life into a permanently unfinished optimisation project.

Declarations

Author Contribution

Conceptualization; writing—original draft; writing—review and editing.

References

1. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC (2018) Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. npj Digit Med 1:39. https://doi.org/10.1038/s41746-018-0040-6
2. Agar N (2013) Truly human enhancement: a philosophical defense of limits. MIT Press, Cambridge, MA. https://doi.org/10.7551/mitpress/9780262026635.001.0001
3. Alfrink K, Keller I, Kortuem G, Doorn N (2023) Contestable AI by design: towards a framework. Minds Mach 33:613–639. https://doi.org/10.1007/s11023-022-09611-z
4. Babic B, Cohen IG, Savulescu J (2026) Justifiability and AI: putting explainability in its place. AI Soc. https://doi.org/10.1007/s00146-026-03030-9
5. Behdadi D, Vanfleteren LEGW, Sundemo D (2026) Ethical risks of AI-enabled remote patient monitoring for COPD: a multi-dimensional use case analysis. AI Soc 41:5645–5659. https://doi.org/10.1007/s00146-025-02806-9
6. Bostrom N (2005) A history of transhumanist thought. J Evol Technol 14(1):1–25. https://jetpress.org/volume14/bostrom.pdf
7. Brennan C (2023) Weak transhumanism: moderate enhancement as a non-radical path to radical enhancement. Theor Med Bioeth 44:229–248. https://doi.org/10.1007/s11017-023-09606-6
8. European Parliament and Council of the European Union (2024) Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Off J Eur Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng
9. Fricker M (2007) Epistemic injustice: power and the ethics of knowing. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780198237907.001.0001
10. García-Barranquero P, Alonso Fernández M (2025) Differentiating human enhancement and transhumanism: better or perfect? Scientia et Fides 13(1):121–142. https://doi.org/10.12775/SetF.2025.007
11. Goodin RE (1985) Protecting the vulnerable: a reanalysis of our social responsibilities. University of Chicago Press, Chicago
12. Hauskeller M (2016) Mythologies of transhumanism. Palgrave Macmillan, London. https://doi.org/10.1007/978-3-319-39741-2
13. Herzog C, Blank S, Stahl BC (2025) Towards trustworthy medical AI ecosystems - a proposal for supporting responsible innovation practices in AI-based medical innovation. AI Soc 40:2119–2139.

https://doi.org/10.1007/s00146-024-02082-z

14. Kempt H, Heilinger JC, Nagel SK (2023) I'm afraid I can't let you do that, Doctor': meaningful disagreements with AI in medical contexts. AI Soc 38:1407–1414. https://doi.org/10.1007/s00146-022-01418-x
15. Lupton D (2016) The quantified self: a sociology of self-tracking. Polity, Cambridge
16. Lyreskog DM, McKeown A (2022) On the (non-)rationality of human enhancement and transhumanism. Sci Eng Ethics 28:52. https://doi.org/10.1007/s11948-022-00410-4
17. Mackenzie C, Stoljar N (eds) (2000) Relational autonomy: feminist perspectives on autonomy, agency, and the social self. Oxford University Press, New York. https://doi.org/10.1093/oso/9780195123333.001.0001
18. Maccaro A (2026) Algorithmic ethics and healthcare pluralism: rethinking care between automation and global inequality. AI Soc 41:4599–4610. https://doi.org/10.1007/s00146-026-02854-9
19. Nissenbaum H (2010) Privacy in context: technology, policy, and the integrity of social life. Stanford University Press, Stanford. https://doi.org/10.1515/9780804772891
20. Nussbaum MC (2011) Creating capabilities: the human development approach. Harvard University Press, Cambridge, MA
21. Obermeyer Z, Powers B, Vogeli C, Mullainathan S (2019) Dissecting racial bias in an algorithm used to manage the health of populations. Science 366(6464):447–453. https://doi.org/10.1126/science.aax2342
22. Ruckenstein M, Schüll ND (2017) The datafication of health. Annu Rev Anthropol 46:261–278. https://doi.org/10.1146/annurev-anthro-102116-041244
23. Sharon T (2017) Self-tracking for health and the quantified self: re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. Philos Technol 30(1):93–121. https://doi.org/10.1007/s13347-016-0215-5
24. Sharon T (2018) When digital health meets digital capitalism, how many common goods are at stake? Big Data Soc 5(2):1–12. https://doi.org/10.1177/2053951718819032
25. Subías-Beltrán P, Unceta I, de Lecuona I, Pujol O (2026) Asking the right questions: a governance approach to uphold human autonomy in artificial intelligence. AI Soc 41:1149–1173. https://doi.org/10.1007/s00146-025-02611-4
26. Thompson E (2007) Mind in life: biology, phenomenology, and the sciences of mind. Harvard University Press, Cambridge, MA
27. van Dijck J (2014) Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. Surveill Soc 12(2):197–208. https://doi.org/10.24908/ss.v12i2.4776
28. van Dijck J, Poell T, de Waal M (2018) The platform society: public values in a connective world. Oxford University Press, Oxford. https://doi.org/10.1093/oso/9780190889760.001.0001
29. van Leersum CM, Peine A (2026) Only practical knowledge or knowing the algorithm? Notions and necessities of explainable artificial intelligence in long-term care. AI Soc 41:1189–1206.

https://doi.org/10.1007/s00146-025-02614-1

30. Varela FJ, Thompson E, Rosch E (1991) The embodied mind: cognitive science and human experience. MIT Press, Cambridge, MA
31. World Health Organization (2015) World report on ageing and health. World Health Organization, Geneva. https://www.who.int/publications/i/item/9789241565042
32. World Health Organization (2021) Ethics and governance of artificial intelligence for health: WHO guidance. World Health Organization, Geneva. https://www.who.int/publications/i/item/9789240029200
33. Zuboff S (2019) The age of surveillance capitalism: the fight for a human future at the new frontier of power. PublicAffairs, New York