

# Future-Building as Maintenance: Rate-Based Health Evaluation for Indefinite Horizons

Daniel S. Goldman 

February 6, 2026

## Abstract

Standard health evaluation sums quality-adjusted life years over a lifetime, but that sum becomes meaningless when biological lifespan has no fixed upper bound. A shift from sum to rate—measuring sustained quality of life per unit time—resolves the problem and, more interestingly, reveals deep connections between aging biology, existential-risk governance, and institutional measurement design. The hallmarks-of-aging framework already describes health in the language of failure modes and maintenance; reliability engineering has long applied similar reasoning to complex systems; and the futurist concern with keeping the long-run future accessible turns out to be the same problem viewed at civilisational scale. The central challenge is governance: any rate worth optimising will be gamed, and the measurement system that supports it needs the same ongoing, adaptive maintenance as the bodies and institutions it tracks. The treatment remains agnostic about mathematical formalisation so that the argument survives any future rework of the underlying formalism.

## 1 A world without a finish line

For most of recorded history, the question “how long will this person live?” had a rough but serviceable answer: not very long, by modern standards, and with an upper bound that everyone understood. Public health could plan around that bound. Insurance could price it. Governments could discount future costs against it. The entire apparatus of health evaluation—cost-effectiveness analysis, technology assessment, resource allocation—grew up inside a world where lives had a characteristic shape: they started, they lasted a few decades, and they ended.

That shape is starting to change. Simon Melov’s review of the geroscience programme describes a research agenda organised around a single shift in target: rather than treating chronic diseases one at a time, go after the aging processes that produce them all ([Melov, 2016](#)). If that programme succeeds—even partially—then the planning horizon for a human life stretches in ways that quietly break the standard evaluation toolkit.

Quality-adjusted life years, the workhorse metric of health technology assessment, are defined as a sum. You take the quality of each year lived, you add them up, and you compare interventions by how many QALYs they buy. That sum presupposes a terminus. When the horizon is open-ended, the sum either diverges or depends on an arbitrary cutoff, and cost-effectiveness ratios built on it inherit the same fragility.

The fix is to measure a rate instead: quality of life per unit time, sustained over an ongoing regime. The long-run average reward per unit time in a regenerative process has clean existence conditions in renewal-reward theory ([Sigman, 2018](#); [Gallager, 2011](#)) and extends naturally to controlled settings through average-cost optimisation ([Feinberg et al., 2012](#)). The mathematics matters less here than the conceptual shift. A rate stays bounded and interpretable whether the horizon is

ten years or ten thousand. It captures how well life is going, how much effort maintenance requires, and how often catastrophe interrupts—in one object.

Call that object QALY flow.

## Ensemble Context

**Paper ID:** d7006442-c67e

This paper is part of a four-paper ensemble on health evaluation under indefinite horizons. The root paper (c2f3de53-f086) defines the core mathematical framework—QALY flow as a renewal-reward ratio, with formal results on existence, cost-effectiveness, perturbation stability, and distributional aggregation. This paper connects aging biology, reliability engineering, existential-risk governance, and institutional measurement design under a unified maintenance framing, showing how the rate-based perspective operates across scales without requiring the formal apparatus. Companion papers explore the civilizational context of longevity (73997063-e1ed) and translate the framework into public health practice (80fbe599-356c).

## 2 Aging as engineering

In 2013, Carlos López-Otín and colleagues published a taxonomy of aging that reads, almost eerily, like a failure-mode-and-effects analysis for a biological system (López-Otín et al., 2013). Genomic instability, telomere attrition, epigenetic drift, loss of proteostasis, mitochondrial dysfunction, cellular senescence, stem cell exhaustion, deregulated nutrient sensing, altered intercellular communication—each hallmark identifies a subsystem whose gradual degradation contributes to the organism’s declining performance. Their 2023 update expanded the list to include chronic inflammation, dysbiosis, and a web of cross-hallmark interactions, making the picture look less like a checklist and more like a complex system with coupled failure modes (López-Otín et al., 2023).

Maxim Finkelstein had already made the engineering parallel explicit. His work at the Max Planck Institute mapped reliability-theory concepts—redundancy depletion, damage accumulation, hazard-rate dynamics—onto biological aging trajectories (Finkelstein, 2006). The conceptual move is to treat mortality and morbidity as the output of a system whose components degrade, whose redundancy erodes, and whose performance can be modelled with the same tools that aerospace and energy engineers use for aircraft engines and power grids.

Once that move is made, geroscience starts to look like a maintenance engineering programme. Instead of waiting for a specific subsystem to fail and then attempting a targeted repair—treating heart disease, excising a tumour—you invest in the conditions that keep multiple subsystems operating within tolerance. Prevention becomes quality assurance. Monitoring becomes condition-based maintenance. And the success criterion shifts from “did we fix the problem?” to “is the system still running well?”—which is exactly what a rate measures.

The analogy is imperfect in one important respect, and the imperfection is itself instructive. In industrial maintenance, Alberto Pliego Marugán and colleagues have studied what happens when repairs are *increasingly imperfect*: each intervention restores less function than the last, and the optimal policy has to adapt, shifting from repair toward replacement as the system ages (Pliego Marugán et al., 2024). Biology almost certainly behaves this way. The first round of a senolytic therapy may clear senescent cells effectively; the tenth round operates on a body that has accumulated other forms of damage, in tissues that have been repeatedly stressed, with compensatory mechanisms that may have shifted. The engineering literature’s response to imperfect repair is adaptive control—adjust the maintenance schedule, change the intervention mix, re-evaluate

the monitoring thresholds—and that response translates directly into a design principle for health systems under indefinite horizons: the maintenance policy itself must be maintained.

Here is where the rate perspective starts to earn its keep. Under a sum-based evaluation, an intervention that adds five healthy years scores well regardless of what happens afterwards. Under a rate-based evaluation, an intervention is valuable only insofar as it contributes to sustained performance. A therapy that produces a burst of improvement followed by accelerated decline looks good in the sum and bad in the rate. A preventive programme that slightly reduces the pace of degradation across decades looks modest in the sum and transformative in the rate. The metric changes what you value, and what you value changes what you build.

### 3 Maintenance all the way up

The engineering framing is compelling for bodies. It becomes more compelling—and more consequential—when you notice that the same framing applies at every scale above the individual.

Eileen Crimmins has argued for a set of “social hallmarks of aging” that parallel the biological ones (Crimmins, 2020). Education, income inequality, chronic stress exposure, environmental toxin burden, social isolation—these upstream conditions shape biological aging trajectories as powerfully as any cellular mechanism. People in deprived neighbourhoods age faster, measured biologically, than people in affluent ones. The hallmarks of aging are real, but they do not operate in a social vacuum; they are *modulated* by the conditions in which people live.

Terrie Moffitt’s work on the geroscience translation agenda sharpens the point (Moffitt, 2020). Even when a biological intervention works in the lab, its population-level impact depends on adherence, access, behaviour, and the social infrastructure that delivers it. The bottleneck between a promising therapy and a sustained improvement in population health is almost never the biology alone. It is the system of behaviour, institutions, and incentives through which the biology is deployed.

From the standpoint of QALY flow, this is a single observation with a long reach: the sustained rate of health quality is a property of the entire sociotechnical system, not of the biomedical subsystem alone. Reducing chronic stress is upstream maintenance on the biological hallmarks. Educational investment that improves health literacy is condition monitoring at scale. Limiting toxin exposure is preventive maintenance on the population’s genome. Each of these interventions changes the steady-state distribution of health states across the population, and a rate-based metric registers that change.

The same logic extends to catastrophic risk. Nick Bostrom’s work on existential risk argues that preserving the long-run future is extraordinarily valuable, because the potential future is vast and the loss from premature truncation is correspondingly enormous (Bostrom, 2013). Hilary Greaves and William MacAskill formalise conditions under which the expected value of the far future dominates present-day considerations (Greaves and MacAskill, 2021). These arguments translate cleanly into rate language: catastrophic events reduce QALY flow both by destroying lives directly and by disrupting the maintenance systems that sustain quality for everyone else.

Bostrom’s earlier paper on “astronomical waste” identified a striking asymmetry between speed and safety in technological development (Bostrom, 2003). Accelerating progress increases the rate of value creation, but increasing catastrophic risk decreases the probability that the value-creating regime persists at all. Under a rate-based objective, that asymmetry becomes a portfolio constraint: you optimise the sustained rate subject to the requirement that the regime survives. Safety investment is infrastructure for the rate, the same way that pandemic preparedness is infrastructure for population health and building inspection is infrastructure for a city’s housing stock.

The World Health Organization’s framework for climate-resilient health systems illustrates what this looks like institutionally ([World Health Organization, 2015](#)). The framework organises adaptation around anticipation, response capacity, and recovery—exactly the dimensions that determine how a shock affects the sustained rate. A health system that can absorb a pandemic or a climate disaster and continue delivering care maintains QALY flow through the disruption. One that collapses under stress creates a period of severely reduced quality that drags the long-run average down, potentially for years after the acute event has passed.

So the maintenance picture scales: bodies are maintained by biological repair, social conditions, and behavioural support; populations are maintained by health systems, social infrastructure, and environmental standards; and the civilisational trajectory is maintained by catastrophe prevention, resilience engineering, and institutional coordination. QALY flow is legible at each of these levels because it measures the same thing at each level: how well the system is performing per unit time, on average, over the long run.

## 4 Managing the measurement

All of this—the maintenance engineering, the social determinants, the civilisational portfolio—assumes that you can actually measure QALY flow well enough to govern with it. The assumption deserves scrutiny, because the history of institutional metrics is a history of measurement corrupting itself.

Michael Stumborg and colleagues at CNA produced an unusually thorough catalogue of the ways that metrics degrade when they become targets ([Stumborg et al., 2022](#)). The mechanisms are varied—proxy manipulation, strategic data omission, threshold gaming, reallocation of effort from unmeasured to measured activities—but the pattern is monotonous. Once a number matters for resource allocation, the people being measured find ways to improve the number that do not improve the thing the number was supposed to track. Hospital readmission rates go down because hospitals reclassify admissions. Test scores go up because teachers teach to the test. Crime statistics improve because police reclassify offences.

A QALY-flow metric deployed at scale would face every one of these pressures, intensified by the fact that it is designed for continuous optimisation rather than periodic reporting. A hospital system evaluated on sustained health quality per patient-year has incentives to select healthier patients, to classify complications as new episodes rather than ongoing ones, and to invest in the measurement apparatus—self-reported quality surveys, biomarker panels—in ways that inflate the reported rate without improving the underlying reality.

This is a genuine problem, and the shape of the solution is by now a familiar one. Measurement integrity is a maintenance problem. The measurement system degrades under incentive pressure the way a biological system degrades under oxidative stress or a building degrades under weather. It requires monitoring, repair, adaptation, and occasional replacement of components that have become too corrupted to salvage.

Holger Strulik’s survey of aging measurement illustrates the challenge at the level of individual assessment: different constructions of “biological age” yield different rankings of individuals and different conclusions about intervention efficacy ([Strulik, 2021](#)). At the population level, the problem multiplies. Which quality-of-life instrument do you use? How do you handle non-response? How do you distinguish genuine improvement from reporting bias? Each of these is a maintenance task for the measurement system, and each requires ongoing attention.

The institutional precedents are instructive. The UK Treasury’s Green Book codifies an appraisal process built around explicit options analysis, structured uncertainty treatment, and manda-

tory post-implementation review (HM Treasury, 2026). NIST’s risk assessment framework organises ongoing evaluation around threat identification, vulnerability analysis, and iterative updating (National Institute of Standards and Technology, 2012). Both frameworks treat evaluation as a *process* rather than an event—something that is never finished, only current. Governing a QALY-flow metric requires the same stance: the measurement is never “done.” It is an ongoing regime of data collection, integrity checking, adversarial auditing, and institutional accountability that must itself be evaluated and maintained.

The distributional dimension adds another layer. When you optimise a single aggregate rate, you naturally concentrate effort where improvements are cheapest. The people who are hardest to help—those with complex needs, poor baseline health, or limited access—become drags on the metric, and a system under performance pressure may quietly abandon them. Miqdad Asaria, Susan Griffin, and Richard Cookson developed distributional cost-effectiveness analysis specifically to make these dynamics visible (Asaria et al., 2013). The technique defines outcomes for identifiable subgroups and applies explicit equity weights, so that the question “who benefits?” is answered alongside “how much benefit?” Under a QALY-flow regime, the distributional analysis operates on group-level rates rather than group-level sums, but the logic is identical: aggregate optimisation without distributional scrutiny produces aggregate numbers that mask concentrated harm.

Conventional tools do not vanish in this picture. The Green Book’s discounting guidance remains entirely appropriate for bounded subproblems—the construction costs of a new hospital, the ramp-up phase of a new intervention, the payback period of a capital investment (HM Treasury, 2026). Discounting and rate-based evaluation are complements: discounting handles the transitional and time-limited components of a policy, while QALY flow handles the steady-state, continuing-effects dimension. The Green Book already distinguishes between these categories; the rate metric fills a slot that the existing framework recognises but cannot currently populate under open-ended horizons.

## 5 What remains

The gaps are honest ones. Measuring a population’s QALY flow under conditions where the measurement itself shapes behaviour—where monitoring changes adherence, where reporting incentives distort data, where the interventions being evaluated alter the social fabric that generates the data—is a problem that adaptive control theory can frame (Tadepalli and Ok, 1998; Wang et al., 2023) but that institutional practice has barely begun to address. Setting safety margins for catastrophic risk when the models themselves may be misspecified is a hard problem in robust optimisation (Wang et al., 2023; Boucherie et al., 2023), and the gap between the theoretical tools and the practical governance of real institutions is wide. Designing portfolio governance that bridges short-horizon budgets with indefinite-horizon objectives requires integrating discounting practice (HM Treasury, 2026) with rate-based steady-state reasoning in ways that have no established template. And sustaining distributional fairness over horizons long enough for political coalitions to dissolve, reform, and dissolve again is a governance challenge that exceeds anything in the current literature on equity weighting (Asaria et al., 2013; Crimmins, 2020).

Building the future, it turns out, is a maintenance job. And maintenance, done well, is what makes everything else possible. A person whose biology is sustained rather than deteriorating has time to master difficult skills, to undertake projects that span decades, to accumulate and transmit knowledge at a depth that a seventy-year window barely permits. A civilisation that has learned to maintain its population’s health, its institutional integrity, and its capacity to absorb shocks has freed up the energy and attention that currently go to managing decline. The headaches of

measurement, fairness, adaptive governance, and catastrophe prevention are real, and they are the price of entry to a world where human ambition is no longer truncated by human fragility.

## References

- Miqdad Asaria, Susan Griffin, and Richard A. Cookson. Distributional cost-effectiveness analysis: A tutorial. CHE Research Paper 92, Centre for Health Economics, University of York, York, UK, 2013.
- Nick Bostrom. Astronomical waste: The opportunity cost of delayed technological development. *Utilitas*, 15(3):308–314, 2003.
- Nick Bostrom. Existential risk prevention as global priority. *Global Policy*, 4(1):15–31, 2013.
- Richard J. Boucherie et al. Average cost markov decision processes: Policy-iteration and value-iteration, 2023. Addendum to Operations Research: Introduction to Models and Methods (PDF creation date: 2023-04-01).
- Eileen M. Crimmins. Social hallmarks of aging: Suggestions for geroscience research. *Ageing Research Reviews*, 63:101136, 2020. doi: 10.1016/j.arr.2020.101136.
- Eugene A. Feinberg, Pavlo O. Kasyanov, and Nina V. Zadoianchuk. Average-cost markov decision processes with weakly continuous transition probabilities, 2012. arXiv:1202.4122.
- Maxim Finkelstein. On engineering reliability concepts and biological aging. Working Paper WP-2006-021, Max Planck Institute for Demographic Research, Rostock, Germany, 2006.
- Robert Gallager. Discrete stochastic processes, chapter 4: Renewal processes, 2011. MIT OpenCourseWare 6.262.
- Hilary Greaves and William MacAskill. The case for strong longtermism. Working Paper 5-2021, Global Priorities Institute, University of Oxford, June 2021.
- HM Treasury. Discounting: Green book supplementary guidance. Supplementary guidance, HM Treasury, February 2026a.
- HM Treasury. The green book: Central government guidance on appraisal and evaluation. Government guidance, HM Treasury, February 2026b.
- Carlos López-Otín, Maria A. Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. The hallmarks of aging. *Cell*, 153(6):1194–1217, 2013. doi: 10.1016/j.cell.2013.05.039.
- Carlos López-Otín, Maria A. Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. Hallmarks of aging: An expanding universe. *Cell*, 2023. doi: 10.1016/j.cell.2022.11.001. Online January 19, 2023.
- Simon Melov. Geroscience approaches to increase healthspan and slow aging. *F1000Research*, 5: 785, 2016. doi: 10.12688/f1000research.7583.1.
- Terrie E. Moffitt. Behavioral and social research to accelerate the geroscience translation agenda. *Ageing Research Reviews*, 63:101146, 2020. doi: 10.1016/j.arr.2020.101146.

- National Institute of Standards and Technology. Guide for conducting risk assessments. NIST Special Publication 800-30 Revision 1, National Institute of Standards and Technology, 2012.
- Alberto Pliego Marugán, Jesús María Pinar-Pérez, and Fausto Pedro García Márquez. A reinforcement learning agent for maintenance of deteriorating systems with increasingly imperfect repairs. *Reliability Engineering & System Safety*, 2024. Preprint available as arXiv:2505.20725 (v1, 2025).
- Karl Sigman. Some basic renewal theory: The renewal reward theorem, 2018. Lecture notes.
- Holger Strulik. Measuring ageing, 2021. Handbook chapter draft (version July 26, 2021).
- Michael F. Stumborg, Timothy D. Blasius, Steven J. Full, Christine A. Hughes, and Jennifer M. Scherer. Goodhart’s law: Recognizing and mitigating the manipulation of measures in analysis. Report COP-2022-U-033385-Final, CNA, September 2022.
- Prasad Tadepalli and DoKyeong Ok. Model-based average reward reinforcement learning, 1998. Technical report / manuscript (PDF dated January 13, 1998).
- Yue Wang, Alvaro Velasquez, George Atia, Ashley Prater-Bennette, and Shaofeng Zou. Robust average-reward markov decision processes, 2023. Preprint PDF.
- World Health Organization. Operational framework for building climate resilient health systems. Framework, World Health Organization, 2015.