INDIAN INSTITUTE OF SCIENCE

BENGALURU

INTRODUCTION TO PLANETARY HEALTH VALUE ASSESSMENT (PHVA) FRAMEWORK

Under the Guidance of

Prof. Sai Siva Gorthi

OPTICS & MICROFLUIDICS INSTRUMENTATION LAB (OMI LAB)

DEPARTMENT OF INSTRUMENTATION &APPLIED PHYSICS

Submitted by

Dr. Venkata Sai Rohit

MPH Intern

PRASANNA SCHOOL OF PUBLIC HEALTH

MANIPLA ACADEMY OF HIGHER EDUCATION

Internship duration

From: 08.06.2026

To: 14.07.2026

From Sustainability Technology Assessment to Planetary Health Value Assessment: History and Methodological Architecture

Preface and Scope:

This monograph traces a single line of intellectual development: from the historical emergence of technology assessment as a policy instrument in the 1960s, through the successive waves of sustainability science that broadened it, through the mature but domain-bound assessment frameworks now in routine use in environmental, social, and health economics practice, to the specific measurement gap that these frameworks leave unfilled, and finally to the proposal, critique, and provisional methodological architecture of Planetary Health Value Assessment (PHVA) as a candidate response to that gap. The account is deliberately continuous rather than modular: each chapter is written to be read as the argument that motivates the next, rather than as an independent summary of a single source document. Chapters 1 through 4 establish, largely through the history and comparative logic of existing frameworks, why a cross-sectoral evaluation architecture is needed at all. Chapters 5 through 7 set out what has been proposed to fill that gap and how it is meant to operate methodologically. Chapter 8 subjects the proposal to sustained critical scrutiny; distinguishing limitations the proposal has already conceded from limitations a rigorous external reviewer would still raise. Chapter 9 closes with a concrete research agenda.

Two terminological clarifications orient everything that follows. First, this monograph treats "Planet Health Technology Assessment" (PHTA) and "Planetary Health Value Assessment" (PHVA) as two successive drafts of the same underlying idea, not as two competing frameworks: PHTA is the earlier, pitch-stage formulation; PHVA is its substantially more disciplined successor, revised specifically in response to the kinds of criticism a methodologist would raise against PHTA. Second, "sustainability technology assessment" (STA) is used throughout as the umbrella term for the existing family of environmental, social, and life-cycle assessment methodologies reviewed in Chapter 2 — life-cycle assessment, environmental impact assessment, social impact assessment, industrial ecology, and resilience assessment — which collectively constitute the evidentiary and methodological substrate on which PHVA is built.

Chapter 1: The Need for Sustainability-Based Evaluation

1.1 The Origins of Technology Assessment

Technology assessment (TA) as a formal governmental function originated in the United States in the 1960s, at a moment when the visible consequences of large-scale industrial and infrastructural technology — from pesticide use to nuclear power to highway construction — had begun to outpace the institutional capacity to evaluate them before deployment. The US Congress established the Office of Technology Assessment (OTA) in 1972 specifically to give legislators an independent, analytically rigorous basis for anticipating the second-order and third-order consequences of technological choices, rather than relying solely on the technical claims of the agencies or industries proposing them. The underlying premise of early TA was straightforwardly consequentialist: technologies have effects beyond their immediate function, those effects are frequently unanticipated by their developers, and a public body could in principle forecast and weigh them before large, difficult-to-reverse commitments were made.

In practice, however, classical TA carried three structural blind spots that would eventually motivate the entire sustainability-assessment tradition surveyed in Chapter 2. It was overwhelmingly oriented toward economic and technical performance criteria — cost, efficiency, reliability, scalability — because these were the criteria that funding agencies and legislators found tractable and because the analytic tools available at the time were best developed for exactly this kind of accounting. It largely bracketed environmental consequences, treating pollution, resource depletion, and ecological disruption as externalities to be handled, if at all, by separate regulatory processes rather than as first-order evaluative criteria. And it systematically undervalued social and distributional dimensions: TA reports of the period rarely asked who specifically would bear a technology's costs, as distinct from what the aggregate cost-benefit ratio looked like at the level of the whole society. Each of these blind spots would later become the organizing concern of a distinct assessment tradition — environmental impact assessment for the second, social impact assessment for the third, and, later, life-cycle and industrial-ecology thinking for a more rigorous treatment of the first.

1.2 The Turn to Sustainability

The environmental crises that intensified through the 1970s and 1980s — acid rain, stratospheric ozone depletion, accumulating evidence of anthropogenic climate change, and a series of high-profile industrial and nuclear accidents — created sustained political pressure for evaluation frameworks that treated environmental consequences as constitutive of an activity's value rather than as an afterthought. The decisive conceptual turning point is the 1987 Brundtland Commission report, Our Common Future, which defined sustainable development as development that meets the needs of the present without compromising the ability of future generations to meet their own needs. This single sentence did two things that continue to structure every framework reviewed in this monograph. It introduced intergenerational equity as a first-order evaluative criterion, requiring that any technology, program, or action assessed today be evaluated not only for its effects on people currently living but for the constraints it places on the choices available to people not yet born — a requirement no classical TA framework had incorporated. And it explicitly refused to resolve the tension between development and environmental preservation by subordinating one to the other, instead defining sustainability as the search for a compatible path between them — a refusal that anticipates, by three decades, PHVA's own explicit refusal to collapse multiple evaluative domains into a single number before their trade-offs have been made visible (Chapter 5).

The Brundtland definition was rapidly operationalized. John Elkington's 1994 formulation of the Triple Bottom Line (TBL) — the proposition that organizations should be evaluated against three simultaneous bottom lines, economic (profit), environmental (planet), and social (people) — gave sustainability thinking a structure that corporate reporting, government procurement, and academic assessment methodology could all adopt without requiring a single composite metric. The TBL's enduring contribution is not any specific indicator set but the structural claim that three qualitatively different kinds of value cannot be netted against one another without loss of information: a technology that generates strong financial returns while degrading environmental and social capital is not, on a TBL reading, unambiguously successful merely because its economic bottom line is positive. This structural insight — that heterogeneous value dimensions should be reported alongside one another rather than merged — is the direct ancestor of the "mandatory dashboard, optional composite" design principle that both Doughnut Economics (Section 2.4) and PHVA (Chapter 6) would later adopt explicitly.

1.3 Planetary Boundaries and the Biophysical Ceiling

Where the Brundtland Commission and the Triple Bottom Line established that environmental and social considerations belong inside the evaluative frame, neither specified how much environmental pressure the planet could safely absorb. Rockström et al. (2009), writing in

Nature, closed this gap by proposing the planetary boundaries framework: nine Earth-system processes — climate change, biosphere integrity, biogeochemical (nitrogen and phosphorus) flows, ocean acidification, land-system change, freshwater use, stratospheric ozone depletion, atmospheric aerosol loading, and the introduction of novel entities such as synthetic chemicals and plastics — within which humanity has, in the authors' phrase, a safe operating space. The framework was updated by Steffen et al. (2015), and, critically, Richardson et al. (2023) provided the first simultaneous quantification of all nine boundaries, finding that six of the nine had already been transgressed, with the degree of transgression increasing across every boundary previously identified as crossed. This is not a peripheral academic detail: it converts planetary boundaries from a conceptual framework into a live, quantitatively monitored ceiling that any environmental evaluation criterion can be referenced against, rather than an arbitrary or self-defined benchmark. Every framework discussed in Chapter 2 that touches environmental effectiveness — life-cycle assessment's impact categories, environmental impact assessment's air- and water-quality standards, and PHVA's own Domain 1 (Chapter 5) — ultimately either measures against, or could in principle be referenced against, some subset of these nine boundaries.

1.4 Why Historical Evolution Matters to the Present Argument

The three developments traced in this chapter — the recognition that technology assessment must anticipate second-order consequences, the Brundtland insight that intergenerational equity and multi-dimensional value cannot be reduced to a single economic criterion, and the planetary boundaries framework's provision of a scientifically anchored environmental ceiling — together define the minimum requirements that any credible successor evaluation framework must satisfy. A framework that ignores second-order consequences repeats classical TA's original error. A framework that forces heterogeneous value dimensions into one number before establishing whether that aggregation is defensible repeats the error the Triple Bottom Line was designed to correct. And a framework that treats environmental effectiveness as self-defined, rather than referenced against an external, updated scientific ceiling, forfeits the single most important methodological advance planetary boundaries science has contributed. Chapter 2 now turns to the family of frameworks that were built, across five decades, to satisfy these requirements within particular sectors — and to the reasons, developed cumulatively across that chapter and made explicit in Chapter 3, that none of them was designed to satisfy all three requirements simultaneously across an unrestricted range of human activity.

Chapter 2: Existing Sustainability Assessment Frameworks

The frameworks reviewed in this chapter were not designed as competitors. Each emerged to answer a specific evaluative question that classical technology assessment could not answer, and each has become, over several decades, the accepted disciplinary standard for that specific question. This chapter treats each framework in turn, giving its historical origin, theoretical basis, methodology and quantification logic, principal equations, a real-world application, and a candid statement of what it does well and where it fails — before Chapter 3 asks what happens when these frameworks are placed side by side and Chapter 4 asks what remains unanswered even after all of them have been applied together.

2.1 Life-Cycle Assessment (LCA)

Historical Origin and Purpose

Life-cycle assessment traces its origin to a 1969 study commissioned by the Coca-Cola Company comparing the environmental burdens of different beverage-container materials — the first documented Resource and Environmental Profile Analysis, and the historical template later formalized into the ISO 14040 (2006) and ISO 14044 (2006) standards that govern LCA practice today (Hunt & Franklin, 1996). LCA's purpose is to quantify the environmental impacts associated with every stage of a product's or service's life, from raw-material extraction through production, use, and end-of-life disposal, on the premise that no life-cycle stage exists in isolation and that burden shifted from one stage to another is not genuine improvement.

Theoretical Basis and Methodology

LCA is standardized around four mandatory phases. Goal and scope definition — the most consequential phase, because every later result is conditioned on the choices made here — specifies the functional unit, the system boundary (cradle-to-grave, cradle-to-gate, or gate-to-gate), the allocation rule for multi-output processes, and cut-off criteria for excluding minor flows. Life-cycle inventory (LCI) collects quantitative data on every input and output crossing the system boundary, typically drawing on secondary databases such as ecoinvent when primary process data cannot be collected directly (Frischknecht et al., 2005). Life-cycle impact assessment (LCIA) translates this inventory into impact-category scores through classification, characterization, optional normalization, and optional, explicitly value-laden weighting. Interpretation closes the cycle by identifying hotspots, testing sensitivity to key assumptions, and drawing conclusions relevant to the original goal; Finnveden et al. (2009) demonstrate that value choices made during interpretation can reverse which of two compared alternatives appears more sustainable, which is why mandatory sensitivity and consistency checks are now standard practice.

Quantification

The core life-cycle impact-assessment equation is:

S_k = sum over i of (Qi times CF_ik)

where S_k is the impact score for category k, Qi is the quantity of elementary flow i, and CF_ik is the characterization factor translating a unit of flow i into its contribution to impact category k. ReCiPe2016, the current harmonized successor to the earlier CML and Eco-indicator 99 methods, supplies characterization factors for seventeen midpoint impact categories and converts these into endpoint damage categories (human health, ecosystems, resource scarcity) (Huijbregts et al., 2017).

Impact Category	**Elementary Flows**	**Quantification Model**	**Unit**	**Representative Characterization Factors**
Global Warming Potential (GWP)	CO_2, CH_4, N_2O, HFCs, SF_6 to air	IPCC AR6 100-yr GWP factors	kg CO_2-eq	GWP(CO_2)=1; GWP(CH_4)=27.9; GWP(N_2O)=273
Acidification (AP)	SO_2, NO_x, NH_3, HCl, HF to air	Fate-exposure factors (CML-2002, ReCiPe)	kg SO_2-eq	AP(SO_2)=1; AP(NO_x)=0.5; AP(NH_3)=1.6
Eutrophication (EP)	NO_x, NH_4^+, NO_3^-, PO_4^{3-}, BOD to water/air	Stoichiometric N/P ratios	kg PO_4-eq or kg N-eq	EP(PO_4)=1; EP(NO_x)=0.13; EP(NH_4)=0.33
Human Toxicity (HTP)	Heavy metals, organics to air/water/soil	USEtox fate x exposure x effect	CTUh (cases/kg emitted)	Ingestion and inhalation pathways integrated
Freshwater Ecotoxicity	Metals, organics to freshwater	USEtox aquatic compartment model	CTUe (PAF.m3.day/kg)	Potentially Affected Fraction of species
Land Use	Land occupation and transformation (m2.yr)	Species-loss (PDF) models	PDF.m2.yr	Probability of Disappeared Fraction of species
Water Scarcity	Water withdrawn by region (m3)	AWARE / Pfister scarcity factors	m3-eq or m3 deprivation	Scarcity CF varies 0.001-100 by basin

Impact Category	Elementary Flows	Quantification Model	Unit	Representative Characterization Factors
Resource Depletion	Fossil energy (MJ), mineral mass (kg)	Abiotic Depletion Potential	kg Sb-eq (minerals); MJ (energy)	ADP based on reserve base or EDIP

Table 2.1. LCA impact categories and their quantification models (adapted from ReCiPe2016 and CML characterization schemes).

Real-World Example

The founding 1969 Coca-Cola container study illustrates the decisiveness of the goal-and-scope phase: the scope choices made directly changed which container material appeared environmentally superior, establishing the template later formalized into ISO 14040/14044's mandatory goal-and-scope requirement.

Advantages, Limitations, and When It Fails

LCA's advantages are standardization, quantitative rigor, and broad database support. Its limitations are equally well documented: normalization and weighting are optional and value-laden; system-boundary and allocation choices can reverse conclusions (Finnveden et al., 2009); and, most relevant to this monograph's argument, LCA has no native capacity to represent equity, cultural value, knowledge contribution, or resilience. Social LCA (S-LCA) and Life Cycle Costing (LCC) were developed specifically to extend LCA into the social and economic dimensions the Triple Bottom Line requires; their combination was formally proposed by Kloepffer (2008) as Life Cycle Sustainability Assessment, LCSA = LCA + S-LCA + LCC, on the explicit argument that environmental LCA alone is insufficient evidence for a sustainability claim. LCSA is, at present, the most comprehensive life-cycle-based sustainability tool in routine use, but it remains a product- and service-level methodology; it does not extend to education, culture, or research contribution, and its social pillar, formalized by the UNEP/SETAC (2009) guidelines around five stakeholder categories (workers, local community, society, consumers, value-chain actors), is considerably less standardized than its environmental counterpart.

2.2 Environmental Impact Assessment (EIA)

Historical Origin and Purpose

EIA was first institutionalized through the US National Environmental Policy Act of 1969, and was adopted in India through the Environment (Protection) Act, 1986, becoming mandatory for large projects under the EIA Notification of 1994, subsequently revised in 2006. Where LCA evaluates a product across its full life cycle irrespective of location, EIA evaluates a specific, sited project against the environmental baseline of the place where it will actually be built, incorporating socioeconomic, cultural, and human-health impacts alongside strictly ecological ones.

Methodology

The EIA process proceeds through screening, scoping, baseline study, impact prediction, mitigation, compilation into an Environmental Impact Statement, independent expert review, a regulatory decision, and post-approval monitoring.

Quantification

Air-quality impact prediction standardly relies on the Gaussian plume model for point sources under steady-state atmospheric conditions:

C(x,y,z) = Q / (2 pi sigma_y sigma_z u) times exp(-y^2 / 2 sigma_y^2) times [exp(-(z-H)^2 / 2 sigma_z^2) + exp(-(z+H)^2 / 2 sigma_z^2)]

where C is ground-level concentration, Q is the emission rate, sigma_y and sigma_z are dispersion coefficients dependent on downwind distance and Pasquill-Gifford atmospheric stability class, u is mean wind speed at stack height, and H is effective stack height. Water-quality prediction typically applies the Streeter-Phelps dissolved-oxygen sag equation, D(t) = [k_d L0 / (k_r minus k_d)] times (e^(-k_d t) minus e^(-k_r t)) plus D0 times e^(-k_r t), yielding the critical minimum dissolved-oxygen point compared against fisheries and drinking-water standards. Noise-impact prediction applies point-source spherical-spreading attenuation, L_eq(r) = L_eq(r0) minus 20 log10(r/r0) minus alpha(r minus r0). Ecological-impact assessment applies the Importance Value Index, IVI = Relative Density + Relative Frequency + Relative Dominance, ranging 0 to 300, together with Shannon diversity, H' = negative sum of (pi times ln pi), and the Habitat Suitability Index, HSI = geometric mean of habitat-suitability variables on a 0 to 1 scale, underlying the Habitat Unit currency used in wetland mitigation banking.

Real-World Example

A proposed NH-48 highway expansion in India illustrates the full chain: predicted construction-phase PM10 emissions of 6,660 kg per day yielded a Gaussian-model maximum ground-level concentration of approximately 340 micrograms per cubic metre at 200 metres downwind under stability class D, roughly four times the 24-hour NAAQS limit of 100 micrograms per cubic metre, triggering mandatory water-spraying and covered-vehicle mitigation; predicted construction noise attenuated to 45.5 dB-A at 100 metres, barely meeting the nocturnal residential limit; and IVI/Shannon-diversity analysis of the affected 80-hectare forest patch, with four Schedule I species recorded, triggered a mandatory 1:2 compensatory-afforestation requirement. The Sardar Sarovar (Narmada) Dam remains India's most-studied EIA case for the full procedural chain: the Morse Committee's Independent Review (1992) documented that resettlement commitments made at the mitigation stage were not subsequently met, a finding that led directly to the World Bank's withdrawal of project funding in 1993.

Advantages, Limitations, and When It Fails

EIA's strength is its integration of physically grounded predictive models with a legally mandated procedural sequence that forces mitigation commitments and independent review before project approval. Its principal weakness, extensively documented in Indian practice, is that project-proponent-funded EIAs create a structural incentive to underestimate impact significance, a regulatory-capture problem documented by Kohli et al. (2011) and the Comptroller and Auditor General's 2011 audit of Indian EIAs, and that monetary valuation is rarely performed systematically even though the underlying toolkit (Value of Statistical Life, contingent valuation, replacement cost, travel cost methods) exists. EIA's biodiversity assessment is site-specific and descriptive rather than globally comparable, in contrast to LCA's standardized, if coarser, metric.

2.3 Social Impact Assessment (SIA)

Historical Origin and Purpose

SIA developed alongside EIA in the 1970s from the recognition that development projects affect communities and social systems in ways environmental assessment cannot capture. Vanclay's (2003) International Principles for Social Impact Assessment, codified over five years of International Association for Impact Assessment consultation across six continents, define SIA as the process of analyzing, monitoring, and managing the intended and unintended social consequences of planned interventions, and establish that social impacts must be considered throughout project design, not assessed only after the fact.

Methodology and Quantification

SIA collects data through key-informant interviews, focus-group discussions, Participatory Rural Appraisal (Chambers, 1994), social network analysis, and Delphi expert-consensus rounds (Dalkey & Helmer, 1963), organized around indicator domains including livelihood disruption, health impact, food security, social cohesion, cultural impact, gender equity, access to services, and psychological stress.

Social Domain	Indicator	Data Collection Method	Quantified Output
Livelihood disruption	% households losing >25% income; # families displaced	Household survey (>=10% affected population)	% change in household income; Gini coefficient change
Health impact	Disease burden change (DALYs); mortality/morbidity rates	Clinical records; model-based projections	DALYs attributable to project; adjusted odds ratio
Food security	HFIAS score (0-27); dietary diversity score	Validated 30-day recall survey	Prevalence of food insecurity (%)
Social cohesion	Putnam social capital index; participation rates	Trust surveys; institutional density mapping	Index score (0-100)
Gender equity	Gender Parity Index; Female Labour Force Participation	Sex-disaggregated survey	GPI change; FLFP % change

Table 2.2. SIA indicator domains and quantification methods (Vanclay, 2003; World Commission on Dams, 2000).

When SIA incorporates health-impact quantification, as increasingly required under WHO Environmental, Social, and Health Impact Assessment guidelines, it applies the standard Disability-Adjusted Life Year framework, DALY = YLL + YLD, where Years of Life Lost is YLL = N times L, and Years Lost to Disability is YLD = I times DW times D, incidence multiplied by a standardized disability weight multiplied by duration.

Real-World Example

An adapted analysis of Sardar Sarovar Dam resettlement across 45,000 project-affected families illustrates SIA's quantitative range: average household income fell from an estimated baseline of INR 42,000 to INR 28,500 post-resettlement, an aggregate annual loss of approximately INR 607.5 crore; the food-insecurity access scale score rose from 4.2 to 11.7 within three years of displacement; and waterborne-disease and malaria-transmission increases were quantified as excess DALYs and an adjusted odds ratio of 3.4 for malaria, directly attributable to project-induced environmental change.

Advantages, Limitations, and When It Fails

SIA's distinctive strength is that it measures what interventions mean to people rather than what people do to ecosystems, which is LCA's and EIA's domain; the two are complementary, not substitutable. Its central limitation is that aggregate cost-benefit framing is structurally inadequate for its own subject matter: a Kaldor-Hicks efficiency criterion is insufficient when the people bearing the costs are the most vulnerable and receive inadequate compensation, because SIA's evaluative purpose is distributional, not aggregative. Monetization approaches available to SIA are each contested (Social Return on Investment's proxy values in particular are widely criticized as arbitrary), and none resolves the underlying tension between aggregation and distributional justice.

2.4 Industrial Ecology (IE)

Historical Origin and Purpose

Industrial Ecology emerged as a formal discipline with Frosch and Gallopoulos's 1989 Scientific American paper, which proposed that industrial systems should be designed so that waste from one process becomes input for another, mirroring closed-loop material cycling in natural ecosystems.

Methodology and Quantification

Material Flow Analysis (MFA) tracks material movement through a defined system over a defined period, implementing physical mass conservation: dS/dt = sum of imports minus sum of exports at national scale, with domestic material consumption DMC = Domestic Extraction plus Imports minus Exports, and material productivity defined as GDP divided by DMC. Substance Flow Analysis (SFA) narrows this to a single substance — for example, phosphorus in Indian agriculture, where P_balance(soil) = P_fertiliser + P_manure + P_atmospheric minus P_crop_harvest minus P_runoff minus P_leaching; a positive balance drives progressive soil phosphorus saturation, increased runoff, eutrophication, and downstream human-health risk, a complete substance-to-ecosystem-to-health causal chain of exactly the kind One Health surveillance is designed to monitor. Environmentally Extended Input-Output Analysis (EE-IOA) extends the Leontief (1941) input-output model with environmental intensity matrices: x = (I minus A) inverse times f for total sectoral output, and e = F-hat times (I minus A) inverse times f for the associated environmental burden vector, where A is the technical coefficient matrix, f is final demand, and F-hat is a diagonalized matrix of per-sector environmental intensities.

Real-World Example

A phosphorus SFA for the Indian Punjab wheat-rice-dairy system estimates a surplus of approximately +82 kg P per hectare per year; the resulting soil phosphorus saturation is linked, via canal-system transport to the Harike Wetland Ramsar site, to a documented threefold increase in algal-bloom frequency since 1990, degraded fisheries affecting an estimated 4,200 fisher livelihoods, and microcystin contamination of drinking water — tracing a single substance from agricultural input through ecosystem damage to human-health impact within one coherent quantitative chain.

Advantages, Limitations, and When It Fails

IE's distinctive strength is its capacity to trace a causal mechanism connecting industrial or agricultural inputs to ecosystem degradation to human-health outcomes, a mechanistic linkage most other frameworks assert rather than trace. Its principal limitation is that IE lacks standardized biodiversity characterization factors analogous to LCA's; the connection between a quantified material-flow surplus and its species-level biodiversity consequence remains conceptual rather than systematically calculable, and IE has no native equity or cultural dimension.

2.5 Resilience Assessment

Historical Origin and Purpose

Ecological resilience was first formally defined by Holling (1973) as the ability of a system to absorb disturbance and reorganize while retaining essentially the same function, structure, and identity — a definition subsequently extended, through Gunderson and Holling's (2002) Panarchy framework, from ecology into social, economic, and institutional systems generally.

Core Concepts

The adaptive cycle describes systems moving through four recurring phases: exploitation (r), rapid growth and resource accumulation; conservation (K), stabilization that increasingly reduces adaptive flexibility; release (Omega), collapse triggered by an accumulating or external disturbance; and reorganization (alpha), during which the system reassembles, often incorporating new structure or knowledge. Panarchy theory extends this by observing that resilience operates simultaneously across nested scales, such that a disease outbreak at farm level is shaped by veterinary policy at national level and international guidance at global level. Bene et al. (2012) operationalize resilience into three empirically distinguishable capacities: absorptive capacity, the ability to withstand a shock without losing function; adaptive capacity, the ability to adjust strategies during or after a shock; and transformative capacity, the ability to restructure fundamentally so as to prevent recurrence. The WHO's post-Ebola health-system resilience agenda (Kruk et al., 2015) operationalizes these into five system functions: aware, diverse, self-regulating, integrated, and adaptive.

Quantification

Three distinct quantification traditions coexist. Engineering resilience (Pimm, 1984) measures how quickly a system returns to equilibrium after a small perturbation: R_engineering = 1 divided by Return_time, empirically estimated by fitting exponential decay to post-perturbation time-series data and extracting the dominant eigenvalue of the system's Jacobian matrix at equilibrium, a measure that assumes linear dynamics near equilibrium and fails for systems with alternative stable states. Ecological resilience instead measures the magnitude of disturbance a system can absorb before shifting to an alternative basin of attraction, operationalized through bifurcation analysis and early-warning indicators: variance in the state variable rises toward infinity, and the lag-1 autocorrelation coefficient AR(1) rises toward 1, as the system approaches a critical threshold, a phenomenon termed critical slowing down (Dakos et al., 2012). For complex social-ecological systems, composite indicator approaches are used instead: the UNDRR Sendai Framework indicators, and, for ecosystems specifically, the Mean Species Abundance index, MSA = average across species of (current population divided by pristine population), where MSA of 1 indicates a pristine baseline and MSA of 0 indicates local extinction.

Real-World Example

Applied to a Rajasthan/Gujarat transboundary pastoral system facing co-occurring brucellosis epizootic and drought risk, a resilience indicator can be constructed directly from the known herd-decapitalization threshold: RI = (Seroprevalence_threshold minus Seroprevalence_current) divided by Seroprevalence_threshold. At an observed current seroprevalence of 18 percent against a documented 25 percent decapitalization threshold, RI = (25 minus 18) divided by 25 = 0.28, indicating the system sits 28 percent of its threshold range from a forced-distress-sale regime shift; a vaccination campaign targeting seroprevalence below 10 percent would raise RI to 0.60, more than doubling the resilience buffer.

Monetization and Limitations

Resilience is inherently difficult to monetize because it is a system property that prevents future losses rather than generating present flows; the standard counterfactual logic is Value of resilience = Expected cost of regime shift multiplied by probability of regime shift without investment, minus the same expected cost multiplied by probability of regime shift with investment. Walker et al. (2009) apply this logic to the Goulburn-Broken Catchment in Australia, estimating resilience value at AUD 200 to 800 million over twenty years while explicitly flagging two-to-three orders of magnitude of uncertainty. Resilience assessment's core limitation for this monograph's purposes is that it measures adaptive capacity and threshold proximity extremely well but has no native mechanism for representing environmental footprint, distributional equity, or knowledge and cultural contribution.

2.6 One Health

One Health is the single most institutionally mature cross-sectoral health framework in existence. The One Health High-Level Expert Panel's consensus definition, adopted by the Quadripartite (the Food and Agriculture Organization, the UN Environment Programme, the World Health Organization, and the World Organisation for Animal Health, formalized in March 2022) in December 2021, defines One Health as an integrated, unifying approach that aims to sustainably balance and optimize the health of people, animals, and ecosystems, recognizing their close interdependence (OHHLEP, 2021). Operationally, One Health is structured around the One Health Joint Plan of Action 2022–2026, organized into six action tracks centred on zoonotic disease, antimicrobial resistance, food safety, and related health-system capacities (Quadripartite, 2022). One Health's strength is institutional: a standing scientific advisory panel, dedicated agency secretariats, and country-level implementation guides give it operational reach that few other frameworks in this chapter possess. Its limitation, for the purposes of the present argument, is scope: One Health was not designed, and does not claim, to function as a general-purpose evaluation tool for arbitrary human products, services, or cultural activities. It answers the question of how human, animal, and ecosystem health interact at the disease interface; it does not answer questions about the value of a documentary film, a teaching curriculum, or an industrial process unconnected to zoonotic or food-safety pathways.

2.7 Planetary Health

Planetary health is a related but distinct field, defined by the 2015 Rockefeller Foundation–Lancet Commission as the health of human civilization and the state of the natural systems on which it depends (Whitmee et al., 2015), a term that originated in a 2014 Lancet manifesto and was subsequently institutionalized through the Planetary Health Alliance. Where One Health centres on the human-animal-ecosystem disease interface, planetary health takes a civilizational-scale view that more readily accommodates non-infectious dimensions — climate, food systems, and, most relevant here, the cultural and knowledge dimensions of human activity. Planetary health is best understood as a conceptual field and social movement rather than a scoring tool: it supplies the definitional object of concern (the health of civilization and its supporting natural systems) without itself specifying a measurement methodology, which is precisely the gap that motivates the methodological architecture developed in Chapters 5 through 7.

2.8 Planetary Boundaries as Reference Standard

As established in Chapter 1, the planetary boundaries framework (Rockström et al., 2009; Steffen et al., 2015; Richardson et al., 2023) provides a scientifically anchored, quantitatively monitored environmental ceiling across nine Earth-system processes, six of which are now documented as transgressed. Its distinctive methodological contribution, beyond the substantive finding itself, is that it supplies an external, non-arbitrary reference standard against which any framework's environmental-effectiveness criterion can be validated, rather than requiring each framework to define its own environmental benchmark from scratch.

2.9 Doughnut Economics

Kate Raworth's Doughnut framework (Raworth, 2012, 2017) combines the planetary boundaries' ecological ceiling with a social foundation of twelve dimensions derived from the UN Sustainable Development Goals — food, water, health, education, income, political voice, and others — defining a safe and just space for humanity as the area between an ecological ceiling that must not be overshot and a social floor below which no one should fall. Empirical applications have consistently found that no country currently meets its social foundation for all citizens without exceeding its share of planetary boundaries. The Doughnut's key methodological contribution is structural rather than substantive: it demonstrates that a ceiling-and-floor dual-boundary representation can communicate a genuinely complex, multi-dimensional reality — environmental limits and social adequacy simultaneously — without forcing the two into a single number, and without requiring the analyst to specify a trade-off rate between an additional unit of ecological pressure and an additional unit of social deprivation. This ceiling/floor logic, more than any specific indicator in the Doughnut's twelve-dimension social foundation, is the design principle later frameworks in this monograph borrow most directly.

2.10 Health Technology Assessment and the QALY

Health Technology Assessment (HTA) is the field the later chapters of this monograph borrow their evaluative logic from most directly. The consensus definition jointly developed by the International Network of Agencies for Health Technology Assessment (INAHTA) and Health Technology Assessment International, published by O'Rourke, Oortwijn, and Schuller (2020), states that HTA is a multidisciplinary process that uses explicit methods to determine the value of a health technology at different points in its lifecycle, in order to promote an equitable, efficient, and high-quality health system. INAHTA itself comprises 54 publicly funded HTA agencies across 33 countries, demonstrating that a structured, multidisciplinary, lifecycle-oriented value-determination process of this kind is institutionally workable at national scale.

HTA's standard outcome metric, the Quality-Adjusted Life Year (QALY), combines length and quality of life into a single number used in cost-per-QALY calculations, and is simultaneously the most thoroughly documented cautionary tale available for any framework that aspires to a single aggregate value metric. A substantial bioethics literature demonstrates that the QALY systematically discriminates against people with disabilities and chronic conditions, because extending the life of someone with a lower baseline quality-of-life score mechanically yields fewer QALYs than extending the life of someone in better health, irrespective of how each person values their own life (Sinclair, 2012, 2018). Schneider (2022) labels the QALY ableist on precisely these grounds. The practical consequence has been severe: United States federal law now prohibits the use of QALY-based cost-effectiveness thresholds in Medicare coverage decisions, and disability-rights organizations have campaigned for an equivalent ban across all federal programs (National Center for Disability, Equity, and Intersectionality, 2021). This history functions, throughout the remainder of this monograph, not as a reason to avoid quantification altogether but as a binding design constraint on any future composite metric — the precise function it serves in Chapter 6's treatment of a possible headline indicator.

2.11 Multi-Criteria Decision Analysis and EVIDEM

In direct response to dissatisfaction with single-ratio cost-effectiveness analysis, health economics developed multi-criteria decision analysis (MCDA) as a complementary approach. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) established an MCDA Emerging Good Practices Task Force in 2014, publishing two widely cited reports establishing best practice for structuring criteria, eliciting weights, and reporting uncertainty (ISPOR MCDA Task Force, Reports 1 and 2, 2016). The most developed operational example is EVIDEM (Evidence and Value: Impact on Decision Making), which defines a standard set of decision criteria, requires structured by-criterion evidence synthesis for each, and combines them through an explicit, deliberative weighting process rather than a fixed formula; EVIDEM has been field-tested by a public payer in Canada and a reimbursement committee in South Africa (Goetghebeur et al., 2011). Independent reviews note that MCDA/EVIDEM models still face real challenges: inconsistent treatment of uncertainty, and difficulty for non-specialist committees to interpret an aggregate score, a limitation documented in a 2018 systematic review in Frontiers in Public Health. EVIDEM's core discipline — structure the criteria, synthesize evidence per criterion, make weighting explicit and contestable, and resist premature aggregation — constitutes the single most important methodological inheritance for the tiered methodology developed in Chapter 6.

2.12 Natural Capital and Ecosystem-Service Valuation

Three established toolkits quantify environmental and natural-capital dimensions with internationally agreed methods. Life-cycle assessment, already reviewed in Section 2.1, provides the product-level methodology. Ecosystem-service valuation, beginning with Costanza et al.'s (1997) widely cited global estimate and updated by Costanza et al. (2014) using land-use data through 2011 and the TEEB valuation database, expresses the economic value healthy ecosystems provide, while its authors are explicit that such valuation is intended to inform decisions rather than to commodify nature. The Natural Capital Protocol, developed collaboratively by 38 organizations and released by the Capitals Coalition in 2016 after a public consultation drawing over 3,200 comments, provides a standardized decision-making framework for organizations to identify, measure, and value their dependencies and impacts on natural capital. IPBES's Nature's Contributions to People framework extends the Millennium Ecosystem Assessment's ecosystem-services concept to explicitly include both beneficial and detrimental contributions of nature and to accommodate multiple value systems, including Indigenous and local knowledge — a scope that goes beyond conventional ecosystem-service valuation but that, as Chapter 8 documents, has not yet been matched by institutional representation of Indigenous knowledge-holders in the governance structures downstream frameworks propose.

2.13 Composite Sustainability Indices

A distinct family of frameworks — the Genuine Progress Indicator, the Inclusive Wealth Index, the Happy Planet Index, and related national-level composite indices — has, since the 1990s, attempted to compress economic, environmental, and social welfare into a single aggregate number, generally as an explicit alternative to GDP. Böhringer and Jochem (2007), surveying this literature in Ecological Economics, identify a recurring structural failure: aggregating heterogeneous value dimensions into one composite number requires implicit value judgments about what trades off against what, judgments that are ethical rather than technical, and that composite-index construction typically buries rather than exposes. This composite-index critique is not a peripheral methodological footnote; it is the single most important reason the frameworks developed later in this monograph (Doughnut Economics, EVIDEM, and, centrally, PHVA in Chapter 6) each explicitly reject a mandatory single-number output in favor of a disaggregated dashboard with an optional, clearly caveated composite.

2.14 The Nascent Planetary Health Impact Assessment Literature

A further complication for any newly proposed framework in this space is that the term Planetary Health Impact Assessment (PHIA) already exists in the peer-reviewed literature. Stefanopoulou et al. (2025), published in Bioelectromagnetics, propose a PHIA approach built on expert-elicited knowledge graphs and AI-assisted evidence synthesis to trace direct and ecologically mediated health pathways, illustrated through a radiofrequency electromagnetic-field exposure case study. In parallel, a postdoctoral research line within the CINPHIA consortium at Wageningen University and Research is explicitly developing a Planetary Health Impact Assessment compass intended for Dutch policy practice. These efforts are methodologically sophisticated but are currently scoped to specific environmental-exposure-to-health pathways embedded in policy practice, rather than to the general, cross-sectoral evaluation of arbitrary products, services, educational programs, and cultural contributions that motivates Chapters 4 through 6 of this monograph.

Framework	Primary Focus	Typical Scale	Aggregation Stance
One Health (Quadripartite/OHHLEP)	Human-animal-ecosystem disease interface; AMR; food safety	National / global policy	No composite metric; coordinated action tracks
Planetary Health (Lancet Commission)	Civilizational health vs. state of natural systems	Global / civilizational	Conceptual field, not a scoring tool
Planetary Boundaries	Nine Earth-system limits	Global / regional	Nine separate boundary variables, not merged
Doughnut Economics	Ecological ceiling + social foundation	National / city	Dual-boundary visual; no single score
HTA / QALY	Value of a health technology over its lifecycle	Health system / payer	Optional single ratio (cost per QALY); contested
MCDA / EVIDEM	Structured multi-criteria health-value judgments	Health system / payer	Explicit weighted score; uncertainty reported
LCA (ISO 14040/44)	Cradle-to-grave environmental footprint	Product / service	Multiple impact categories; review required for claims
Natural Capital Protocol	Organizational dependencies/impacts on nature	Organization / sector	Multiple capital accounts; valuation optional
Composite indices (GPI, IWI, HPI)	Aggregate welfare or sustainability	National	Single index; widely critiqued on commensurability
PHIA (Stefanopoulou et al., 2025; CINPHIA)	Specific environmental-exposure health pathways	Case study / national policy	Knowledge-graph evidence synthesis, not a score

Table 2.3. Positioning of the frameworks reviewed in this chapter, adapted from the comparative table developed in the PHVA proposal.

Chapter 3: Comparative Critical Analysis

Placed side by side, the frameworks reviewed in Chapter 2 divide cleanly along two axes that matter for everything that follows: the dimension of value each is built to measure, and the scale and object of analysis each assumes. LCA and Industrial Ecology measure what a product, process, or material system does to the physical environment, through mass- and energy-balance accounting that is agnostic to who is affected and why it matters to them. Environmental Impact Assessment measures the same broad territory but anchors it to a specific site and a specific project, adding legally mandated procedural discipline that LCA, as a comparative product-level tool, does not require. Social Impact Assessment measures what an intervention means to the people affected by it — livelihoods, food security, cultural attachment, psychological wellbeing — a genuinely distinct evaluative object from what LCA and EIA measure, even when both are applied to the same physical project. Resilience Assessment measures a system property, adaptive capacity and threshold proximity, that is largely orthogonal to both the environmental-footprint and the social-consequence questions: a project can have a modest environmental footprint and a well-compensated social-impact profile while still leaving the community or ecosystem it touches less able to absorb the next shock. One Health and Planetary Boundaries anchor a specific subset of these concerns — zoonotic and ecosystem-health interfaces, and the biophysical ceiling respectively — to institutionally mature, internationally coordinated monitoring infrastructure. HTA and MCDA/EVIDEM contribute the disciplined, multidisciplinary, evidence-graded evaluation logic that none of the environmental or social frameworks above possesses natively, but they operate within a single, comparatively well-bounded sector — health technologies — where the criteria set, while multi-dimensional, is far narrower than the eight-domain architecture Chapter 6 will describe. Doughnut Economics and the composite-index literature contribute, respectively, a structural solution to premature aggregation (the ceiling/floor dashboard) and a cautionary catalogue of what happens when that structural solution is not adopted.

The overlaps and conflicts among these frameworks are as instructive as their coverage gaps. LCA and Industrial Ecology overlap substantially in their environmental accounting logic — both trace material and energy flows against physically grounded balance equations — but IE's substance-flow tracing offers causal-mechanism detail (input to soil to water to ecosystem to human health) that LCA's characterization-factor approach, being a static multiplier rather than a dynamic model, does not provide; conversely, LCA's characterization factors are globally standardized and comparable across studies in a way IE's case-specific substance-flow models are not. LCA and EIA overlap in environmental impact prediction but conflict in scope and comparability: LCA compares product systems irrespective of location, while EIA evaluates one sited project against its local baseline, meaning the same technology can score well under an LCA comparison (low global-average footprint) while triggering serious local EIA findings (high local exceedance of an air- or water-quality standard) — a tension neither framework alone resolves, because each is answering a different, equally legitimate question. SIA and the equity dimension of Doughnut Economics overlap conceptually (both foreground distributional consequence) but SIA operates at project scale with primary survey data, while Doughnut's social-foundation indicators operate at national or city scale with secondary Sustainable Development Goal data — the two are complementary at different levels of aggregation rather than substitutable. HTA/QALY and the composite-index literature are in direct methodological tension: QALY is precisely the kind of single aggregate metric Böhringer and Jochem's (2007) composite-index critique warns against, and the QALY's own documented disability-discrimination history (Sinclair, 2012, 2018; Schneider, 2022) is the most concrete, highest-stakes illustration available of that critique's practical consequences.

Across economic, environmental, social, cultural, knowledge, equity, resilience, and planetary-health dimensions, the pattern that emerges is consistent: each framework achieves real rigor within its designated dimension, and each framework's rigor is purchased specifically by narrowing its scope to that dimension. LCA's characterization factors are rigorous because LCA does not attempt to represent equity. EVIDEM's structured weighting is workable because EVIDEM operates within health technology's comparatively bounded criteria set, not across the full breadth of human activity. The QALY is quantitatively tractable because it compresses two dimensions — length and quality of life — that are at least conceptually commensurable in a way that, for example, carbon emissions and cultural cohesion plainly are not. No framework reviewed in this chapter was designed, or claims, to hold economic, environmental, social, cultural, knowledge, equity, resilience, and planetary-health value simultaneously in view for an arbitrary human action. This is not a failure of any individual framework; each does precisely what it was built to do. It is, however, precisely the condition that produces the measurement gap Chapter 4 now develops in detail: a government, university, or funder comparing options that span several of these dimensions at once has no framework that was built for that comparison, and is left either applying several incompatible frameworks in parallel and reconciling their outputs informally, or defaulting to whichever single dimension (typically the economic one) happens to have the most tractable existing metric.

Evaluative Dimension	LCA	EIA	SIA	Industrial Ecology	Resilience Assessment	HTA/QALY	MCDA/EVII
Environmental footprint	Strong (ISO characterization factors)	Strong (site-specific prediction)	Absent (indirect only)	Strong (substance-flow tracing)	Weak (not primary focus)	Absent	Absent (impo LCA where u
Economic cost	Via LCC extension only	Weak in Indian practice	Weak (contested monetization)	Strong (material/EE-IOA costing)	Weak (high uncertainty)	Strong (cost-per-QALY)	Strong (explic weighting)
Social/distributional equity	Absent (S-LCA extension only)	Weak (HIA/SIA add-on)	Strong (primary purpose)	Absent	Weak (indirect via livelihoods)	Weak (documented equity failure)	Depends on criteria set ch
Cultural/knowledge value	Absent	Absent	Weak (cultural-impact indicator only)	Absent	Absent	Absent	Depends on criteria set ch
Resilience/adaptive capacity	Absent	Absent	Absent	Weak (circularity as proxy)	Strong (primary purpose)	Absent	Depends on criteria set ch
Aggregation stance	Multiple categories; review required	No aggregation; procedural sequence	No aggregation; distributional reporting	No aggregation; physical balance	No aggregation; indicator battery	Optional single ratio; contested	Explicit weig score; uncerta reported

Table 3.1. Cross-framework evaluative-dimension matrix. Cells reflect the maturity of each framework's native methodology for the row dimension, not a judgement of overall framework quality; a framework recorded as "Absent" on a given dimension is, by design, not intended to measure it.

Table 3.1 makes the argument of this chapter visible in a single view. Reading down any column shows a framework that is strong on at most two or three of the six dimensions and silent on the rest — not because the framework is deficient, but because each was built to answer one evaluative question rigorously rather than several questions adequately. Reading across any row shows that some evaluative dimensions (economic cost, environmental footprint) are well served by multiple existing instruments, while others (cultural and knowledge value, and to a lesser extent resilience) are served natively by at most one framework each, and in the case of cultural value, only weakly. This asymmetry is itself diagnostic: it is not that sustainability science has failed to develop cultural- or knowledge-value metrics through inattention, but that these dimensions resist the kind of physically or actuarially grounded quantification that makes LCA's characterization factors or HTA's QALY tractable in the first place — a fact Chapter 5 treats as a design constraint rather than a problem to be engineered away.

Chapter 4: The Measurement Gap

Consider four decision-makers, each facing a genuine, unavoidable choice, none of whom is served by any single framework reviewed in Chapter 2. A district health department in a brucellosis-endemic pastoral region must decide between funding a mass vaccination campaign for livestock and funding a community health-worker surveillance program: HTA methodology can, in principle, price the vaccination campaign in cost-per-DALY-averted terms, but has no native mechanism for representing the pastoral system's resilience gain (Section 2.5's RI calculation) or the campaign's contribution to the epidemiological knowledge base that surveillance would build. A university tenure and research-evaluation committee must decide how to weigh a lab-scientist's citation-heavy but narrowly disciplinary publication record against an interdisciplinary colleague's community-engaged research program with modest citation counts but documented policy influence: bibliometric evaluation captures the first case well and the second case not at all, and no framework reviewed in this monograph — not even HTA's own recent shift toward "impact beyond output" thinking — yet has an operational methodology for the comparison. A national agriculture ministry must decide between subsidizing drip-irrigation retrofits (which reduce pressure on a transgressed planetary boundary, freshwater use, but whose uptake, as Section 6.1's later sketch will show, is often concentrated among large landholders with capital to invest) and subsidizing a smallholder-targeted extension program with a smaller direct environmental footprint but a stronger equity profile: LCA can score the first option's environmental performance and SIA can score the second option's distributional performance, but no framework scores both options against the same eight-dimensional canvas so that the equity-versus-environment trade-off is visible to the decision-maker rather than buried inside two separate, non-comparable reports. A national broadcaster must decide whether a climate-themed documentary series justifies its production budget relative to a direct reforestation grant of the same size: the documentary's production emissions are LCA-tractable, but its intended value — shifting public understanding and behavior at scale — sits entirely within Domain 7 of the architecture Chapter 6 will describe, a dimension no framework reviewed in Chapter 2 quantifies, and that Section 2.6's introductory PHVA public explainer material explicitly and deliberately declines to force into a number.

These four cases share a structure. In each, the decision-maker's problem is not a lack of rigorous measurement — HTA, LCA, SIA, and resilience assessment are each, within their domain, mature and well-validated — but the absence of any shared evidentiary architecture that would let outputs from different domains be placed on a common, auditable footing without either (a) forcing a false commensurability that collapses ecological, social, and cultural value into one contestable number, repeating the composite-index failure documented in Section 2.13, or (b) leaving the comparison entirely informal and therefore untraceable, unauditable, and vulnerable to whichever domain's advocates argue most persuasively in the room. The literature on this specific gap is thin but growing: the 2024 scoping-review protocol on incorporating planetary-health considerations into health guidelines and HTAs (referenced in the PHVA proposal's Section 3.2) documents the HTA field's own recognition that its traditional criteria set does not capture planetary-boundary consequences; Stefanopoulou et al.'s (2025) PHIA work documents a parallel recognition within environmental-health science that exposure-pathway analysis needs a broader evaluative frame than any single existing tool supplies; and the persistent recurrence, across research, policy, health, education, and agriculture, of exactly the kind of cross-domain comparison problem illustrated by the four cases above, is itself the strongest available evidence that the gap is real and not merely a rhetorical device introduced to motivate a new framework.

It is worth being precise about what would, and would not, count as closing this gap, because the temptation — visible in the earlier Planet Health Technology Assessment concept reviewed critically in Chapter 8 — is to conflate "closing the gap" with "inventing a single new master number." A genuine solution to the measurement gap does not require that carbon tonnes, DALYs, Gini-coefficient change, and documentary viewership become commensurable in some literal arithmetic sense; Section 2.13's composite-index critique establishes that no amount of methodological sophistication makes that commensurability problem disappear, because the underlying difficulty is an ethical value-judgment problem, not a measurement-precision problem. What a genuine solution requires is a shared architecture — a common set of domains, a common evidentiary discipline for populating each domain, and a common, transparent process for surfacing (not resolving) the trade-offs between domains when a real decision requires comparing options that score differently across them. This is the specific, narrower, and more defensible ambition Chapter 5 now turns to.

Chapter 5: Development of Planetary Health Value Assessment

5.1 Why PHVA Was Conceived, and Why the Name Changed from PHTA

The idea that eventually became Planetary Health Value Assessment first appeared under the name Planet Health Technology Assessment (PHTA), a document that read, on its own explicit terms, as a pitch rather than a scientific proposal: it contained no citations, no worked methodology, no numerical example, and no limitations section, and it proposed a single composite output — the Quality-Adjusted Planetary Year (QAPY) — modeled directly on the QALY without engaging any of the QALY's own well-documented equity failures (Section 2.10; Chapter 8 develops this comparison in full). PHTA's diagnosis was sound: it correctly identified that fragmented, domain-specific evaluation tools produce systematically inconsistent decisions across sectors, and that non-market contributions such as education, culture, and research impact are frequently invisible to the funding and policy metrics that actually govern resource allocation. But its execution repeated, in compressed form, nearly every failure mode Chapters 2 through 4 have catalogued: an unbounded scope folding industrial products, teaching, research, and entertainment into a single evaluative frame; a composite metric proposed with no conversion methodology or worked example; no structured weighting process; no evidence-grading discipline; no engagement with any of the nine adjacent bodies of literature — One Health, Planetary Boundaries, Doughnut Economics, LCA, the Natural Capital Protocol, MCDA/EVIDEM — that already occupy substantial parts of the space PHTA claimed to fill for the first time.

The revision to Planetary Health Value Assessment addressed the naming problem first, because the naming problem was not merely cosmetic. PHTA's own name sat close enough to the pre-existing, well-established planetary health field (Whitmee et al., 2015) that a careful reviewer would immediately ask whether the proposal was an inadvertent reinvention of existing science under a new label — and the original document neither cited nor distinguished itself from that field. Adopting Planetary Health as the object of concern, rather than inventing a parallel "Planet Health" terminology, situates the framework correctly within the existing scientific landscape rather than beside it, and signals — accurately — that PHVA is proposed as an interoperability layer across existing science, not as a wholesale replacement for it. The shift from Technology Assessment to Value Assessment in the name serves an equivalent function: Technology Assessment already has a specific, narrower meaning in medicine (Section 2.10), and PHVA's ambition — to evaluate the planetary-health value of any human action, not only technologies — is more accurately signalled by the broader term.

5.2 Why PHVA Is More Scientifically Defensible

The substantive gap between PHTA and PHVA is best measured against the specific criteria Chapter 3's comparative analysis and Chapter 4's measurement-gap argument establish as necessary for a credible cross-sectoral evaluation architecture. PHVA cites and positions itself honestly against the nine bodies of literature reviewed in Chapter 2, rather than presenting its domain architecture as a novel invention (Section 2.14's discussion of the pre-existing PHIA literature is a direct example of this honesty: PHVA explicitly declines to reuse the PHIA name, precisely to avoid implying a false identity with an already-claimed, more narrowly scoped term). It adopts a tiered, disaggregated methodology, described in full in Chapter 6, that defers any single composite output behind a mandatory dashboard, directly implementing the lesson Section 2.13's composite-index critique establishes as binding rather than optional. It builds the QALY's documented disability-discrimination history into its design from the outset as a binding constraint on any future headline metric, rather than treating that history as background reading, extensively engaging Sinclair (2012, 2018) and Schneider (2022) at the point where a composite indicator is discussed (Chapter 6, Section 6.5). And it includes its own candid limitations section, discussed in full in Chapter 8, acknowledging that weighting remains inescapably political, that the commensurability problem is managed rather than solved, and that the framework itself remains an unvalidated proposal rather than a tested instrument.

5.3 How PHVA Extends HTA and Complements One Health and Planetary Health

PHVA's working definition is deliberately modeled on the structure of the INAHTA/HTAi consensus HTA definition reviewed in Section 2.10: it is a multidisciplinary process that uses explicit, evidence-graded, multi-criteria methods to characterize the planetary-health consequences and societal value of a human action, product, service, policy, or program across its full lifecycle, with the purpose of informing decision-making so as to promote an equitable, ecologically safe, and resilient relationship between human civilization and the natural systems on which it depends. This is a direct generalization of HTA's own logic — multidisciplinary, evidence-based, lifecycle-oriented, decision-supporting — extended from the single sector (health technologies) HTA was built for to any human action whose planetary-health consequences an institution wishes to assess.

PHVA's relationship to One Health and Planetary Health is one of extension and complementarity rather than substitution. Where One Health's institutional mandate is organized around six action tracks addressing zoonotic disease, antimicrobial resistance, and food safety at the human-animal-ecosystem interface, PHVA is designed to be a natural extension of that institutional infrastructure into a broader set of human activities — industrial production, education, research, cultural production — that One Health was never intended to cover, and Chapter 7's governance pathway explicitly recommends formal engagement with the Quadripartite Secretariat and OHHLEP rather than parallel institution-building. Where Planetary Health supplies the definitional object of concern (the health of civilization and the natural systems it depends on) without itself specifying a measurement methodology, PHVA adopts that definitional object while supplying exactly the operational methodology Planetary Health, as a conceptual field, does not provide.

5.4 The Eight Domains

PHVA organizes evidence into eight domains, chosen against two explicit criteria: each domain must map onto existing measurement infrastructure reviewed in Chapter 2, so that adoption costs are low and no institution is asked to build a measurement capability from zero; and the eight domains, taken together, must cover the environmental ceiling, social floor, safety, contribution, and value dimensions identified across that literature without duplicating any one of them. The architecture is structurally indebted to the EUnetHTA Core Model and the EVIDEM criteria set (Section 2.11), adapted from individual health technologies to any human action, product, or service.

Domain	Definition	Primary Evidentiary Instruments	Possible Quantitative Approach	Possible Qualitative Approach
D1. Environmental & Ecological Effectiveness	Degree to which the action reduces pressure on, or helps restore, the nine planetary boundaries	Planetary boundary control-variable data; GHG/LCA impact categories; biodiversity-intactness indices	LCA characterization scores; distance-to-boundary metrics	Expert narrative on boundary-relevance where data are sparse

Domain	Definition	Primary Evidentiary Instruments	Possible Quantitative Approach	Possible Qualitative Approach
D2. Resource & Material Efficiency	Material and energy intensity per unit of function delivered; circularity; waste avoided	ISO 14040/44 life-cycle inventory data; circular-economy material-flow indicators	Material productivity; recycling input rate; MCI	Qualitative circularity narrative for informal-sector cases
D3. Ecological & Health Safety	Likelihood, severity, and reversibility of unintended harm to ecosystems, animals, or human health	PHIA-style exposure-pathway analysis; One Health surveillance and good-practice guidance	DALYs; exposure-response modelling; SFA-traced pathways	Precautionary narrative where exposure pathways are undercharacterized
D4. Equity & Distributional Justice	Effect on meeting the social foundation for the most affected populations	Doughnut Economics' twelve SDG-derived social-foundation indicators; disaggregated SDG data	Gini coefficient change; SDG indicator scores	Participatory equity narrative where disaggregated data are absent
D5. Resilience & Adaptive Capacity	Contribution to a community's, institution's, or ecosystem's capacity to absorb shocks and adapt	Disaster-risk-reduction and livelihood-diversity indicator sets	Resilience indicator (threshold-proximity); MSA index	Absorptive/adaptive/transformative capacity narrative
D6. Knowledge, Educational & Research Contribution	Contribution to scientific understanding, skills, and public literacy beyond output counts	Outcome/impact evidence, not only publication or patent counts	Citation and downstream-policy-influence tracing where available	Narrative account of demonstrated influence on practice or policy
D7. Cultural & Social-	Contribution to social trust, shared	Qualitative narrative evidence;	Engagement/participation metrics where available	Deliberately not forced into a number

Domain	Definition	Primary Evidentiary Instruments	Possible Quantitative Approach	Possible Qualitative Approach
Cohesion Value	identity, and public engagement through art, sport, media, community life	engagement and participation data		
D8. Economic Value & Cost Implications	Lifecycle cost, avoided cost, and, where available, monetized ecosystem-service or natural-capital impact	Ecosystem-service valuation; Natural Capital Protocol accounts; standard cost-effectiveness data	Net present value; avoided-cost calculation	Reported separately; never substituted for the other seven domains

Table 5.1. The eight PHVA domains, adapted from the domain architecture set out in the PHVA proposal, Section 4.

Domain 8 is reported separately from, and never substituted for, the other seven. This follows directly from the Natural Capital Coalition's own caution, reviewed in Section 2.12, that valuing nature is not the same as commodifying it: a high monetary avoided-cost figure must never be presented as if it resolved a poor score on Domain 3 (safety) or Domain 4 (equity). Domain 7 is the domain most clearly informed by the composite-index critique of Section 2.13 and the Triple Bottom Line's structural insight of Section 1.2: its evidence is deliberately kept qualitative, on the explicit judgment that forcing cultural value into a number would do more harm than good by implying a false precision the underlying evidence — narrative, participatory, engagement-based — cannot support.

5.5 Expected Outputs

A completed PHVA assessment produces, at minimum, a materiality statement identifying which of the eight domains are relevant to the case at hand (Chapter 6, Tier 1); a graded evidence synthesis for each material domain, drawing on the instruments listed in Table 5.1; and a mandatory disaggregated dashboard reporting each material domain's finding alongside its evidence-quality grade. Where an institution elects to proceed to Tier 3 (Chapter 6), the output additionally includes an explicit, documented weighting rationale and a sensitivity analysis across plausible weight scenarios. A single composite headline number is an optional, clearly caveated output, treated in Chapter 6 as the framework's most provisional component rather than its central deliverable.

Chapter 6: Methodological Architecture

PHVA is operationalized as a four-tier process, deliberately structured so that an assessment can stop at Tier 1 or 2 for low-stakes, low-resource cases, and proceed to Tier 3 or 4 only where the decision stakes justify the additional cost. Each tier corresponds to a specific methodological decision, and each decision is made for a specific, traceable reason drawn from the frameworks reviewed in Chapter 2.

6.1 Tier 1: Scoping and Materiality Screening

Following the goal-and-scope phase already required by ISO 14044 (Section 2.1), every PHVA assessment begins by defining the functional unit and system boundary, the comparator, and, critically, which of the eight domains are material to the case at hand. Not every domain is material to every action: Domain 7 may be immaterial to an industrial chemical process while highly material to a media campaign, exactly as Section 6.3's illustrative sketches below demonstrate. This materiality screening keeps the framework proportionate rather than bureaucratically heavy for simple cases, directly answering the scalability concern Chapter 8 documents as a criticism of the earlier PHTA concept — that it was unclear whether the same evidentiary burden applied to evaluating one professor's teaching impact and a national infrastructure project. Materiality screening is why it need not.

6.2 Tier 2: Domain-Level Evidence Synthesis

For each material domain, evidence is synthesized using the existing instrument identified in Table 5.1: an LCA report for Domains 1 and 2, a planetary-boundary distance-to-target calculation for Domain 1, exposure-pathway analysis for Domain 3, disaggregated SDG/social-foundation data for Domain 4, and so on. Each domain-level finding is graded for evidence quality using GRADE-style categories already familiar from HTA practice — high, moderate, low, and very low confidence — rather than presented as a false-precision point estimate. This evidence-grading discipline is the specific methodological device that lets a PHVA dashboard report a high-confidence D1 finding alongside a necessarily low-confidence D7 finding without implying that the two are equally certain.

6.3 Tier 3: Multi-Criteria Synthesis

Where a decision requires comparing options across domains, PHVA follows the EVIDEM/MCDA discipline reviewed in Section 2.11: a deliberative panel — including affected communities and technical experts and ethicists, not technical experts alone — elicits explicit weights for each material domain using a structured method such as swing weighting or discrete-choice methods documented by the ISPOR MCDA Task Force; a sensitivity analysis is run across plausible weight scenarios; and uncertainty is reported as a range, never as false precision. This directly addresses the documented weakness of earlier MCDA/EVIDEM applications — inconsistent uncertainty treatment and difficulty for non-specialist committees to interpret an aggregate score — though Chapter 8 documents that the mechanism offered (sensitivity ranges) may not fully resolve the interpretability problem it is intended to solve.

6.4 Tier 4: Reporting

Every PHVA assessment reports a mandatory, disaggregated dashboard: a score or qualitative rating for each material domain, explicitly traceable to its underlying evidence and instrument. A single composite headline indicator is optional and, where used, must be reported alongside — never instead of — the dashboard, following the headline-plus-dashboard precedent already established by instruments such as the OECD Better Life Index and the UN Sustainable Development Goal indicator framework.

6.5 On a Possible Headline Metric: Planetary Health Years

If an institution chooses to compute a single headline number for high-level communication, the working label Planetary Health Years (PHY) is recommended in place of the earlier Quality-Adjusted Planetary Year, precisely to avoid suggesting a false methodological equivalence with the QALY. The QALY's calculation procedure — health-state utility weighting through standardized public-preference elicitation — does not transfer cleanly to ecological or cultural domains, and attempting to force that transfer is the exact failure mode PHVA is designed to avoid. PHY is treated as a research-stage, aspirational metric, not an operational decision rule, until three conditions are met: its weighting procedure has been subjected to the same disability- and equity-focused scrutiny that exposed the QALY's flaws (Sinclair, 2012, 2018; Schneider, 2022); independent critical review, analogous to the mandatory critical review ISO 14040 requires before any public comparative LCA claim, has been conducted; and it has been piloted in at least one low-stakes domain before any higher-stakes application, such as resource allocation, is considered. No finished PHY calculation formula is specified at this stage, because specifying one prematurely is precisely the failure mode documented in the composite-index literature (Section 2.13).

6.6 Why Each Methodological Decision Was Chosen

Every architectural choice in this chapter traces back to a specific documented failure elsewhere in this monograph. Materiality screening (Tier 1) exists because Chapter 4's four illustrative decision-maker cases each involve a different subset of the eight domains, and forcing every assessment through all eight regardless of relevance would reproduce PHTA's unbounded-scope problem in a more elaborate form. Evidence grading (Tier 2) exists because Chapter 2's frameworks vary enormously in their evidentiary maturity — LCA's characterization factors versus D7's necessarily narrative cultural evidence — and reporting both without a confidence qualifier would imply a false parity. Deliberative, explicit weighting (Tier 3) exists because Section 2.13's composite-index critique establishes that weighting is an ethical choice that cannot be delegated to a formula, and EVIDEM's field-tested discipline (Section 2.11) is the most mature available precedent for making that choice transparent and contestable rather than opaque. The mandatory dashboard (Tier 4) exists because Section 1.2's Triple Bottom Line insight and Section 2.9's Doughnut Economics precedent both establish that heterogeneous value dimensions lose essential information when merged, and because Section 2.13's composite-index literature documents, across three decades of national-level attempts, what happens when that lesson is not heeded. And the provisional, three-condition status of any headline metric (Section 6.5) exists because Section 2.10's QALY history is not a hypothetical risk but a documented, legally consequential case of exactly this failure occurring in the single field — health technology assessment — from which PHVA borrows its evaluative logic most directly.

Chapter 7: Quantification Strategy

Chapter 6 established the procedural architecture within which quantification occurs; this chapter addresses the narrower and, in many respects, harder question of what can actually be measured, by what method, and with what confidence, domain by domain. The governing principle, established in Chapter 5 and reinforced throughout Chapter 6, is that PHVA does not invent new measurement instruments where mature ones already exist, and does not force quantification onto domains whose underlying evidence base cannot support it.

7.1 Environmental Indicators (Domains 1 and 2)

Environmental quantification is, by a wide margin, the most methodologically mature domain in the entire architecture, because it inherits directly from LCA's ISO-standardized characterization-factor methodology (Section 2.1) and from the planetary boundaries' quantified distance-to-target metrics (Section 2.8). The core LCIA equation, S_k = sum over i of (Qi times CF_ik), converts raw inventory flows into comparable impact-category scores, and Richardson et al.'s (2023) simultaneous quantification of all nine planetary boundaries supplies an external ceiling against which any D1 finding can be referenced rather than left as a self-defined benchmark. The principal quantification challenge in this domain is not measurement itself but double-counting: as Chapter 8 documents in detail, a single unit of avoided carbon emission can plausibly be counted under Domain 1 (as environmental effectiveness), Domain 5 (as a contribution to climate resilience), and Domain 8 (once monetized as an avoided cost) unless an explicit accounting rule prevents the same physical benefit from being summed across domains.

7.2 Economic Indicators (Domain 8)

Economic quantification draws on cost-effectiveness methodology already standard in HTA and health economics, together with ecosystem-service valuation (Costanza et al., 1997, 2014) and Natural Capital Protocol accounts (Section 2.12) for the monetization of environmental and natural-capital dimensions specifically. Net present value and standard avoided-cost calculations are directly applicable wherever a comparator scenario is well defined. The domain's central methodological caution, following directly from the Natural Capital Coalition's own guidance, is that monetary valuation of nature is intended to inform decisions, not to establish a market price for it, and that Domain 8 figures must never be netted against, or used to offset, findings in the other seven domains.

7.3 Social and Equity Indicators (Domain 4)

Equity quantification draws on Doughnut Economics' twelve Sustainable Development Goal-derived social-foundation indicators (Section 2.9) and on SIA's household-survey-based methodology (Section 2.3) for project-level distributional analysis. Where DALY-based health quantification is incorporated, the standard formula DALY = YLL + YLD applies directly, using GBD reference life tables for YLL and standardized disability weights for YLD. The central limitation, developed fully in Chapter 8, is that SDG indicator completeness varies enormously by country and is documented to be systematically weaker on disability disaggregation specifically — meaning Domain 4, despite being explicitly designed as PHVA's equity safeguard, could inherit a milder version of the very invisibility problem the framework is designed to guard against, unless the underlying SDG data source is itself audited.

7.4 Resilience Quantification (Domain 5)

Resilience quantification, as Section 2.5 establishes in detail, faces a genuine conceptual difficulty that the other domains do not share to the same degree: resilience is a system property that manifests fully only under perturbation, meaning any pre-perturbation measurement is necessarily a proxy rather than a direct observation. Three proxy approaches are available, each suited to different evidentiary conditions: engineering resilience ($R = 1/\text{Return\_time}$) where time-series data around a known perturbation exist; ecological resilience and threshold-proximity indicators (the $RI = (\text{threshold} - \text{current})/\text{threshold}$ construction illustrated in Section 2.5's brucellosis example, together with early-warning indicators such as rising AR(1) autocorrelation) where a known critical threshold can be specified; and composite indicator batteries (Sendai Framework indicators, the Mean Species Abundance index) where neither a clean time series nor a well-defined threshold is available. Monetization of resilience, where attempted, follows the counterfactual expected-cost-avoided logic documented by Walker et al. (2009), with the explicit caution that even the most developed applications of this logic report two-to-three orders of magnitude of uncertainty.

7.5 Knowledge and Research Contribution (Domain 6)

Knowledge-domain quantification is deliberately built around the same "impact beyond output" principle INAHTA's own recent evaluation agenda has adopted: publication and patent counts are treated as necessary but insufficient evidence, and the domain's evidentiary target is downstream influence on practice or policy, which is considerably harder to establish causally. Where quantifiable, this can draw on citation and policy-uptake tracing; where such tracing is unavailable, the domain relies on documented, narrative evidence of demonstrated influence rather than a substituted output count. Chapter 8 documents that this domain shares a specific, unresolved attribution problem with Domain 7: the gap between what was produced and what changed as a result of it.

7.6 Cultural and Social-Cohesion Value (Domain 7)

Domain 7 is the clearest example in the entire architecture of a domain that should never be forced into a single number. Its evidentiary base — narrative and participatory accounts of shifted public understanding, engagement and participation data where available — is not merely currently under-quantified but is, on the framework's own explicit design judgment, better served by qualitative synthesis than by a manufactured numerical proxy, because a numeric score would imply a precision the underlying evidence cannot support and would risk exactly the kind of false-commensurability failure Section 2.13's composite-index critique warns against.

7.7 Where Quantification Is Currently Possible, Where It Is Not, and Which Domains Should Never Be Forced Into a Single Number

Domains 1, 2, and 8 sit at the mature end of this spectrum, inheriting directly from LCA, planetary-boundary science, and established economic valuation methodology. Domain 3 sits close behind, inheriting from PHIA-style exposure-pathway analysis and DALY methodology, though its evidentiary completeness depends heavily on whether an exposure pathway has already been characterized in the literature. Domains 4 and 5 are quantifiable in principle but constrained in practice by the completeness and disaggregation of the underlying secondary data (SDG indicators for D4; time-series or threshold data for D5). Domain 6 is quantifiable for output counts but only weakly quantifiable for the downstream-influence evidence the domain is actually meant to capture. Domain 7 is, by explicit design choice rather than by current data limitation alone, treated as a domain that should never be collapsed into a single number: forcing quantification here would not resolve an evidentiary gap so much as manufacture a false precision that a mandatory dashboard is specifically designed to avoid.

Domain	**Evidentiary Maturity**	**Primary Quantitative Method**	**Confidence in Current Practice**
D1 Environmental & Ecological Effectiveness	High	LCIA characterization; planetary-boundary distance-to-target	High for GWP, moderate for biosphere-integrity boundary
D2 Resource & Material Efficiency	High	Material productivity; circular-economy indicators	High
D3 Ecological & Health Safety	Moderate	DALY; exposure-response modelling	Moderate, dependent on existing exposure-pathway literature
D4 Equity & Distributional Justice	Moderate	SDG indicator scores; Gini coefficient change	Moderate, weaker on disability disaggregation
D5 Resilience & Adaptive Capacity	Low to Moderate	Threshold-proximity indicators; MSA index	Low, proxy-dependent
D6 Knowledge, Educational & Research Contribution	Low	Citation/policy-uptake tracing where available	Low, attribution-limited
D7 Cultural & Social-Cohesion Value	Deliberately Qualitative	Narrative synthesis; engagement data	Not applicable by design
D8 Economic Value & Cost Implications	High	NPV; ecosystem-service valuation; avoided-cost calculation	High, contingent on comparator specification

Table 7.1. Summary of quantification maturity across the eight PHVA domains, synthesizing the domain-by-domain discussion above.

This maturity gradient has a direct implication for Tier 3 weighting (Section 6.3): a deliberative panel comparing a high-confidence D1 finding against a low-confidence D6 finding is not comparing two numbers of equal epistemic standing, and Chapter 8 identifies the absence of a specified rule for how evidence-quality grade should interact with domain weight as one of the more consequential open methodological questions the framework has not yet resolved.

Chapter 8: Critical Review

A framework of this ambition earns credibility not by defending itself against every possible objection but by stating plainly which objections it has already absorbed into its design and which remain open. This chapter follows that discipline, treating PHTA and PHVA as two successive drafts of one underlying proposal and asking, in turn, what a rigorous, exceptionally critical reviewer would find in each.

8.1 PHTA: A Severe but Warranted Assessment

Read as a scientific proposal rather than a pitch document, PHTA fails on nearly every axis a methodologist would apply. Conceptually, its central definitional claim — that a structured evaluation framework applying engineering, scientific, and economic principles can measure the holistic value of human actions — describes an aspiration rather than an operational construct: it specifies no unit of analysis, no system boundary, and no functional unit, in direct contrast to the mandatory goal-and-scope phase Section 2.1 establishes as foundational to LCA practice. Its scope is functionally unbounded, folding industrial products, teaching, research, and entertainment into one evaluative frame with a single output metric, a pattern the sustainability-metrics literature (Ebert & Welsch, 2004) identifies as the classic signature of an under-specified construct. Its assumption that all human activity can be placed on one continuous worthiness scale is asserted rather than argued, despite this being among the most contested unsolved problems in the composite-index literature reviewed in Section 2.13. Methodologically, no evidence-synthesis method is specified; quantification is entirely unaddressed, with no conversion methodology or worked example for turning a "cultural dividend" or "knowledge dividend" into anything comparable with carbon tonnes; weight assignment across pillars is entirely unspecified, with no elicitation process of the kind ISPOR's MCDA Task Force (Section 2.11) documents as necessary; and the proposed Quality-Adjusted Planetary Year is modeled explicitly on the QALY without engaging either the QALY's own documented equity harms (Sinclair, 2012, 2018; Schneider, 2022) or the general composite-index failure literature (Böhringer & Jochem, 2007) — a striking omission given the direct analogy the document itself draws. Zero validation exists anywhere in the document: no pilot, no numeric example, no sensitivity analysis. Ethically and in terms of governance, equity appears only as a passing measurement example under one pillar, with no equity-weighting safeguard and no engagement with the QALY's documented equity failures despite the QAPY metric being explicitly modeled on it; Indigenous and local knowledge are not mentioned anywhere; and Goodhart's Law and gaming risk are not mentioned once, despite the framework explicitly targeting academic evaluation and funding decisions — a textbook setup for exactly that failure mode. Compared against each of the nine adjacent bodies of literature reviewed in Chapter 2, PHTA is weaker than every named comparator on methodological rigor, evidentiary grounding, and awareness of its own risk profile. What survives this criticism intact is the underlying diagnosis: that fragmented, domain-specific evaluation tools produce inconsistent cross-sectoral decisions, and that non-market contributions are frequently invisible to the metrics that actually govern funding, is correct and shared by serious scholars in the composite-index and MCDA literatures; it is the specific execution, not the underlying goal, that is unready.

8.2 PHVA: A More Mature Proposal, Examined Against a Higher Bar

PHVA is substantially more mature than PHTA: it cites real literature, positions itself honestly against nine existing frameworks, adopts the tiered, disaggregated methodology described in Chapters 6 and 7, and includes its own candid limitations section. Precisely because it is more mature, the appropriate critical task is harder: to identify what a rigorous reviewer would flag beyond what the proposal has already conceded, rather than to repeat criticisms it has already absorbed.

Several conceptual limitations remain unaddressed even in the revised proposal. Domain-boundary ambiguity persists one level up from PHTA's pillar-overlap problem: Domain 1 (environmental and ecological effectiveness) and Domain 3 (ecological and health safety) both reference ecological harm or benefit from different angles — effectiveness versus safety — inviting the same indicator-overlap risk PHTA exhibited, now more precisely labeled but not structurally resolved. Domains 6 and 7 share an attribution problem — output counted is not the same as downstream impact — that the proposal's own Section 6.2 explicitly flags for Domain 6 but does not equally flag for Domain 7, even though causal attribution of cultural influence on behavior is at least as difficult to establish as attribution of research influence on policy. Materiality screening at Tier 1 (Section 6.1) is performed by the assessor before any independent check exists; because which domains are screened as immaterial determines which risks get surfaced at all, this is a first-order gaming vector that is conceptually prior to, and not covered by, the Goodhart's Law risk the proposal explicitly flags for domain-level reporting. GRADE-style evidence grading (Section 6.2), developed in clinical epidemiology for hierarchies of experimental and observational study designs, is applied to Domain 7's qualitative cultural evidence and Domain 6's attribution-weak educational evidence without demonstrated transferability — a conceptual-borrowing problem structurally similar to the HTA/QALY borrowing that undermined PHTA, one layer removed. And no discount-rate stance is specified anywhere for the intergenerational trade-offs Domains 1 and 5 implicitly weigh against Domain 4's near-term equity concerns and Domain 8's near-term cost concerns — a well-known, ethically loaded methodological choice, familiar from the Stern-Nordhaus discount-rate debate in climate economics, that materially changes which options score well.

Methodological limitations are similarly real. No inter-rater reliability data exists yet for Tier 3 domain weighting; the proposal itself lists this as future work, meaning it cannot currently demonstrate that two independent deliberative panels would reach comparable weights for the same case — a basic reliability bar any HTAi or ISPOR reviewer would require before endorsement. No method specifies how domains of differing evidence-quality grades are then compared or weighted at Tier 3: whether a "very low confidence" Domain 7 finding should carry reduced weight, a reduced score, both, or neither is left unspecified. Sensitivity analysis is prescribed across plausible weight scenarios with no method for defining what counts as plausible, risking a sensitivity analysis that is itself arbitrarily bounded, in contrast to the global sensitivity-analysis conventions documented by Saltelli et al. (2008). And no validation plan compares PHVA domain scores or dashboard outputs against realized real-world outcomes before pilot testing begins — predictive validity is deferred to the research agenda (Chapter 9) rather than designed into the pilot phase itself.

Practical implementation limitations compound these. The full Tier 3/4 process is resource-intensive, but no cost estimate distinguishes Tier 1/2 costs from Tier 3/4 costs, so it cannot currently be verified whether materiality screening genuinely keeps small assessments proportionate or whether resource-poor institutions would simply default to skipping Tier 3/4 altogether — reintroducing, for institutional rather than data capacity, the same global-measurement-inequity risk the proposal explicitly flags for data infrastructure. Open publication of domain-level dashboard data is recommended without addressing the commercial-confidentiality barrier that already limits voluntary LCA and Natural Capital Protocol disclosure among private firms. Cross-country applicability of Domain 4's equity criteria assumes reasonably complete, disaggregated SDG data that is systematically weaker in precisely the lower-income and conflict-affected settings the proposal is most concerned to protect. And the governance pathway (Chapter 9) assumes voluntary convening of at least eight independent institutions with different mandates and funding cycles, with no discussion of a fallback pathway if that convening stalls — the single most common failure mode for multi-institutional efforts of this kind.

Ethical and governance limitations include a genuine, still-open representation gap: IPBES's Nature's Contributions to People framework is cited as an evidentiary source for Domains 1 and 3, but no Indigenous or local-knowledge-holder organization is named among the proposed Phase 1 convening group, despite the framework's explicit reliance on that literature. Domain 4's SDG-derived indicators are documented in the Washington Group on Disability Statistics literature to have poor disability disaggregation, meaning Domain 4 could reproduce a milder version of the very disability-invisibility problem the proposal explicitly and correctly worries about for any future PHY metric, without the proposal itself saying so. No conflict-of-interest policy is specified for who sits on Tier 3 deliberative panels, relevant because the institutions proposed to convene the framework may also become its evaluators. And no discussion addresses the distinct risk of "policy-based evidence" — a national agency invoking PHVA domain scores to justify a decision actually made for unrelated reasons — which is a different failure mode from the Goodhart's Law gaming risk the proposal does flag.

Scientifically, no pilot has been conducted despite the framework being detailed enough, in principle, to run at least a low-stakes Tier 1/2 pilot before wider dissemination; no causal model connects Tier 3 weighted dashboard scores to the planetary-health consequences the framework claims to characterize, meaning predictive validity cannot yet even be defined, let alone tested; and internal validity — whether, for example, the Domain 7 evidence-synthesis process reliably distinguishes a documentary that genuinely shifts public understanding from one that is merely widely viewed — is entirely unaddressed.

Placed against each of its named comparators, PHVA is markedly stronger than PHTA and admirably self-aware about its own positioning, but real extrapolation risk remains in applying single-sector tools at cross-sectoral scale: EVIDEM's structured-weighting discipline is a correct methodological borrowing, but EVIDEM has not been demonstrated to scale from health technology's comparatively bounded criteria set to an eight-domain architecture spanning ecology to culture, an extrapolation the proposal does not flag as such. And PHVA's eight-domain dashboard, while it correctly borrows the Doughnut's ceiling/floor structural logic (Section 2.9), risks being considerably less legible to a lay policymaker than the Doughnut's two-boundary visual — a usability regression relative to the framework PHVA itself cites as design inspiration.

8.3 What PHVA Currently Addresses, and Where Gaps Remain

Balanced against this catalogue of open questions, several genuine strengths survive scrutiny and are worth stating without qualification. PHVA's explicit refusal to collapse into a single composite metric before independent validation is a genuinely rare, mature methodological choice relative to almost all composite-sustainability-index precedents. Its literature review honestly identifies prior art — PHIA, One Health, Doughnut Economics — rather than claiming false originality, a substantial improvement over PHTA. Its QALY-informed equity safeguard, designing explicitly against the QALY's documented disability-discrimination history before building any aggregate metric, is a legitimately novel contribution relative to most sustainability-index proposals, which rarely engage health-economics equity literature at all. And its phased governance pathway, while ambitious, correctly identifies that a framework of this kind needs institutional co-ownership rather than parallel institution-building, explicitly recommending collaboration with the existing PHIA and CINPHIA research line rather than duplicating it. The honest summary a critical reviewer would offer is this: PHVA has already done the harder half of the work required of a credible cross-sectoral evaluation architecture — stating its own limitations plainly, refusing premature aggregation, and grounding its domain architecture in established instruments — while the harder empirical half, validation, inter-rater reliability, and at least one completed real-world pilot, remains entirely to be done, and Chapter 9 sets out the concrete steps that doing it would require.

Chapter 9: Future Research Agenda

The preceding chapter's catalogue of open questions translates directly into a research agenda, sequenced deliberately to avoid the premature-aggregation failure mode that recurs throughout this monograph as the single most consistent cause of framework failure, from the composite-index literature of Section 2.13 through the QALY's history in Section 2.10 to PHTA's own execution in Section 8.1.

9.1 Validation and Pilot Testing

The single highest-priority item is completing at least one real, not illustrative, low-stakes pilot — the university-curriculum sketch outlined in Section 5.5's discussion of expected outputs is a natural candidate, given its high data availability and low decision stakes — with published, open dashboard data, before any wider dissemination or higher-stakes application is attempted. Directly following this, the inter-rater reliability study on Tier 3 domain weighting that Chapter 8 identifies as currently absent should be commissioned and its results published before any Tier 3 or Tier 4 application is used in a real funding or policy decision; this single validation study would resolve the most basic reliability objection a rigorous methodologist would raise.

9.2 Case Studies Across Sectors

Beyond the university-curriculum pilot, at least two or three additional low-stakes, high-data-availability sectors should be selected to test the full tiered methodology of Chapter 6 without attempting a headline metric: agricultural technologies already tracked through existing LCA studies are a natural second candidate, given the density of existing Domain 1 and Domain 2 evidence documented in Chapter 2; a national HTA agency's pilot of a planetary-boundary screening module as an additional, optional domain within one existing HTA report category is a natural third, building directly on the 2024 scoping-review trend already documented in the HTA literature (Section 5.3).

9.3 International Collaboration

Consensus-building should be convened as a joint task group, modeled explicitly on the INAHTA/HTAi process that produced the 2020 consensus HTA definition reviewed in Section 2.10, bringing together the Quadripartite/OHHLEP Secretariat, the Planetary Health Alliance, the Stockholm Resilience Centre, the ISPOR MCDA Special Interest Group, the Capitals Coalition, IPBES, and the CINPHIA/Wageningen University PHIA research team, to refine the domain architecture and weighting governance described in Chapters 5 and 6 before any pilot begins. Chapter 8's representation gap should be addressed directly and early in this process: at least one Indigenous or local-knowledge-holder organization should be named as a Phase 1 convener, not referenced only as an evidentiary source through IPBES, given the framework's explicit reliance on the Nature's Contributions to People literature.

9.4 Governance

A standing, independent review function should be established, modeled on the mandatory critical review ISO 14040 requires for public comparative LCA claims, before any PHVA output, and especially any future PHY headline number, is used in a public comparative claim or a binding policy or funding decision. A documented, independently reviewed materiality-screening justification requirement should be added to close the Tier 1 gaming vector Chapter 8 identifies, modeled on LCA's own scope-documentation requirements. And an explicit discount-rate policy, or at minimum a disclosed sensitivity range across plausible discount rates, should be adopted for domains involving long-horizon outcomes, closing the intergenerational-trade-off gap Chapter 8 identifies as entirely unaddressed in the current proposal.

9.5 Software Development and AI Integration

The evidence-synthesis burden at Tier 2, particularly for Domain 3's exposure-pathway analysis, is a natural candidate for the kind of expert-elicited knowledge-graph and AI-assisted evidence-synthesis methodology Stefanopoulou et al. (2025) have already developed for PHIA; direct collaboration with that research line, rather than parallel tool development, follows the same institutional-co-ownership principle Chapter 6 applies to governance generally. Dashboard-reporting software should be developed jointly with, rather than independently of, the standard-setting bodies (ISO, the Global Reporting Initiative, the Capitals Coalition) whose existing LCA and Natural Capital Protocol reporting infrastructure Domain 1, Domain 2, and Domain 8 already depend on, to avoid producing a reporting format incompatible with the systems institutions already use.

9.6 Policy Translation and Institutional Adoption

Universities are the most tractable early adopter, piloting PHVA-informed research-impact evaluation as an explicit, opt-in supplement to, not replacement for, existing publication- and patent-based evaluation, consistent with the impact-beyond-output lesson HTA itself has already learned (Section 5.3). National HTA and regulatory agencies should commission an independent ethics review of any proposed PHY-style headline metric before allowing its use in coverage or funding decisions, applying the same scrutiny the QALY received only after years of documented harm (Section 2.10) — reversing, deliberately, the order in which that scrutiny was historically applied to the QALY itself. Funders should require that any funded pilot publish its domain-level dashboard data openly, to support the independent validation work described in Section 9.1, and should fund Tier 1/2 pilots before funding any attempt to operationalize a cross-domain headline metric.

9.7 Future Methodological Improvements

Beyond the near-term agenda above, three longer-horizon methodological questions remain genuinely open and are flagged here rather than resolved: whether an adapted, qualitative-evidence confidence framework such as GRADE-CERQual can be piloted specifically for Domain 6 and Domain 7 evidence before GRADE's clinical-epidemiology-derived categories are applied more widely to qualitative cultural and educational-impact evidence; whether a minimum-viable, single-institution pilot pathway can be defined independently of full Quadripartite, IPBES, and Capitals Coalition buy-in, so that the framework's development does not stall entirely if multilateral convening proves slower than the four-phase governance pathway anticipates; and whether a costed worked example contrasting a Tier 1/2-only assessment against a full Tier 1 through 4 assessment can be published, so that the proportionality claim implicit in the tiered design — that materiality screening genuinely keeps small assessments cheap — can be evaluated empirically rather than asserted.

Concluding Synthesis

This monograph has traced one continuous argument across nine chapters. Technology assessment emerged in the 1960s with structural blind spots — economic and technical bias, environmental neglect, social and equity undervaluation — that the sustainability turn of the following decades, from the Brundtland Commission through the Triple Bottom Line to planetary boundaries science, was built specifically to correct. Five decades of subsequent work produced a family of frameworks — life-cycle assessment, environmental impact assessment, social impact assessment, industrial ecology, and resilience assessment, alongside One Health, planetary health, Doughnut Economics, health technology assessment, and multi-criteria decision analysis — each of which achieves genuine rigor within a specific, deliberately bounded evaluative domain. Placed side by side, these frameworks reveal a consistent structural pattern: rigor is purchased by narrowing scope, and no framework was built, or claims, to hold economic, environmental, social, cultural, knowledge, equity, resilience, and planetary-health value simultaneously in view for an arbitrary human action. This is the measurement gap that motivates everything that follows: not a failure of any individual framework, but the absence of a shared architecture for comparing outputs across them without either forcing a false commensurability or abandoning the comparison to informal, unauditable judgment.

Planetary Health Value Assessment is examined in this monograph as a candidate response to that gap, in two successive drafts. The earlier Planet Health Technology Assessment concept correctly diagnosed the problem but failed nearly every methodological test a rigorous reviewer would apply — unbounded scope, unaddressed quantification, no structured weighting, no evidence grading, no engagement with the QALY's documented equity failures despite proposing a metric modeled directly on it, and zero empirical validation. The revised PHVA proposal addresses the majority of these failures directly: it positions itself honestly against nine adjacent bodies of literature, adopts a tiered, disaggregated methodology that defers any composite metric behind a mandatory dashboard, and builds the QALY's disability-discrimination history into its design as a binding constraint from the outset. Real limitations remain, catalogued in full in Chapter 8: domain-boundary ambiguity, unresolved attribution problems in the knowledge and cultural domains, an unaudited materiality-screening gaming vector, absent inter-rater reliability data, no discount-rate policy for intergenerational trade-offs, and — most consequentially — zero completed pilots and no validated causal model connecting dashboard scores to real-world planetary-health outcomes.

The honest position this monograph arrives at, and the position a critical reviewer would be right to insist on, is that PHVA is a credible, well-reasoned proposal for a methodological architecture, not yet a validated evaluation instrument. Its most defensible feature is precisely the feature that distinguishes it most sharply from the composite-index tradition it explicitly declines to repeat: a refusal to claim more certainty than its current evidence base supports. Chapter 9's research agenda — validation, pilot testing, inter-rater reliability studies, international collaboration through existing institutions rather than parallel institution-building, and a governance structure built around independent review from the outset — is not a list of afterthoughts appended to a finished framework. It is the specific, sequenced work that separates a well-reasoned proposal from a validated instrument, and it is the work this monograph identifies as the necessary next step before any of the framework's more ambitious applications — resource allocation, faculty evaluation, national policy comparison — can be responsibly attempted.

References

Bene, C., Wood, R. G., Newsham, A., & Davies, M. (2012). Resilience: New utopia or new tyranny? *IDS Working Paper*, 2012(405), 1–61.

Böhringer, C., & Jochem, P. E. P. (2007). Measuring the immeasurable — A survey of sustainability indices. *Ecological Economics*, 63(1), 1–8.

Capitals Coalition (formerly Natural Capital Coalition). (2016). *Natural Capital Protocol*. Capitals Coalition.

Chambers, R. (1994). The origins and practice of participatory rural appraisal. *World Development*, 22(7), 953–969.

Chertow, M. R. (2000). Industrial symbiosis: Literature and taxonomy. *Annual Review of Energy and the Environment*, 25, 313–337.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, 357(4), 370–379.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R. G., Sutton, P., & van den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*, 387, 253–260.

Costanza, R., de Groot, R., Sutton, P., van der Ploeg, S., Anderson, S. J., Kubiszewski, I., Farber, S., & Turner, R. K. (2014). Changes in the global value of ecosystem services. *Global Environmental Change*, 26, 152–158.

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*, 9(3), 458–467.

Ebert, U., & Welsch, H. (2004). Meaningful environmental indices: A social choice approach. *Journal of Environmental Economics and Management*, 47(2), 270–283.

Ekvall, T., & Weidema, B. P. (2004). System boundaries and input data in consequential life cycle inventory analysis. *International Journal of Life Cycle Assessment*, 9(3), 161–171.

European Parliament and Council. (2001). *Directive 2001/42/EC on the assessment of the effects of certain plans and programmes on the environment*. Official Journal of the European Communities.

Ferguson, N. M., Cummings, D. A. T., Fraser, C., Cajka, J. C., Cooley, P. C., & Burke, D. S. (2006). Strategies for mitigating an influenza pandemic. *Nature*, 442(7101), 448–452.

Finnveden, G., Hauschild, M. Z., Ekvall, T., Guinée, J., Heijungs, R., Hellweg, S., Koehler, A., Pennington, D., & Suh, S. (2009). Recent developments in life cycle assessment. *Journal of Environmental Management*, 91(1), 1–21.

Fischer-Kowalski, M. (1998). Society's metabolism: The intellectual history of materials flow analysis, Part I, 1860–1970. *Journal of Industrial Ecology*, 2(1), 61–78.

Folke, C., Carpenter, S. R., Walker, B., Scheffer, M., Chapin, T., & Rockström, J. (2010). Resilience thinking: Integrating resilience, adaptability and transformability. *Ecology and Society*, 15(4), Article 20.

Food and Agriculture Organization of the United Nations. (2016). *RIMA-II: Resilience index measurement and analysis-II*. FAO.

Frischknecht, R., Jungbluth, N., Althaus, H.-J., Doka, G., Dones, R., Heck, T., Hellweg, S., Hischier, R., Nemecek, T., Rebitzer, G., & Spielmann, M. (2005). The ecoinvent database: Overview and methodological framework. *International Journal of Life Cycle Assessment*, 10(1), 3–9.

Geissdoerfer, M., Savaget, P., Bocken, N. M. P., & Hultink, E. J. (2017). The circular economy — A new sustainability paradigm? *Journal of Cleaner Production*, 143, 757–768.

Goetghebeur, M. M., et al. (2011). Bridging health technology assessment (HTA) with multicriteria decision analyses (MCDA): Field testing of the EVIDEM framework for coverage decisions by a public payer in Canada. *BMC Health Services Research*, 11, 329.

Gunderson, L. H., & Holling, C. S. (Eds.). (2002). *Panarchy: Understanding transformations in human and natural systems*. Island Press.

Huijbregts, M. A. J., Steinmann, Z. J. N., Elshout, P. M. F., Stam, G., Verones, F., Vieira, M., Zijp, M., Hollander, A., & van Zelm, R. (2017). ReCiPe2016: A harmonised life cycle impact assessment method at midpoint and endpoint level. *International Journal of Life Cycle Assessment*, 22(2), 138–147.

Hunt, R. G., & Franklin, W. E. (1996). LCA — How it came about: Personal reflections on the origin and the development of LCA in the USA. *International Journal of Life Cycle Assessment*, 1(1), 4–7.

International Network of Agencies for Health Technology Assessment (INAHTA). (2026). *HTA definition and network overview*. https://www.inahta.org/

International Organization for Standardization. (2006). *ISO 14040:2006 — Environmental management — Life cycle assessment — Principles and framework*.

International Organization for Standardization. (2006). *ISO 14044:2006 — Environmental management — Life cycle assessment — Requirements and guidelines*.

ISPOR MCDA Emerging Good Practices Task Force. (2016). Multiple criteria decision analysis for health care decision making — An introduction: Report 1. *Value in Health*, 19(1), 1–13.

ISPOR MCDA Emerging Good Practices Task Force. (2016). Multiple criteria decision analysis for health care decision making — Emerging good practices: Report 2. *Value in Health*, 19(2), 125–137.

Kloepffer, W. (2008). Life cycle sustainability assessment of products. *International Journal of Life Cycle Assessment*, 13(2), 89–95.

Kruk, M. E., Myers, M., Varpilah, S. T., & Dahn, B. T. (2015). What is a resilient health system? Lessons from Ebola. *The Lancet*, 385(9980), 1910–1912.

Morse, B., & Berger, T. R. (1992). *Sardar Sarovar: Report of the Independent Review*. Resource Futures International.

National Center for Disability, Equity, and Intersectionality. (2021). *It is time to let go of the QALY's legacy of discrimination*.

Oezesmi, U., & Oezesmi, S. L. (2004). Ecological models based on people's knowledge: A multi-step fuzzy cognitive mapping approach. *Ecological Modelling*, 176(1–2), 43–64.

One Health High-Level Expert Panel (OHHLEP). (2021). *Operational definition of "One Health."* World Health Organization.

O'Rourke, B., Oortwijn, W., & Schuller, T. (2020). The new definition of health technology assessment: A milestone in international collaboration. *International Journal of Technology Assessment in Health Care*, 36(3), 187–190.

Quadripartite (FAO, UNEP, WHO, & WOAH). (2022). *One Health Joint Plan of Action (2022–2026)*.

Raworth, K. (2012). *A safe and just space for humanity: Can we live within the doughnut?* Oxfam Discussion Paper.

Raworth, K. (2017). A doughnut for the Anthropocene: Humanity's compass in the 21st century. *The Lancet Planetary Health*, 1(2), e48–e49.

Raworth, K. (2017). *Doughnut economics: Seven ways to think like a 21st-century economist*. Random House Business Books.

Richardson, K., Steffen, W., Lucht, W., Bendtsen, J., Cornell, S. E., Donges, J. F., Drüke, M., Fetzer, I., Bala, G., von Bloh, W., Feulner, G., Fiedler, S., Gerten, D., Gleeson, T., Hofmann, M., Huiskamp, W., Kummu, M., Mohan, C., Nogués-Bravo, D., … Rockström, J. (2023). Earth beyond six of nine planetary boundaries. *Science Advances*, 9(37), eadh2458.

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., … Foley, J. A. (2009). A safe operating space for humanity. *Nature*, 461, 472–475.

Schneider, P. (2022). The QALY is ableist: On the unethical implications of health states worse than dead. *Quality of Life Research*, 31, 1545–1552.

Scoones, I. (1998). *Sustainable rural livelihoods: A framework for analysis* (IDS Working Paper No. 72). Institute of Development Studies.

Sinclair, S. (2012). How to avoid unfair discrimination against disabled patients in healthcare resource allocation. *Journal of Medical Ethics*, 38(3), 158–162.

Sinclair, S. (2018). Improved health state descriptions will not benefit disabled patients under QALY-based assessment. *Journal of Medical Ethics*, 44(11), 797–798.

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., de Vries, W., de Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B., & Sörlin, S. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science*, 347(6223), 1259855.

Stefanopoulou, M., Sonnenschein, T. S., Poulletier de Gannes, F., Scheider, S., Vermeulen, R., Röösli, M., & Huss, A. (2025). Towards a planetary health impact assessment framework: Exploring expert knowledge and artificial intelligence for a RF-EMF exposure case-study. *Bioelectromagnetics*. https://doi.org/10.1002/bem.70038

Swarr, T. E., Hunkeler, D., Klöpffer, W., Pesonen, H.-L., Ciroth, A., Brent, A. C., & Pagan, R. (2011). Environmental life-cycle costing: A code of practice. *International Journal of Life Cycle Assessment*, 16(5), 389–391.

United Nations Environment Programme, & Society of Environmental Toxicology and Chemistry. (2009). *Guidelines for social life cycle assessment of products*. UNEP/SETAC Life Cycle Initiative.

Van Boeckel, T. P., Brower, C., Gilbert, M., Grenfell, B. T., Levin, S. A., Robinson, T. P., Teillant, A., & Laxminarayan, R. (2015). Global trends in antimicrobial use in food animals. *PNAS*, 112(18), 5649–5654.

Vanclay, F. (2003). International principles for social impact assessment. *Impact Assessment and Project Appraisal*, 21(1), 5–12.

Wageningen University & Research. (2026). *Postdoc: Developing a Planetary Health Impact Assessment in policy practice (CINPHIA consortium)* [Position description].

WHO European Centre for Health Policy. (1999). *Health impact assessment: Main concepts and suggested approach* (Gothenburg Consensus Paper). WHO Regional Office for Europe.

Whitmee, S., Haines, A., Beyrer, C., Boltz, F., Capon, A. G., de Souza Dias, B. F., et al. (2015). Safeguarding human health in the Anthropocene epoch: Report of The Rockefeller Foundation–Lancet Commission on planetary health. *The Lancet*, 386(10007), 1973–2028.

World Business Council for Sustainable Development. (2000). *Eco-efficiency: Creating more value with less impact*. WBCSD.

World Commission on Dams. (2000). *Dams and development: A new framework for decision-making*. Earthscan.

Planetary Health Value Assessment (PHVA)

Applying the Eight-Domain, Four-Tier PHVA Architecture to an AI-Enabled Zoonotic Disease Surveillance System for Brucellosis in Rural India

Purpose and How to Read This Example

This document is a worked example, not a validated empirical assessment. No primary data collection has been carried out for the case described here; where figures are given, they are drawn from the published literature on human and livestock brucellosis in India and from the published performance characteristics of comparable AI-enabled surveillance systems, and are clearly flagged as such. Its purpose is methodological: to demonstrate, on a single concrete case, how the PHVA domain architecture and four-tier process (Sections 4 and 5 of the PHVA proposal) would actually structure an assessment — including the trade-offs and evidentiary gaps a genuine application would surface rather than hide.

The case was chosen deliberately. It sits inside the same One Health/HTA space the PHVA proposal uses to motivate the framework, it connects directly to an active area of applied brucellosis epidemiology in India, and it is exactly the kind of intervention — part digital health tool, part animal-health program, part public-health surveillance system — that existing sector-bound frameworks struggle to evaluate as a single object, which is the argument developed in Section 2 below.

1. The Intervention

AI-enabled zoonotic disease surveillance system for brucellosis (rural India). The intervention is a digital surveillance platform that ingests four existing data streams — (i) livestock serological screening results (RBPT and i-ELISA) collected under India's National Animal Disease Control Programme (NADCP), which since 2019 has targeted universal S19 vaccination of female calves aged four to eight months; (ii) veterinary field-reporting and milk-cooperative records; (iii) human syndromic and confirmed-case data from the Integrated Disease Surveillance Programme (IDSP) and primary health centres; and (iv) agro-climatic and grazing-pattern data. A machine-learning risk-prediction model, deployed through a mobile application used by village-level veterinary field staff and ASHA (Accredited Social Health Activist) workers, combines these streams to flag herds, villages, and occupationally exposed households (dairy farmers, animal attendants, veterinarians, abattoir workers) for prioritized RB51/S19 vaccination, confirmatory serological testing, and human case-finding, and surfaces results on a district-level dashboard for animal husbandry and public-health officers.

The case is chosen because published seroprevalence and burden estimates already establish that this is a non-trivial problem at scale: human brucellosis has been estimated to cause a median of 177,601 DALYs annually in India (0.15 DALYs per 1,000 persons per year, rising to 0.29 per 1,000 in occupationally exposed adults), at an estimated annual economic loss of roughly USD 10.46 million, while bovine seroprevalence has been estimated at a pooled 16.6% in cattle nationally and has been shown to vary sharply by agro-climatic zone — from 1.45% to 8.73% within a single state in one study — meaning any surveillance tool's performance is likely to be geographically uneven rather than uniform.

2. Why Existing Frameworks Cannot Fully Evaluate This Intervention

Table 1 summarizes, framework by framework, what each established methodology would capture about this intervention and what it would leave out. The pattern is the same one that motivates the PHVA proposal generally (Section 2 of the proposal): each framework is disciplined and well-validated within its own domain, and each is structurally blind to at least three of the eight PHVA domains that this particular intervention implicates simultaneously.

Framework	What it would capture	What it would miss
Life-Cycle Assessment (LCA)	Could quantify the ICT hardware, cloud-compute, and connectivity footprint of the platform (Domain 1/2) with ISO 14040/44 rigor.	No native treatment of diagnostic accuracy, health outcomes, equity of access, or cultural trust — the dimensions that actually determine whether the system works as a health intervention.
Environmental Impact Assessment (EIA)	Would screen out the intervention almost immediately as having no significant physical footprint (no construction, no land-use change, no emissions of regulatory concern).	That correct screening decision tells a decision-maker nothing about whether the system is a good use of public health resources — EIA has no mechanism to say anything at all about Domains 3–8.
Social Impact Assessment (SIA)	Directly relevant to distributional and livelihood questions (Domain 4) — who gains access, who is excluded, effects on occupationally exposed groups.	Weak on quantitative epidemiological outcomes (Domain 3) and has no natural way to price avoided disease burden against system cost (Domain 8).
Health Technology Assessment (HTA) / QALY	Its cost-per-DALY-averted logic (Section 2.5 of the PHVA proposal) is the most directly transferable tool here, and is used explicitly in Domain 8 below.	Standard HTA practice does not systematically incorporate animal-health, environmental, or cultural-trust dimensions, and the QALY's documented equity failures (Sinclair, 2012, 2018; Schneider, 2022) are a specific risk for a tool whose highest-value targets are occupationally exposed, often lower-income livestock keepers.
One Health (Quadripartite / OHHLEP)	This is the closest institutional home for the intervention conceptually — it is precisely a human–animal–environment surveillance tool of the kind One Health was built to coordinate.	One Health's Joint Plan of Action (2022–2026) is an institutional coordination and governance framework built around action tracks, not a scoring or comparative-value methodology; it can house the intervention but cannot itself tell a funder whether this specific AI system is a better use of

Framework	What it would capture	What it would miss
		resources than, say, scaling up conventional serosurveillance.
Planetary Health (Lancet Commission)	Provides the civilizational framing — zoonotic spillover risk as a planetary-health concern — that motivates taking the intervention seriously at all.	Is explicitly a conceptual field and social movement, not an assessment methodology (Section 2.2 of the PHVA proposal); it has no domain architecture, indicators, or tiered process to apply to a specific system.

The structural gap is therefore not that any one framework is weak — each is strong within its remit — but that no single one of them is built to hold the environmental footprint, diagnostic safety, equity of targeting, system resilience, workforce capacity-building, community trust, and cost-effectiveness of this intervention inside one auditable, comparable evaluation. That is precisely the interoperability-layer role PHVA is designed to play (Section 3 of the PHVA proposal), which Section 4 below demonstrates domain by domain.

3. Domain-by-Domain Application

Each domain below follows the same structure requested for a worked example: indicators, data requirements, quantitative methods, qualitative methods where quantification is not defensible, and expected benefits and risks. Materiality judgments (Tier 1) are stated for each domain and are revisited formally in Section 4.1.

Domain 1: Environmental and Ecological Effectiveness

Degree to which the action reduces pressure on, or helps restore, the nine planetary boundaries.

Materiality (Tier 1): Low. The intervention is a software and data-integration layer running on existing mobile and cloud infrastructure; it has no construction, land-use, or direct emissions footprint. It is retained in the assessment, at low weight, because indirect effects on antimicrobial use are plausible and because Tier 1 screening decisions should be shown, not merely assumed.

Indicators	• Cloud-compute and data-transmission energy use (kWh per year) attributable to the platform • Device-level embodied footprint of additional mobile handsets or sensors procured for field staff • Avoided-antimicrobial-use proxy: change in empirical (non-diagnosis-based) antibiotic treatment courses in livestock following improved diagnostic targeting

Data Requirements	• Server/cloud provider energy and carbon disclosures for the hosting infrastructure • Procurement records for any new field devices • Veterinary treatment records pre- and post-deployment, disaggregated by whether treatment followed a confirmatory test or was empirical
Quantitative Methods	• Screening-level life-cycle assessment (ISO 14040/44) of the ICT stack, expressed in kg CO_2-eq per year • Simple before/after comparison of empirical antimicrobial treatment courses per 1,000 animals, as an indirect antimicrobial-resistance-pressure indicator
Qualitative Methods	• Narrative statement of why Domain 1 is screened as low-priority (Tier 1), so the omission is a documented judgment rather than a silent gap
Expected Benefits	• Negligible direct environmental footprint relative to physical infrastructure alternatives • Plausible small positive effect on antimicrobial stewardship through better-targeted, diagnosis-led treatment
Risks / Limitations	• E-waste from handset turnover and possible sensor deployment, not currently tracked by any existing instrument • Energy footprint of model training and periodic retraining, small in absolute terms but currently unmeasured

Domain 2: Resource and Material Efficiency

Material and energy intensity per unit of function delivered; circularity; degree to which the system reuses rather than duplicates existing infrastructure.

Materiality (Tier 1): Moderate. The system's central efficiency claim — fewer, better-targeted diagnostic tests and vaccine doses per case detected — is a genuine PHVA-relevant efficiency question, distinct from Domain 1's environmental framing.

Indicators	• Serological tests conducted per confirmed case, before versus after AI-assisted targeting • Vaccine doses (S19/RB51) delivered per epidemiologically material herd, versus blanket delivery • Proportion of the platform built on reused NADCP/IDSP infrastructure versus newly commissioned systems

Data Requirements	• NADCP and IDSP testing and vaccination logs, before and after deployment, at district level • System architecture documentation distinguishing reused from newly built components
Quantitative Methods	• Material/test efficiency ratio: confirmed cases per 100 serological tests, compared against a pre-intervention baseline • A simple circularity indicator: proportion of infrastructure inputs (servers, connectivity, field devices) reused from existing government digital-health systems rather than newly procured
Qualitative Methods	• Narrative account of interoperability with NADCP and IDSP data standards, since a system built as a parallel silo would score well on paper efficiency metrics while adding real duplication cost the indicators would not capture
Expected Benefits	• Reduced redundant serological testing and better-targeted vaccine allocation, if the risk-prediction model is well calibrated • Lower marginal infrastructure cost where existing NADCP/IDSP systems are reused
Risks / Limitations	• Vendor or platform lock-in if the AI system is built on proprietary, non-interoperable infrastructure • Maintenance burden shifting local government staff time away from other animal-health functions

Domain 3: Ecological and Health Safety

Likelihood, severity, and reversibility of unintended harm to ecosystems, animals, or human health; and, conversely, the core diagnostic-safety case for the intervention.

Materiality (Tier 1): High. This is the intervention's central materiality domain: its entire rationale is improving the safety and timeliness of zoonotic disease detection.

Indicators	• Sensitivity and specificity of the AI risk-prediction model against confirmed serological and clinical outcomes • Time from exposure or symptom onset to confirmatory testing, before versus after deployment • DALYs averted, modelled against the existing national burden estimate of 177,601 DALYs annually (0.15 per 1,000 persons per year; 0.29 per 1,000 in occupationally exposed adults)

	• False-negative rate and its downstream clinical consequence (delayed treatment, chronic sequelae)
Data Requirements	• Model validation datasets with ground-truth confirmatory test results, stratified by agro-climatic zone given documented seroprevalence heterogeneity (1.45%–8.73% across zones in one state-level study) • IDSP case-timeline data (symptom onset to diagnosis) pre- and post-deployment • Existing national DALY and economic-loss burden estimates as the counterfactual baseline
Quantitative Methods	• ROC/AUC-based model performance evaluation, reported separately by agro-climatic zone rather than as a single national figure • DALY-averted modelling using standard YLL/YLD decomposition, referenced against the existing national burden estimate • An extension of pathogen substance-flow analysis (SFA) tracing Brucella transmission from livestock reservoirs to human occupational exposure, comparing modelled flow before and after AI-assisted targeting
Qualitative Methods	• A precautionary narrative addressing false negatives specifically in agro-climatic zones or population subgroups underrepresented in the model's training data • Documentation of clinical escalation pathways when the model flags a case, so safety does not rest on the algorithm's output alone
Expected Benefits	• Earlier case detection and more precisely targeted RB51/S19 vaccination in high-risk herds • Potential reduction in the occupational DALY burden, which is nearly double the general-population rate
Risks / Limitations	• False negatives in data-sparse zones or subgroups, creating a false sense of security that is worse than no algorithmic triage at all • Automation bias: frontline veterinary and ASHA workers deferring to model output even where local clinical judgment would flag a case the model missed

Domain 4: Equity and Distributional Justice

Effect on meeting the social foundation (following the Doughnut Economics logic adopted in the PHVA proposal, Section 2.4) for the most affected populations.

Materiality (Tier 1): High. The intervention's target population — dairy farmers, animal attendants, landless livestock labourers, and tribal households — is precisely the population documented in the literature as bearing the highest occupational seroprevalence (up to 20% in some occupational subgroups) and the least existing digital infrastructure.

Indicators	• Coverage of AI-flagged case-finding and vaccination activity, disaggregated by landholding size, gender, and tribal status • Smartphone ownership and network connectivity by district, as a precondition for the mobile-application component to reach a household at all • Distribution of app usability outcomes among low-literacy users
Data Requirements	• Disaggregated NADCP/IDSP uptake data by caste/tribal status, gender, and landholding size • District-level smartphone and mobile-network penetration statistics • Household-level occupational exposure data (farm workers, veterinary pharmacists, animal attendants, abattoir workers), whose seroprevalence has been separately documented in the literature
Quantitative Methods	• Gini-coefficient-style comparison of the distribution of early-detection benefit across landholding and occupational strata • SDG-indicator-style disaggregated coverage scores (proxy for SDG 3 and SDG 10 targets)
Qualitative Methods	• Participatory equity assessment conducted directly with tribal and landless livestock-keeping households on app usability, language, and gender access to the household smartphone • Documentation of who within a household actually operates the device the system depends on, since ownership is frequently concentrated in male heads of household
Expected Benefits	• If deliberately targeted, the system could close an existing detection gap for the highest-risk occupational subgroups rather than merely reaching those already best served
Risks / Limitations	• A digital-equity paradox: an AI system requiring smartphone and network access risks systematically under-detecting the most marginalized households — landless labourers, women, and tribal communities with lower digital literacy — who are simultaneously the population with the highest documented occupational exposure

	• Risk that vaccination and testing resources are drawn toward well-connected, better-resourced herds simply because they generate more usable data for the model

Domain 5: Resilience and Adaptive Capacity

Contribution to the surveillance systems, and the wider animal- and human-health system's, capacity to absorb shocks and adapt.

Materiality (Tier 1): Moderate to high. A surveillance tool is only as valuable as its performance under exactly the conditions monsoon-season connectivity loss, staff turnover, local outbreaks that most threaten conventional systems.

Indicators	• System uptime and function during monsoon-season connectivity disruption • Time-to-recovery after a false alarm, model failure, or data-pipeline outage • Retention of trained veterinary and ASHA field staff, since resilience depends on people as much as software
Data Requirements	• Platform uptime/downtime logs and offline-mode functionality test results • Staff training and retention records over the deployment period
Quantitative Methods	• A threshold-proximity indicator estimating how close routine operating conditions (connectivity, staffing) sit to the point of system failure • An adaptation of the R0–resilience equivalence approach — estimating the reduction in effective reproduction achievable per unit of surveillance responsiveness — to compare the system's contribution to outbreak containment against a conventional serosurveillance baseline
Qualitative Methods	• An absorptive/adaptive/transformative capacity narrative: does the system degrade gracefully to manual paper-based reporting during network failure, or does surveillance capacity collapse entirely with the platform
Expected Benefits	• Potential to strengthen the wider zoonotic early-warning system, provided it is designed to degrade gracefully rather than as a single point of failure
Risks / Limitations	• Over-centralization: if district dashboards become the sole trusted source of truth, a technical outage during an actual outbreak is the worst possible failure timing

	• Dependency on connectivity infrastructure that is itself weakest in the tribal and remote rural belts the system is meant to serve

Domain 6: Knowledge, Educational, and Research Contribution

Contribution to scientific understanding, workforce skills, and public epidemiological literacy, beyond simple output counts.

Materiality (Tier 1): Moderate. Real, but evidentially thin unless deliberately tracked, echoing the same output-versus-impact gap the PHVA proposal notes for HTA's own "impact" agenda (Section 2.5).

Indicators	• Veterinary and ASHA worker training completion and, separately, skill retention at follow-up • Peer-reviewed publications, datasets, or policy briefs generated from system outputs • Evidence of uptake by NADCP or IDSP policy or planning processes
Data Requirements	• Training records and follow-up competency assessments • Bibliometric and grey-literature tracing of downstream publications and policy documents citing system data
Quantitative Methods	• Citation and downstream-policy-uptake tracing where available, explicitly reported as low-confidence and attribution-limited rather than as a precise count
Qualitative Methods	• Narrative assessment of whether frontline workers are gaining transferable digital-epidemiology skills, or are functioning primarily as data-entry operators for a system whose reasoning they cannot see or interrogate
Expected Benefits	• Potential to build durable local digital-epidemiology capacity among veterinary and community health workers
Risks / Limitations	• An extractive-data risk: analytical insight and publication credit accruing to external developers or research partners rather than the local health system generating the underlying data • Deskilling if field staff are trained only to act on algorithmic flags rather than to understand the underlying epidemiological reasoning

Domain 7: Cultural and Social Cohesion Value

Contribution to, or erosion of, social trust, shared identity, and community engagement with animal- and public-health institutions.

Materiality (Tier 1): Moderate, and asymmetric: this domain is as likely to register a risk as a benefit for this particular intervention.

Indicators	• Community trust in veterinary and public-health services, before and after deployment • Participation rates in vaccination and testing camps organized around AI-flagged hotspots • Reported stigma associated with a household's herd or village being algorithmically flagged as high-risk
Data Requirements	• Structured community interviews and focus groups in flagged versus non-flagged villages • Camp attendance and refusal records where available
Quantitative Methods	• Engagement and participation metrics where available (camp attendance, repeat testing uptake), reported descriptively rather than as a composite index
Qualitative Methods	• This domain is deliberately not forced into a single number, consistent with Table 2 of the PHVA proposal. A narrative account is required of whether AI-driven flagging is experienced by communities as a locally owned early-warning tool or as an opaque external imposition, since the two produce very different participation outcomes even where the underlying epidemiology is identical
Expected Benefits	• Potential to strengthen trust in animal husbandry and public-health services if flagging and follow-up are conducted transparently and with community involvement
Risks / Limitations	• Stigmatization of flagged villages or tribal communities, particularly if flagging is followed by mandatory culling or movement restrictions rather than voluntary testing and treatment • Erosion of trust if the model's reasoning is not explainable to the communities it flags

Domain 8: Economic Value and Cost Implications

Lifecycle cost, avoided cost, and cost-effectiveness of the intervention, reported separately from, and never substituted for, the other seven domains.

Materiality (Tier 1): High, and directly interfaces with existing HTA cost-effectiveness practice.

Indicators	• Cost per case detected, and cost per DALY averted (an incremental cost-effectiveness ratio, ICER, using the HTA logic reviewed in Section 2.5 of the PHVA proposal) • System procurement, deployment, and ongoing maintenance cost • Avoided economic loss, referenced against the existing national estimate of roughly USD 10.46 million in annual losses from human brucellosis
Data Requirements	• System procurement and maintenance budgets over the evaluation period • Existing national and state-level economic-loss and DALY burden estimates as the counterfactual
Quantitative Methods	• Net present value of avoided cost, and an ICER expressed as cost per DALY averted, benchmarked against standard HTA cost-effectiveness thresholds • Sensitivity analysis of the ICER across plausible ranges of model sensitivity/specificity, since diagnostic performance directly drives the DALY-averted term
Qualitative Methods	• None required for this domain by design; Domain 8 is reported as a standalone monetary account, following the PHVA proposal's explicit rule that a favourable Domain 8 figure must never be presented as if it resolved a weak score on Domain 3 or Domain 4
Expected Benefits	• A plausibly favourable ICER given the scale of the existing documented burden, if diagnostic performance holds up outside the zones in which the model was trained
Risks / Limitations	• Procurement and vendor costs continuing after donor or pilot funding ends • Opportunity cost: the same budget could instead scale up conventional serosurveillance and S19/RB51 vaccination coverage directly, and Domain 8 alone cannot settle which use is preferable — that comparison is exactly what Tier 3 synthesis (Section 4.3) is for

4. The Four-Tier PHVA Process Applied to This Case

4.1 Tier 1 — Materiality Screening

Following the goal-and-scope logic PHVA borrows from ISO 14044 (Section 5.1 of the PHVA proposal), Tier 1 sets the functional unit and system boundary and screens which domains are material enough to justify Tier 2 evidence synthesis. Functional unit: one district-year of AI-assisted zoonotic surveillance coverage. Comparator: continuation of conventional NADCP/IDSP serosurveillance and vaccination without AI-assisted targeting. Geography and time horizon: a representative multi-district rollout, evaluated over a three-year deployment window to capture at least one full monsoon cycle.

Domain	Materiality	Screening rationale
D1 Environmental & Ecological Effectiveness	Low	No physical infrastructure footprint of consequence; retained at low weight for ICT footprint and indirect antimicrobial-use effects only.
D2 Resource & Material Efficiency	Moderate	Testing- and vaccine-targeting efficiency is a genuine, measurable claim of the intervention.
D3 Ecological & Health Safety	High	This is the intervention's core rationale; diagnostic performance and DALYs averted are decision-critical.
D4 Equity & Distributional Justice	High	Target population is disproportionately low-income, occupationally exposed, and digitally under-served.
D5 Resilience & Adaptive Capacity	Moderate–High	System performance under connectivity loss and staff turnover materially affects real-world safety benefit.
D6 Knowledge, Educational & Research Contribution	Moderate	Genuine capacity-building claim, but evidentially thin unless tracked deliberately.
D7 Cultural & Social-Cohesion Value	Moderate	Material as a risk (stigma) as much as a benefit (trust); not reducible to a score.
D8 Economic Value & Cost Implications	High	Direct interface with existing HTA cost-effectiveness practice and funder decision-making.

4.2 Tier 2 Domain-Level Evidence Synthesis

For each material domain, evidence is synthesized using the instrument identified in Section 3 and graded for confidence using GRADE-style categories, exactly as HTA practice already does and as Section 5.2 of the PHVA proposal specifies. Because no primary deployment data exist yet for this illustrative case, the confidence grades below describe the maturity of the underlying evidence base (published brucellosis burden data, published AI-surveillance performance literature) rather than the performance of this specific, as-yet-unbuilt system.

Domain	Evidence base drawn on	GRADE-style confidence
D1	Screening LCA methodology; no system-specific data yet	Very low (system-specific), high (methodology)

Domain	Evidence base drawn on	GRADE-style confidence
D2	NADCP/IDSP testing and vaccination logs (methodology only, pre-deployment)	Low (system-specific), moderate (baseline data)
D3	National DALY burden estimates; published AI-surveillance and AI-diagnostic performance literature	Moderate (burden estimates), low (system-specific diagnostic performance, not yet validated locally)
D4	Published occupational seroprevalence and digital-penetration statistics	Moderate (occupational risk data), low (equity-of-access data specific to this system)
D5	General resilience-assessment literature; no system-specific uptime data yet	Low
D6	General literature on frontline health-worker capacity-building via digital tools	Low, attribution-limited
D7	General literature on community trust and algorithmic flagging in health surveillance	Low, qualitative by design
D8	Published national DALY and economic-loss estimates; comparable AI cost-effectiveness studies in other domains	Moderate (baseline burden), low (system-specific ICER, not yet estimated)

The pattern across the confidence column is itself an important finding: every domain's system-specific evidence is currently low or very low, because the case is illustrative rather than deployed. A genuine PHVA report at this stage would recommend a Tier-2-only, pre-implementation assessment, with Tier 3 synthesis deferred until at least a pilot district-year of real performance data exists — precisely the proportionality principle the PHVA proposal builds in (Section 5.1).

4.3 Tier 3 — Multi-Criteria Synthesis

Where a genuine decision requires comparing this intervention against an alternative use of the same budget (for example, scaling conventional serosurveillance instead), PHVA follows the EVIDEM/MCDA discipline reviewed in Section 2.6 of the PHVA proposal: a deliberative panel — including tribal and landless livestock-keeper representatives, veterinary and public-health experts, and an independent ethicist, not technical experts alone — elicits explicit domain weights using a structured method such as swing weighting, and a sensitivity analysis is run across plausible weight scenarios.

The weights below are illustrative only, constructed to show how the process would work, not as a recommended or validated weighting for this intervention:

Domain	Illustrative panel weight	Rationale for illustrative weight
D3 Ecological & Health Safety	30%	Core rationale of the intervention

Domain	Illustrative panel weight	Rationale for illustrative weight
D4 Equity & Distributional Justice	20%	Reflects the QALY-equity caution built into PHVA design (Section 2.5 of the proposal)
D8 Economic Value & Cost Implications	20%	Funder decision-relevance
D5 Resilience & Adaptive Capacity	12%	Determines whether the safety benefit is realized in practice
D2 Resource & Material Efficiency	8%	Secondary efficiency claim
D6 Knowledge, Educational & Research Contribution	5%	Real but evidentially thin
D7 Cultural & Social-Cohesion Value	5%	Reported narratively; a nominal weight only for illustration
D1 Environmental & Ecological Effectiveness	0%	Screened as immaterial at Tier 1; retained in the dashboard, not the weighted synthesis

A sensitivity analysis across at least three weighting scenarios — a safety-prioritizing scenario, an equity-prioritizing scenario, and a cost-prioritizing scenario — would then be run before any comparative statement is made, and the result reported as a range rather than a single number, following Section 5.3 of the PHVA proposal.

4.4 Tier 4 — Reporting

Every PHVA assessment reports a mandatory, disaggregated dashboard (Section 5 below) rather than collapsing straight to a composite score. A single headline number — in the PHVA proposal's terminology, a provisional "Planetary Health Years" figure — is explicitly not computed here, consistent with Section 5.5 of the proposal, which treats any such figure as a research-stage aspiration requiring independent equity review before use, not a deliverable of a first worked example.

5. PHVA Dashboard: Domain-Level Summary

The dashboard below is the mandatory Tier 4 output. It reports each material domain's finding, confidence grade, and headline direction of effect, without aggregating them into a single score.

Domain	Materiality	Headline finding (illustrative)	Confidence	Direction
D1 Environment	Low	Negligible direct footprint; small possible AMR co-benefit	Very low	Neutral / slightly positive
D2 Resource Efficiency	Moderate	Fewer tests and doses per case, if targeting is well-calibrated	Low	Positive
D3 Health Safety	High	Earlier detection possible; false-negative risk in data-sparse zones	Moderate/Low	Positive, with a flagged risk

Domain	Materiality	Headline finding (illustrative)	Confidence	Direction
D4 Equity	High	Highest-risk occupational groups are also least digitally reached	Low	Contested — risk of exclusion
D5 Resilience	Moderate–High	Untested under monsoon connectivity loss and staff turnover	Low	Uncertain
D6 Knowledge Contribution	Moderate	Plausible capacity-building; attribution not yet demonstrable	Low	Tentatively positive
D7 Cultural Value	Moderate	Trust effect depends entirely on how flagging is implemented	Qualitative	Bidirectional risk
D8 Economic Value	High	Plausibly favourable ICER against a large documented burden	Moderate/Low	Tentatively positive

6. Trade-Offs Between Environmental Sustainability, Public-Health Benefit, Equity, Resilience, and Cost

The dashboard's value, as the PHVA proposal argues (Section 5.4), is precisely what it refuses to hide. Four tensions are visible in this case that a single composite score would tend to average away.

6.1 Health benefit versus equity

Domain 3 and Domain 4 pull in different directions for the same reason: the intervention's diagnostic-safety benefit is largest, in principle, for occupationally exposed, low-income, and tribal populations, precisely because their documented burden is highest — but its delivery mechanism (a smartphone application) is least accessible to exactly that population. A strong D3 score achieved by reaching only well-connected, better-resourced herds would be a distributional failure masquerading as a public-health success, the same pattern the PHVA proposal's Drip-Irrigation sketch (Section 6.1 of the proposal) identifies for a different sector.

6.2 Cost-effectiveness versus resilience

A favourable Domain 8 ICER is easiest to demonstrate for a lean, centralized system with minimal offline redundancy — which is also the weakest possible design from a Domain 5 resilience standpoint. Building genuine offline fallback and graceful degradation into the system is likely to raise its unit cost and therefore weaken its headline ICER, even though it strengthens the intervention's actual, real-world safety value during exactly the network outages most likely to coincide with monsoon-season disease spikes.

6.3 Environmental minimalism versus equity infrastructure

Domain 1's low materiality score is itself contingent on the system remaining a light, software-only layer. Closing the Domain 4 digital-access gap credibly — for example, through subsidized devices or offline data-collection kiosks for landless and tribal households — would very slightly increase Domain 1's environmental footprint (additional hardware, additional embodied emissions) in exchange for a potentially much larger equity gain. A rigid Domain 1 minimalism constraint should not be allowed to foreclose that trade.

6.4 Cultural trust versus algorithmic efficiency

The most efficient way to operate the system, from a pure Domain 2/8 standpoint, is to act automatically on model output with minimal community consultation. That is also the design most likely to produce the Domain 7 stigma and trust-erosion risk identified above. A genuinely defensible implementation is likely to cost more in Domain 2/8 terms — community engagement takes staff time — in exchange for a Domain 7 finding that is positive rather than a documented risk.

None of these tensions has a single correct resolution; PHVA's contribution is to make them visible to whoever ultimately weighs them (Tier 3), rather than resolving them silently inside an aggregation formula.

7. Uncertainty and Methodological Limitations

- Illustrative, not empirical. No primary deployment data exist for this specific system; every system-specific confidence grade in Section 4.2 is low or very low by construction. This worked example demonstrates the PHVA process, not a validated finding about any real AI surveillance system.
- Geographic heterogeneity of the underlying epidemiology. Published seroprevalence estimates vary substantially by agro-climatic zone (from roughly 1.45% to 8.73% within a single state in one study), meaning a model trained and validated in one region cannot be assumed to generalize; Domain 3's confidence grade would need to be assessed zone by zone in a real deployment, not nationally.
- Attribution risk. Any observed reduction in human or livestock brucellosis burden following deployment would be difficult to attribute cleanly to the AI system as distinct from the concurrent, nationally scaled NADCP S19 vaccination campaign already under way; a genuine Tier 2 synthesis would require a comparison arm or staggered rollout design to separate the two.
- Global and within-country inequity in measurement capacity. The PHVA proposal's own limitations section (Section 8) flags that under-monitored regions risk being scored as "low evidence" rather than correctly identified as "high impact, poorly measured." This risk is acute here: the tribal and remote districts most relevant to Domain 4 are also the districts where baseline surveillance and connectivity data are thinnest, so an evidence-graded assessment could systematically under-credit exactly the populations the intervention is meant to prioritize.
- Weighting is a political, not only a technical, choice. The illustrative Tier 3 weights in Section 4.3 are for demonstration only; a genuine application would require the same participatory weight-elicitation exercise, including affected tribal and landless livestock-keeping communities directly, that Section 8 of the PHVA proposal identifies as unresolved even in principle.

- Goodhart's Law / gaming risk. If Domain 3 sensitivity/specificity or Domain 8 ICER figures become tied to continued funding, there is a documented risk — already noted for LCA system-boundary selection and ESG scoring more broadly (Section 8 of the PHVA proposal) — that reporting boundaries or validation datasets could be selected to favour a good score rather than an accurate one.
- Double-counting risk. Care would be required in a real assessment to avoid counting the same avoided disease burden under both Domain 3 (as a safety finding) and Domain 8 (once monetized as an avoided cost), consistent with the general double-counting caution in Section 8 of the PHVA proposal.
- No headline metric is computed. Consistent with Section 5.5 of the PHVA proposal, this worked example deliberately stops at the disaggregated dashboard and does not compute a "Planetary Health Years" or equivalent composite figure, since doing so before Tier 3 weighting has been validated would repeat exactly the premature-aggregation failure mode the PHVA proposal is designed to avoid.

8. Conclusion

This worked example shows what the PHVA proposal's methodology section describes in the abstract: an eight-domain, four-tier process that can hold the environmental footprint, diagnostic safety, equity of access, system resilience, workforce capacity, community trust, and cost-effectiveness of a single, real intervention inside one auditable evaluation, without collapsing them into a number before their trade-offs have been examined. For this particular case, the dashboard's most decision-relevant finding is not any single domain score but the tension between Domain 3 and Domain 4: an intervention whose public-health case is strongest for the populations it is currently least likely to reach through a smartphone-dependent delivery channel. Surfacing that tension, rather than resolving it prematurely inside an aggregate score, is precisely the function PHVA is designed to perform.

References

Godfroid, J. (2017). Brucellosis in livestock and wildlife: zoonotic diseases without pity. Microbes and Infection, 19(3), 158–160.

International Organization for Standardization. (2006). ISO 14040:2006 — Environmental management — Life cycle assessment — Principles and framework.

International Organization for Standardization. (2006). ISO 14044:2006 — Environmental management — Life cycle assessment — Requirements and guidelines.

ISPOR MCDA Emerging Good Practices Task Force. (2016). Multiple criteria decision analysis for health care decision making — An introduction: Report 1. Value in Health, 19(1), 1–13.

ISPOR MCDA Emerging Good Practices Task Force. (2016). Multiple criteria decision analysis for health care decision making — Emerging good practices: Report 2. Value in Health, 19(2), 125–137.

One Health High-Level Expert Panel (OHHLEP). (2021). Operational definition of "One Health." World Health Organization.

O'Rourke, B., Oortwijn, W., & Schuller, T. (2020). The new definition of health technology assessment: A milestone in international collaboration. International Journal of Technology Assessment in Health Care, 36(3), 187–190.

Quadripartite (FAO, UNEP, WHO, & WOAH). (2022). One Health Joint Plan of Action (2022–2026).

Raworth, K. (2017). A doughnut for the Anthropocene: Humanity's compass in the 21st century. The Lancet Planetary Health, 1(2), e48–e49.

Richardson, K., Steffen, W., Lucht, W., et al. (2023). Earth beyond six of nine planetary boundaries. Science Advances, 9(37), eadh2458.

Rockström, J., Steffen, W., Noone, K., et al. (2009). A safe operating space for humanity. Nature, 461, 472–475.

Schneider, P. (2022). The QALY is ableist: On the unethical implications of health states worse than dead. Quality of Life Research, 31, 1545–1552.

Singh, B. B., Khatkar, M. S., Aulakh, R. S., Gill, J. P. S., & Dhand, N. K. (2018). Estimation of the health and economic burden of human brucellosis in India. Preventive Veterinary Medicine, 154, 148–155.

Sinclair, S. (2012). How to avoid unfair discrimination against disabled patients in healthcare resource allocation. Journal of Medical Ethics, 38(3), 158–162.

Sinclair, S. (2018). Improved health state descriptions will not benefit disabled patients under QALY-based assessment. Journal of Medical Ethics, 44(11), 797–798.

Whitmee, S., Haines, A., Beyrer, C., et al. (2015). Safeguarding human health in the Anthropocene epoch: Report of The Rockefeller Foundation–Lancet Commission on planetary health. The Lancet, 386(10007), 1973–2028.

World Health Organization / South-East Asia Regional Office data cited in: Dean, A. S., et al. Epidemiology of brucellosis in cattle and dairy farmers of rural Ludhiana, Punjab. PLOS Neglected Tropical Diseases.