

Reassessing the economic impact of reaching Universal Health Coverage

An audit of the 2019 WHO EPIC estimates

Jeremy A. Lauer
Unexia Association

18 September 2026

Abstract

In a 2019 analysis, Han, Arcand and Lauer linked projected health improvements (‘achieving UHC by 2030’) to economic output through an augmented Solow model. This response, by a co-author, reassesses that contribution using archived manuscripts, spreadsheets and source exports. Alongside an inconsistent capital valuation and a questionable balanced-growth calibration, this audit also establishes a disease-category mismatch in the archived material: Across all 120 inspected country–year–scenario inputs, the series labelled ‘NCD deaths’ reproduces the source workbook’s all-cause deaths series, not deaths that are NCD specific. In several countries, most of the difference in all-cause deaths between the baseline and ambitious UHC scenarios occurs among children under five, and cannot be interpreted as deaths averted through NCD interventions. Thus, using this difference to infer reductions in working-age NCD mortality and disability can produce productivity gains not supported by the underlying data. Holding reconstructed health inputs and archived spending assumptions constant across simulations, the alternative specifications tested here do not reproduce the magnitude of published GDP benefit. Under the preferred reconstruction, the ten countries show US\$223 billion of additional GDP during 2019–2023, or 0.19% of cumulative baseline GDP, compared with the reported US\$578 billion and 0.54%. Nine countries show positive gains in the reconstruction, compared with 10 in the original paper. These findings warrant qualification of policy-facing claims about the spending–health–growth chain and also show the importance of the establishment of a clear chain of provenance, reproducibility, auditability and contestability for complex model estimates.

Keywords: universal health coverage; economic impact; capital valuation; Solow model; reproducibility; health investment.

JEL classifications: I15, I18, O41, C63.

1 Introduction

The problem

Better health can preserve lives, reduce disability, release time spent providing care and increase productive capacity. Relating these effects to gross national output is nevertheless demanding: epidemiological scenarios must be translated across populations, age groups and causes of illness; productive inputs must be valued consistently; and the resulting counterfactual must be distinguished from both an economic prediction and a full-income welfare assessment.

The analysis published in 2019 [1] examined whether the additional health expenditure required to reach UHC could “pay for itself” in the limited sense that its increase in effective labour would exceed the loss of output associated with lower capital accumulation. In that construct, additional health expenditure reduces investment by an assumed share of spending; lower investment slows the accumulation of physical capital. Better health reduces mortality and disability and, through specified age, labour-force-participation and caregiver assumptions, increases effective labour. The estimated GDP effect is the balance between these two effects.

That balance cannot be observed directly. The augmented Solow production function assigns output shares to physical capital and effective labour through assumed production elasticities. Country savings rates and depreciation rates are taken from the archived macroeconomic inputs. The primary reconstruction assumes, as in the original paper, that 10% of incremental health expenditure displaces investment. It holds the production relationship, including productivity and factor elasticities, constant over the projection horizon. It also assumes that the modelled health improvements can be translated into changes in effective labour through mortality, disability, labour-force participation, age-allocation and caregiver-time rules.

The relationship between spending and health outcomes is itself modelled rather than observed. It draws on the OneHealth Tool scenario framework and the resource-needs and health-outcome projections reported by Stenberg and colleagues in *The Lancet Global Health* (2017) [2], not on realised country-specific responses to expenditure.

The attraction of this approach is parsimony. It uses internationally available inputs and can be applied beyond the original ten countries without the data requirements and calibration burden of a full computable general-equilibrium model. Its results are transparent and reproducible, but conditional: changing the health scenario, the morbidity-to-work mapping, the investment-displacement assumption, the initial capital stock, the valuation basis or the production parameters changes the estimated GDP gain. A companion paper applies a common health-response scenario to a broader group of 111 countries [3]; the present paper focuses on the audit of the 2019 analysis.

The audit reported here identifies several problems in the 2019 exercise, as well as modelling choices that require clearer justification. Under the reconstructed inputs and a consistent capital valuation, the estimated gains are materially smaller. They do not support a general claim that UHC expenditure will pay for itself through short-horizon GDP gains. They do support the narrower finding that better health can increase productive capacity: in the preferred reconstruction, nine of ten countries have positive cumulative GDP differences. This finding does not establish fiscal self-financing or a social return on investment. A comprehensive welfare analysis would also value longer lives, better health, financial protection and distributional effects within an explicit table of accounts.[4]

The reference point

Han, Arcand and Lauer's December 2019 SSRN paper addressed the narrow question of increased health expenditure to meet UHC targets and resulting economic gains for ten countries over 2019–2023.[1] Its framework connected Universal Health Coverage scenarios to a human-capital-augmented Solow model, using international demographic, disease-burden, labour and macroeconomic sources. The authors deserve credit for undertaking an ambitious project across a heterogeneous collection of legacy tools and platforms. The ten-country exercise was a pilot within a broader programme of work. Neither the scale of that programme nor the practical challenges of assembling it should be understated.

The present author was the supervising author for WHO and takes full responsibility for the weaknesses of the 2019 paper. This response is a reassessment of shared work, prompted by a subsequent audit. Its purpose is to retain the conceptual framework while making the numerical evidence more defensible.

The policy stakes are significant. WHO's investment-case report, published in 2018 for the 2019–2023 programme, prominently displayed “2–4% economic growth”.[5] The 2019 paper states that its work directly supported that investment case. The quantitative work reported in 2019 and the policy communications from 2018 both belonged to a related programme of work conducted at the WHO, but they use different versions, coverage and aggregation choices.

2 Evidence and scope of the reassessment

The evidence comprises the December 2019 manuscript, a related 2021 Lancet Global Health submission, the archived EPIC UHC spreadsheets, childhood and non-communicable disease (NCD) source exports, disease-summary files, and the Penn World Table (PWT) 9.0 documentation and capital detail.[1, 6, 7, 8] Table 5 of the 2019 paper was checked directly: its country totals and cumulative percentages match the comparator used here. A one-dollar discrepancy in summing Nigeria’s rounded annual entries has no bearing on the reported results.

The R implementation used to produce the two manuscripts was not available for execution. This response does not claim to reproduce that implementation, identify every operation it performed, or establish which spreadsheet errors it inherited. The independently executable reconstruction implements the disclosed economic structure with explicit repairs and synthetic health-input matching.

The exercise preserves historical input vintages wherever possible. It does not substitute today’s GDP observations for the projected macroeconomic setting, re-estimate every country production function, or claim to evaluate, with the benefit of hindsight, what actually happened to economic output during 2019–2023. Its estimand is the difference between two internally specified scenarios under a shared historical macroeconomic framework.

Evidence is strongest for the arithmetic and source correspondences that can be inspected directly. It is weaker for the productivity interpretation of health changes and for the equivalence of reconstructed scenarios to the final OneHealth projects. Throughout, verified source defects, mathematical implications and conditional simulation results are kept separate.

3 Capital must share the valuation basis of output

3.1 Why initial output matching does not solve the problem

Write country output, suppressing the country index, as

$$Y_t = \gamma A_t K_t^\alpha (H_t L_t)^{1-\alpha}, 0 < \alpha < 1. \quad (1)$$

Here K is physical capital, L effective labour, H the schooling-related human-capital index, A total factor productivity and γ a level normalisation. The latter is set once to match baseline GDP in the initial year and is shared across scenarios.

The archived capital input draws on PWT 9.0’s ck , a capital stock at current PPPs, expressed in millions of 2011 US dollars. The country output and health-cost series used for the comparison are in constant 2014 US dollars on a market-exchange-rate basis.[7, 8] Labour is measured in people or productivity-weighted worker equivalents; it is not itself a dollar-valued PWT stock. The immediate incompatibility concerns the monetary basis of capital, output and investment flows.

A change in the units of K_0 can be offset by γ in equation (1), leaving initial output unchanged. This does not make a mixed-unit accumulation equation valid. With saving rate s , depreciation δ , health expenditure C and investment-displacement share r , the relevant diagnostic is

$$\frac{K_{t+1}}{K_t} = 1 - \delta + s \frac{Y_t}{K_t} - r \frac{C_{t+1}}{K_t}. \quad (2)$$

The ratios must use a common valuation basis. Rescaling the stock while leaving investment flows unchanged changes the economic transition, not merely its notation. An initial GDP fit can therefore conceal a consequential capital-to-output mismatch.

3.2 An explicit conversion and date bridge

For the 2014 anchor, the implemented conversion is

$$K_{2014}^{\text{USD},2014} = 10^6 \text{ck}_{2014} \text{pl_k}_{2014}. \quad (3)$$

The capital-specific price-level variable is essential; a GDP price level or a general inflation adjustment is not an equivalent operation. At this anchor the conversion was checked against the sum of PWT's asset-level capital values divided by the exchange rate. Across the ten countries, relative discrepancies were below 1.4×10^{-7} .

Because PWT 9.0 ends in 2014, the stock is advanced to 2018 using observed GDP inputs and the archived saving and depreciation parameters:

$$K_{t+1} = (1 - \delta) K_t + s Y_t^{\text{obs}}, t = 2014, \dots, 2017. \quad (4)$$

The resulting stock is empirically anchored but model-bridged; it is not a direct observation of capital in 2018. Fixed relative investment prices are assumed after the anchor.

A controlled diagnostic, separate from the preferred final reconstruction, held the demographic and macroeconomic assumptions fixed and changed the capital basis. With the experience function, India's projected aggregate GDP fell by 2.57% between 2018 and 2023 when the unconverted PPP-valued capital stock was used, but rose by 25.87% with the converted, date-bridged stock. No country in the latter ten-country diagnostic had an annual aggregate GDP decline. This establishes that the valuation mismatch is sufficient to generate an apparent decline in GDP in at least one reconstructed case.

4 Balanced growth is a benchmark, not a measurement of initial capital

The 2019 paper's Appendix B explains that projected GDP or GDP per capita declined under an earlier calibration and introduces a capital level associated with balanced growth.[1] In the paper's notation, with $k = K/(HL)$, the expression is

$$k_t^{\dot{c}} = \left(\frac{sy A_t}{n+h+\delta} \right)^{1/(1-\alpha)}, \quad (5)$$

where n and h denote labour and human-capital growth. The theoretical motivation is understandable: a stable transition can be useful when a short projection is otherwise dominated by questionable initial conditions. It is also legitimate to use a balanced-growth benchmark as a sensitivity case, or to constrain latent initial capital within an explicitly estimated measurement model.

The difficulty is treating that benchmark as an empirical correction merely because a projected path declines. An economy can be above or below its steady-state capital intensity, and imposing the steady state removes the transition one is trying to estimate.

There is also a technical qualification. Equation (5) is a frozen-technology steady-state expression if A is constant. Under the Hicks-neutral production function in equation (1), a growing A_t makes the balanced-growth target move. In continuous time, defining $E = A^{1/(1-\alpha)} HL$ and $\tilde{k} = K/E$ gives the corresponding stationary benchmark

$$\tilde{k}^{\dot{c}} = \left[\frac{sy}{\delta+n+h+g_A/(1-\alpha)} \right]^{1/(1-\alpha)}, \quad (6)$$

under constant growth rates and without the health-cost term.

A further identification issue arises when capital and the output normalisation are calibrated together. Combining the frozen-technology condition with the initial GDP identity implies

$$K_0^i = \frac{s Y_0}{n+h+\delta}. \quad (7)$$

Capital is then determined by GDP and assumed rates, rather than independently measured.

Declining GDP per person is also weak evidence of a calibration failure when an intervention retains survivors and enlarges population. Even aggregate GDP need not rise in every admissible transition. In this audit, consistent empirical valuation removes the declining aggregate path in the controlled ten-country diagnostic without imposing balanced growth. It therefore supplies a sufficient practical reason to prefer the empirical anchor for the reconstruction.

The controlled experiments reported in Section 7.4 separate these capital choices from the revised health and spending inputs. Under those fixed inputs, the tested legacy capital alternatives yield lower aggregate gains than the preferred empirical anchor. They therefore do not explain the larger published benefit.

5 Health inputs and verified defects in the matched reconstruction

5.1 What the files establish

The source audit identified several discrepancies whose direction and location can be checked directly. [8, 9, 10, 11]

First, all 120 country–year–scenario values imported as NCD deaths for 2018–2023 match the source workbook’s broader total-deaths row. A contrast between these totals cannot identify the NCD intervention effect. In 2023, under-five deaths account for 87.1% of the imported difference in Nigeria, 93.6% in India and 96.4% in Pakistan; Table 3 documents the ten-country pattern. Children can die from NCDs, but the total row is not disease specific.[8, 9]

Second, Pakistan’s diarrhoea baseline uses neonatal deaths while its ambitious arm uses post-neonatal deaths in all six cached years, 2018–2023. In 2019 the imported values, 3,746 and 9,886, imply 6,140 additional deaths. The source post-neonatal baseline is 35,586; including neonatal deaths in both arms gives 39,332 versus 10,807, or 28,525 deaths averted before reference-year matching. This understates the benefit attributable to reduced diarrhoeal disease and cannot be invoked as an explanation for inflated gains.[8, 9]

Third, Angola’s imported baseline post-neonatal pneumonia series equals its ambitious series in every cached year, 2018–2023. For 2019 both imported arms are 19,664, although the flatline source gives 25,701. For 2023 the corresponding figures are 15,703 versus a true flatline value of 29,545. The import therefore erases the post-neonatal benefit in those comparisons. The preferred construction corrects it. A bounded comparison of the malaria, diarrhoea and pneumonia imports in the other audited countries found no further category-import mismatch of these kinds.[8, 9]

Fourth, childhood exports describe under-five outcomes and come from different Spectrum versions embedded in the OneHealth tool. The paper itself explicitly restricts childhood effects to ages 0–4 and 5–14. Restricting the reconstructed effects to under-fives is supported by the available export; extending them to ages 5–14 would require an additional bridge. An all-age childhood variant examined during reconstruction was rejected.

Finally, legacy file versions contain an experience-function sign issue and other age/sex reference problems. The corrected experience profile used here is

$$\rho(a) = 0.8 + 0.02(a - 15) - 0.0002(a - 15)^2. \quad (8)$$

The manuscript prints the concave form, and the archived UHC workbook's version history records an earlier correction.

5.2 A common mortality baseline and intervention contrast

The reconstruction places both arms on a common country-year population, age/sex structure and GBD cause-specific baseline. The archived GBD population distribution is scaled to the macro population total. Health effects are then transported to this common exposure; extra population arising from survivors is handled subsequently. The aim is internal comparability, not recovery of unavailable OneHealth project files. Stenberg and colleagues' modelling programme supplies the broader scenario context.[2]

For childhood disease, HIV and tuberculosis, let B_{dt}^s and A_{dt}^s denote source baseline and ambitious deaths for disease d , and let $r_0 = 2015$ be the common reference year. The implemented contrast is

$$R_{dt} = \frac{A_{dt}^s / B_{dt}^s}{A_{dr_0}^s / B_{dr_0}^s}, \quad (9-10)$$

$$f_{dt} = \max\{0, 1 - R_{dt}\}, D_{gdt} = B_{gdt}^{common} f_{dt},$$

where g indexes age and sex and D denotes directly averted deaths. Childhood fractions apply only to ages 0–4. Both source arms include neonatal and post-neonatal categories on the same basis.

This common-reference operation removes initial level differences. Six early malaria country-year cells, in Nigeria and Ethiopia, still imply harm after rebasing. The non-harm bound is an explicit synthetic assumption, not an empirical finding about intervention efficacy. Keeping those signed discrepancies changes cumulative GDP gains by only US\$0.044 billion. For China's and Brazil's zero flatline malaria series, the ambitious proportional decline from the reference year is transferred to the small positive GBD baseline, avoiding an invented denominator floor.

For NCDs, paired no-scale-up and scale-up blocks within the same ambitious-population file replace the imported total. For component/age row j ,

$$D_{jt}^{NCD} = B_{jt}^s - A_{jt}^s \frac{B_{jr_0}^s}{A_{jr_0}^s}. \quad (11)$$

The 16 component contrasts are allocated to the common age/sex groups using GBD NCD death weights and proportional age-band overlap. Zero reference cells are handled explicitly in the code. Original population denominators are unavailable, so this is transport of absolute counts, not exact standardisation of source hazards. Non-targeted NCD mortality is left unchanged in aggregate. A population-weighted allocation is examined separately.

5.3 Morbidity, work and survivor retention

Following the EPIC approach, matched mortality reductions are used to scale baseline years lived with disability (YLD):

$$\Delta YLD_{gdt} = YLD_{gdt}^B \frac{D_{gdt}}{B_{gdt}^{common}}. \quad (12)$$

For NCDs this applies a broad NCD mortality fraction to broad NCD YLD. Reliable, complete component-specific morbidity responses are not available in the retained source cache. This is a major unresolved structural assumption. A YLD is neither observed work absence nor a directly measured productivity-year. The one-for-one translation before labour participation and experience adjustment should therefore be read as an assumed mechanism.

Working ages are 15–69. Adult gains are weighted by archived labour-force participation and equation (8). Caregiver equivalents are one year per under-five YLD and half a year per YLD at ages 5–14, weighted by pooled male/female participation at ages 30–44. The broad older-age participation crosswalk retains a nearest-band approximation.

Saved lives remain in the model. If S_t^{all} is the previous survivor stock after ageing, q_t^A the intervention all-cause death risk on the common population, and D_t new directly averted deaths, then, componentwise,

$$S_t = S_t^{\text{all}} (1 - q_t^A + D_t) \Delta P_t^{\text{years}} \quad (13-14)$$

The risk includes remaining deaths from the targeted diseases: survivors are not permanently immune. Cohorts advance in single years, enter work at 15 and leave at 70; ages within broad input groups are initially uniform. Additional disability and care needs among survivors are deducted. Morbidity gains are annual flows and are not accumulated as stocks. Births and migration induced by the intervention are omitted.

Health exposure begins in 2018, including improvements already reached that year relative to the 2015 reference. No saved-life stock from 2016–2017 is imported. This distinguishes the effect-size reference year from the economic exposure window.

6 Economic specification and validation

The primary simulation uses common empirical initial capital and a single baseline output anchor. In annual notation, for $t \geq 2019$,

$$\begin{aligned} K_t^B &= (1 - \delta) K_{t-1}^B + s Y_{t-1}^B, \\ K_t^A &= (1 - \delta) K_{t-1}^A + s Y_{t-1}^A - r (C_t^A - C_t^B), r = 0.1. \end{aligned} \quad (15-16)$$

Country-specific savings, depreciation, capital elasticity, human capital and productivity assumptions are retained from the archived inputs. Saving reflects baseline expenditure, so only additional health spending displaces investment in the primary case.

The ambitious scenario output is calculated from its changed labour input; it is not normalised back to baseline. Archived incremental expenditure is zero in 2018, so there is no omitted differential initial capital charge. Costs of earlier scale-up have not been reconstructed. Accordingly, the exercise compares flows over a defined window rather than estimating the full historical return on all resources required to reach the intervention state.

The headline measures are

$$G = \sum_{t=2019}^{2023} (Y_t^A - Y_t^B), p = 100 \frac{G}{\sum_{t=2019}^{2023} Y_t^B}. \quad (17)$$

These are undiscounted cumulative output differences and their baseline-weighted percentage. They are not annual growth rates. Per-person results use the additional retained population in the ambitious denominator.

The executable checks cover positive capital and labour, bounded disease reductions, age allocation, survivor accounting and cohort boundaries at ages 14/15 and 69/70. A null intervention gives zero gains; a cost-only intervention cannot raise GDP. Independent checks of the incremental capital budget and production ratios verify the numerical implementation. For example, common technology and human capital imply

$$\frac{Y_t^A}{Y_t^B} = \left(\frac{K_t^A}{K_t^B} \right)^\alpha \left(\frac{L_t^A}{L_t^B} \right)^{1-\alpha} . \quad (18)$$

These checks establish internal consistency of the specified calculation.

7 Results

7.1 Country differences and aggregate magnitude

Table 1 compares the published country gains with the matched reconstruction. The latter gives US\$223 billion of additional output, against cumulative baseline GDP of US\$120 trillion, or 0.19%. The absolute gain is 38.6% of the published US\$577 billion.

Table 1. Additional GDP, 2019–2023: published estimates and reconstruction.

	2019 paper		Audited reconstruction	
Country	US\$ bn	%	US\$ bn	%
India	102.48	0.64	29.15	0.185
China	45.78	0.07	118.10	0.154
Nigeria	129.19	3.06	9.54	0.232
Pakistan	71.18	4.34	6.03	0.355
Indonesia	59.22	0.89	17.19	0.253
DRC	10.81	3.72	1.43	0.507
Bangladesh	29.09	2.13	6.21	0.445
Ethiopia	10.38	1.11	-0.32	-0.076
Angola	59.94	5.74	2.12	0.254
Brazil	59.39	0.50	33.37	0.277
Ten-country total	577.46	0.54	222.82	0.186

Constant 2014 US dollars; undiscounted sums of annual GDP gains. Published percentages are the paper’s reported cumulative statistics (Table 5; aggregate percentage from its summary), not recomputed on the reconstructed baseline. Proxy percentages follow equation (17). DRC: Democratic Republic of the Congo. Different baseline paths and synthetic health matching limit direct attribution of differences to any single correction.

Nine countries retain positive cumulative gains. Ethiopia’s small negative difference reflects the assumed investment displacement exceeding the short-horizon labour benefit. China has a larger gain in the reconstruction than reported in the paper, so the present reassessment is not a uniform downward rescaling. Country rankings change substantially.

The health accounting yields 9.2 million directly averted deaths during 2019–2023, or 8.742 million averted deaths net of subsequent deaths among retained survivors, and 64.52 million directly averted YLD. These quantities are synthetic scenario outputs. Table 5 expresses the corrected direct contrasts relative to their explicit baseline denominators.

7.2 How much depends on unresolved assumptions?

Table 2 presents selected alternative specifications. Each changes a stated assumption relative to the primary specification.

Table 2. Selected structural sensitivities, 2019–2023.

Specification	GDP gain, US\$ bn	Gain, %
Primary reconstruction	222.82	0.186
Health effects lagged	175.73	0.146
Exclude 2018 health exposure	212.44	0.177
Half the morbidity response	119.66	0.100
No NCD morbidity benefit	64.84	0.054
Mortality only, all incremental costs	16.40	0.014
No caregiver benefit	202.23	0.169
NCD allocation by population	238.65	0.199
2017 reference (different contrast)	157.63	0.131
50% investment displacement	166.72	0.139
No investment displacement	236.83	0.197
Full-cost financing convention	222.42	0.186
80% mortality-effect intensity	181.49	0.151
120% mortality-effect intensity	262.42	0.219

Each case changes one primary assumption. Full results and definitions are in the computational supplement.

Morbidity is the dominant unresolved mapping. Halving its response gives US\$119.656 billion; removing NCD morbidity gives US\$64.839 billion. A mortality-only case, still bearing the full incremental cost, gives US\$16.402 billion.

7.3 Capital assumptions with corrected health and spending held fixed

Nine capital-initialization variants were evaluated for all ten countries, retaining identical revised mortality, morbidity, survivor, labour and cost inputs. Initial baseline GDP is held common; the output normalization is recalculated once for each capital initialization and then shared by both scenarios. Reintroducing unconverted PPP capital with the same 2014–2018 bridge lowers cumulative additional GDP from US\$222.824 billion to US\$198.314 billion. Joint calibration using equation (7) gives US\$194.506 billion; using the archived “optimal” stock gives US\$182.775 billion. Each arm has its own subsequently generated baseline path.

Across the nine specified variants, gains range from US\$178.996 billion to US\$243.307 billion, well below the published GDP gain of US\$577.463 billion. Under equation (7), empirical-capital valuation drops out once initial capital and normalization are jointly solved. These results establish that the tested capital choices do not explain the larger published benefit with corrected health and spending held fixed.

7.4 Corrected health contrasts and their denominators

Table 3 reports the preferred matched health and expenditure contrasts, using ratios of sums over 2019–2023. These values already incorporate the mortality-import corrections. In particular, Nigeria’s 13.1% direct mortality reduction is compatible with its preferred US\$9.54 billion GDP gain.

Table 3. Preferred expenditure and health contrasts, 2019–2023.

Country	Spending increase	All-cause deaths reduction	Modelled morbidity reduction
India	14.08%	5.38%	3.83%
China	2.84%	3.47%	3.45%
Nigeria	18.41%	13.10%	4.78%
Pakistan	67.90%	7.35%	4.27%
Indonesia	20.77%	3.82%	3.61%
DRC	160.77%	12.20%	3.91%
Bangladesh	36.50%	5.14%	4.08%
Ethiopia	113.67%	7.52%	3.78%
Angola	26.00%	14.07%	5.00%
Brazil	1.14%	3.35%	3.34%

Spending: additional modelled cost divided by baseline modelled cost, not GDP or a newly retrieved national expenditure total. Mortality: directly averted deaths across the six disjoint disease groups divided by all-cause baseline deaths, on the common standardized population. Morbidity: modelled YLD averted divided by baseline YLD within the six modelled disease groups, not all-cause disability. Each statistic is a ratio of five-year totals. Morbidity is inferred under equation (12), not independently observed; the differing denominator scopes prevent direct interpretation of the columns as comparable elasticities.

Deaths subsequently occurring among retained survivors are excluded from the direct contrast in Table 3 but included in the economic model. Remaining disability among survivors is likewise deducted when calculating net labour benefits. A lower net reduction reflects the additional people alive and exposed to risk.

8 Policy interpretation and a constructive research standard

The broad proposition that health improvements can raise productive capacity survives this exercise. Under the primary assumptions, nine of ten countries have positive output differences. The published magnitude, country ordering and a strong short-horizon self-financing interpretation do not survive as robust conclusions.

WHO’s prominently reported 2–4% figure needs qualification.[2] The revised 0.19% is *not* a replacement estimate for all countries or all components of WHO’s programme. It is a baseline-weighted ten-country result under a reconstructed intervention, whereas the policy materials in [2] use a distinct aggregation.[12]

The constructive response is to revise the evidential status of the numerical claims and strengthen the computational chain. Four requirements are inseparable.

1. *Provenance* records each source version, scenario, unit, age group, disease definition and transformation.
2. *Reproducibility* lets another analyst regenerate the tables from frozen inputs using a documented environment.

3. *Auditability* renders legible intermediate quantities and checks that can locate a failure.
4. *Contestability* allows a reader to replace uncertain assumptions and see their consequences without reverse-engineering the entire model.

For an impact estimate of the kind at stake here, that chain should include the source epidemiological project or a sufficient export; the population denominators behind counts; explicit capital price and exchange-rate conventions; the initialisation sequence; mappings from causes and ages to labour; and a register separating repairs from discretionary assumptions. A headline number should link to a versioned result and its uncertainty analysis. A spreadsheet, an R script or a Python package can each meet this standard.

This is an attainable standard, not a demand for perfect data prior to analysis. The accompanying implementation uses simple tools and makes the most consequential substitutions visible. Its limitations identify the next priorities: recover final source population denominators and component morbidity outputs; estimate or externally validate morbidity-to-work responses; examine savings and investment accounting; re-estimate production parameters under consistent historical measurement; and extend the horizon with fuller demography.

9 Limitations and conclusion

The reconstruction is conditional on inherited macroeconomic parameters and historical GDP projections. The 2018 capital-stock bridge, fixed relative investment price, domestic-saving assumption and simplified financing mechanism remain approximations. The common GBD baseline may incorporate secular health changes and is not a literal rerun of the OneHealth flatline population model.

Health matching remains synthetic. It combines source vintages, transports NCD counts without their original population denominators, and applies a broad morbidity proxy. Age-band overlap and labour participation use approximations. Survivor retention is explicit, but demographic feedback, employment constraints and wider general-equilibrium responses are not. The controlled reversals identify the consequences of specified mappings within the executable reconstruction.

Subject to these limits, the audit supports qualification of the 2019 numerical headline. Valuation and empirical initialization remain necessary for a coherent model, but the controlled tests do not support treating them as the explanation for the larger published gain. The preferred matched estimate remains US\$222.8 billion and supports positive productive effects in nine countries. It should be presented as a conditional estimate within a family of results whose empirical anchors, synthetic bridges and disputed mechanisms are visible and replaceable.

The companion extension uses the corrected economic foundation but applies a common pooled health-response scenario; its broader-country results are reported separately and should not be read as a validation of the historical estimates.[3]

Acknowledgements

The author acknowledges Seoni Han, Jean-Louis Arcand and the WHO modelling teams for undertaking the original integration of epidemiological and economic evidence. OpenAI ChatGPT and Codex were used to assist with source extraction, code development, computational checks, drafting and typesetting during this audit. The author retains responsibility for the manuscript's interpretation, claims and any remaining errors.

Author relationship to the original work

Jeremy A. Lauer was a co-author of the 2019 paper under the name Jeremy Addison Lauer and contributed to the related WHO technical work. This response represents the author's reassessment; it should not be taken necessarily to express the views of the other original authors (Arcand and Han) or the WHO.

Data and code availability

The accompanying computational supplement contains frozen derived inputs, Python routines, validation checks, country and sensitivity tables, and the editable manuscript source. The main simulation requires Python 3.10 or later and only its standard library; it does not require Excel, R, network access or access to the unavailable historical R implementation. Original project files are identified in Appendix A and linked from the supplement. Access to the project archive may require permission. The supplement reproduces this reconstruction; it does not recover the original epidemiological projects or establish public access to every underlying source. The broader-country extension has a separate run kit and provenance record.

References

- [1] Han S, Arcand J-L, Lauer JA. *Economic Impacts of Investing in Universal Health Coverage for Low- and Lower-Middle Income Countries*. Manuscript dated 12 December 2019. SSRN 3343767. [SSRN record](#); [audited manuscript copy](#). See Table 5 and Appendices A–C.
- [2] Stenberg K, Hanssen O, Edejer TT-T, et al. Financing transformative health systems towards achievement of the health Sustainable Development Goals: a model for projected resource needs in 67 low-income and middle-income countries. *Lancet Global Health*. 2017;5:e875–e887. [doi:10.1016/S2214-109X\(17\)30263-2](#).
- [3] Lauer JA. Extending health expenditure and economic output projections to 90 countries. Companion manuscript. 2026.
- [4] Lauer JA, Morton A, Culyer AJ, Chalkidou K. What counts in economic evaluations in health? Benefit-cost analysis compared to other forms of economic evaluations. Geneva: World Health Organization; 2020. WHO/UHC/HGF/HEF/WorkingPaper/20.2. <https://www.who.int/publications/i/item/9789240008878>
- [5] World Health Organization. *A Healthier Humanity: The WHO Investment Case*. Brochure for the 2019–2023 programme. Geneva: WHO; 2018. WHO/DGO/CRM/18.1. [Official brochure](#), second panel/page.
- [6] Arcand J-L, Han S, Lauer JA. UHC economic-impact manuscript submitted to *The Lancet Global Health*, 2021. Unpublished submission, archived project copy. [Manuscript](#); Table 5 and methodological appendix.
- [7] Groningen Growth and Development Centre. *Penn World Table 9.0: data, user guide and capital detail*. [Release page](#); [user guide](#); [capital-detail data](#). Variables used for the valuation check: ck, pl_k, exchange rate and asset capital values.
- [8] EPIC project archive. UHC economic model workbook, version 0.7. [Archived workbook](#). Macro inputs, GBD burden, costs, formula and version-history evidence; original source-file access may be restricted.
- [9] EPIC/WHO source exports. [Flatline childhood deaths](#) and [ambitious childhood deaths](#). Archived Spectrum exports; neonatal and post-neonatal cause rows.

- [10] EPIC/WHO source export. [NCD mortality, ten countries](#). Country sheets, total-deaths row and paired component blocks beginning in columns T and AL.
- [11] EPIC/WHO source exports. [Disease summary v3](#); [HIV source](#); [tuberculosis source](#). Archived country/scenario series.
- [12] World Health Organization. *Investing global, investing local: supporting value for money towards the health SDGs*. Department of Health Systems Governance and Financing; draft July 2018. [Technical report](#), especially Annex 2.

Appendix A Audit register and computational companion

Issue and evidence	Operation in this reconstruction	Evidential limit
Capital basis: PWT and archive [5,7]	Multiply 2014 ck by capital price level and 10^6 ; verify against asset values/exchange rate; bridge to 2018.	Common currency is verified; 2018 stock remains model-based. Fixed relative prices and GDP-vintage differences remain.
Initial calibration: paper Appendix B [1]	Retain empirical capital; calibrate output normalisation once; no steady-state replacement.	Published formula and motivation are known; full historical R execution order is not.
NCD totals [7,9]	Replace all-cause total contrasts with 16 paired NCD component contrasts; transport counts to common baseline.	All 120 imported values match the total row. Final R use is not observed; population denominators are missing.
Child cause imports [7,8]	Match neonatal/post-neonatal scope in both arms; repair Pakistan diarrhoea and Angola pneumonia contrasts.	Source discrepancies are directly established. Epidemiological versions still differ.
Child age scope [1,8]	Confine source-export reductions to under-fives.	Paper specifies children 0–14, not adults. Ages 5–14 require additional evidence.
Reference matching [8,10]	Common 2015 ratios; bound six early malaria cells at no harm; test signed version.	Synthetic baseline equality does not recover original model trajectories.
Experience and references [1,7]	Concave experience profile; corrected broad-age and female-population references.	Sign already corrected in archived version history and printed manuscript; inheritance by R cannot be assumed.
Morbidity [1,7,9]	Apply disease mortality fractions to matched YLD; deduct survivor disability; test reduced responses.	Broad NCD response and YLD-to-work conversion are assumptions, not validated intervention effects.
Survivors and timing	Age retained cohorts; apply subsequent all-cause mortality; midpoint person-years; include 2018 effects.	No pre-2018 survivor stock, endogenous births, migration or employment response.
Financing	Subtract 10% of incremental costs from capital formation; test alternative shares and full-cost convention.	Not a country fiscal estimate; GDP gains are not budget receipts.

The supplement is organised to separate runnable analysis from archival extraction. In `model/`, `inputs.json` contains frozen inputs; `run.py` rebuilds all scenario tables; `verify.py` checks identities and boundary conditions; and `README.md` explains assumptions and operation. Run the following from that directory:

```
python run.py
python verify.py
```

`comparison.csv` gives country results, `annual.csv` annual levels, `health.csv` health mappings, `mortality_rebase.csv` reference-year diagnostics and `sensitivity.csv` alternative specifications. Monetary entries are dollars, not billions, unless labelled otherwise. Percent entries are percentages: 0.186 denotes 0.186%, not a fraction of 0.186.

The `valuation/` directory preserves the earlier controlled capital/sign diagnostic, its input snapshot, conversion checks and runnable routines. Its results isolate initialisation choices under its stated baseline assumptions; they are not additional UHC estimates to combine with the primary results. Extraction scripts in `audit/` document operations requiring the original workspace source files and are not needed to reproduce the frozen-input simulations. A file manifest records checksums. The manuscript tables are generated directly from the frozen CSV outputs; displayed rounding is intentional.

Separate audit companions retain `aggregate.csv` and `countries.csv` for the nine capital variants, and country-specific health-mapping diagnostic tables for the total-row reversals. These extend the original

computational supplement; they are not described here as a public repository. Table 3 records the source-cell reconciliation; Table 4 reports the one-mapping reversals; Table 5 is regenerated from `corrected_health_table.csv` and its annual numerators and denominators. Capital diagnostics verify identical health, labour and cost records across initializations. The health-mapping tests instead hold baseline economics and spending fixed while changing the named mapping. These experiments should not be pooled as independent additive corrections.

This companion makes the present analysis rerunnable. Full independent source reconstruction additionally requires the linked original files and their dependencies. Neither a checksum nor successful execution substitutes for that distinction, or for substantive external validation.