PROBABILITY FOR AI · FINAL PROJECT 2

Does National Aging & HDI Predict Years Lived with Disability?

João Carlos da Palma Soares · Probability for AI, Fall 2026 · App: pai.inclui.me/global-disability/

1. The Question

Global life expectancy has climbed steadily, yet recent research across 183 WHO member states shows healthy life expectancy has not kept pace, producing an expanding "healthspan–lifespan gap" (Garmany & Terzic, 2024). Public-health planners need to know whether a single global average accurately captures national disability burdens or whether development and aging drive predictable variation. Across 182 sovereign countries, we test whether conditioning expected years lived with disability (**LE – HALE**) on the Human Development Index (HDI) and population share aged 65+ beats a pooled global baseline out-of-sample.

2. Pre-Data Guess (Locked Before Inspecting Country Data)

> *"Before inspecting the dataset, I am **80% sure** that modeling country-level disability burden conditioned on development index (HDI) and older-adult population share (65+) will substantially beat a pooled global baseline model on out-of-sample $\log_{10}$ likelihood across the locked test countries."*

3. Data, Stratified Split & Probabilistic Models

Open Data & Split: We merged four open-licensed (CC BY) Our World in Data tables by ISO-3 code—WHO GHO Life Expectancy (LE) and Healthy Life Expectancy (HALE), UNDP HDI and world regions, and UN WPP age groups—yielding **182 sovereign countries** with zero missing values for 2021 (mean 9.42 yrs, SD 1.30 yrs; range 6.84 yrs in Lesotho to 12.50 yrs in Australia) plus 2019 as a pre-COVID check. Before fitting parameters, a deterministic region-stratified hash split countries into **127 training (70%)** and **55 locked test countries (30%)** across all six world regions.

Model 1 (Pooled Global Baseline): Models every nation as an independent draw from one global Normal distribution fitted on training data: **$Y \sim N(\mu_0 = 9.53 \text{ yrs}, \sigma_0 = 1.30 \text{ yrs})$**, ignoring development and demographics.

Model 2 (Conditioned Normal Linear Model): Models expected disability years given national HDI and 65+ share, fitted strictly on the 127 training countries: **$\hat{y} = 4.903 + 6.277 \cdot \text{HDI} + 0.0062 \cdot \text{Age65+}$, $\sigma_2 = 0.804$ yrs** (Training R^2 = 62.3%).

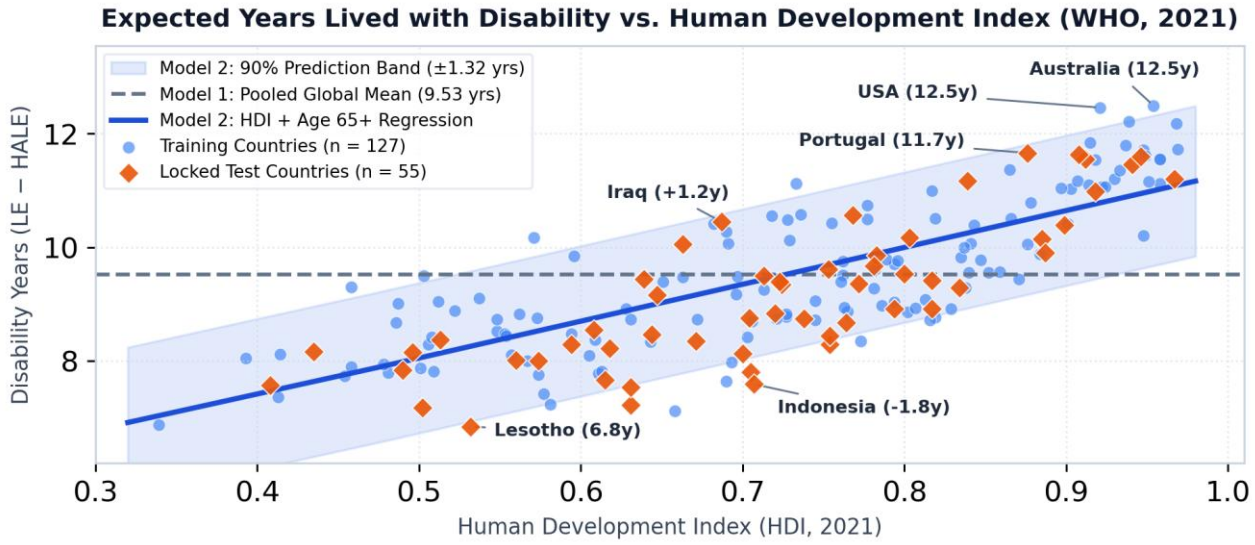

Figure 1. Expected years lived with disability (LE – HALE) vs. HDI across 182 countries (WHO/UNDP/UN, 2021). Blue circles: 127 training countries; orange diamonds: 55 locked test countries. Model 2 captures the upward gradient with a tighter ±1.32 yr 90% band versus Model 1's flat mean (9.53 yrs).

4. Calibration & Locked Test Set Results

Opening the 55 sealed test countries confirmed the locked 80% guess: Model 2 gained **$\Delta \log_{10} L = +11.49$** (≈ 3.1 × 10^{11}× higher likelihood), cut RMSE by 38.2%, and nearly halved calibration error across 10%–90% intervals. The pre-COVID 2019 check yielded nearly identical results ($\Delta \log_{10} L = +11.54$, test $R^2 = 0.599$).

Metric (Locked Test Set, n = 55)	Model 1: Pooled	Model 2: HDI + Age 65+	Improvement
Total $\log_{10}$ Likelihood (higher is better)	−40.57	**−29.07**	**+11.49 (3.1×10¹¹× more likely)**
Mean $\log_{10}$ Likelihood / Country	−0.738	**−0.529**	+0.209 per country
Out-of-Sample RMSE	1.32 yrs	**0.82 yrs**	−38.2% prediction error
Out-of-Sample R^2	−0.081	**0.587**	+58.7% variance explained
Calibration Avg Gap (10%–90% Intervals)	6.36%	**3.54%**	Nearly halved (−2.82 pp)
Observed 90% Interval Coverage	94.5%	**89.1%**	Within 0.9 pp of 90% target

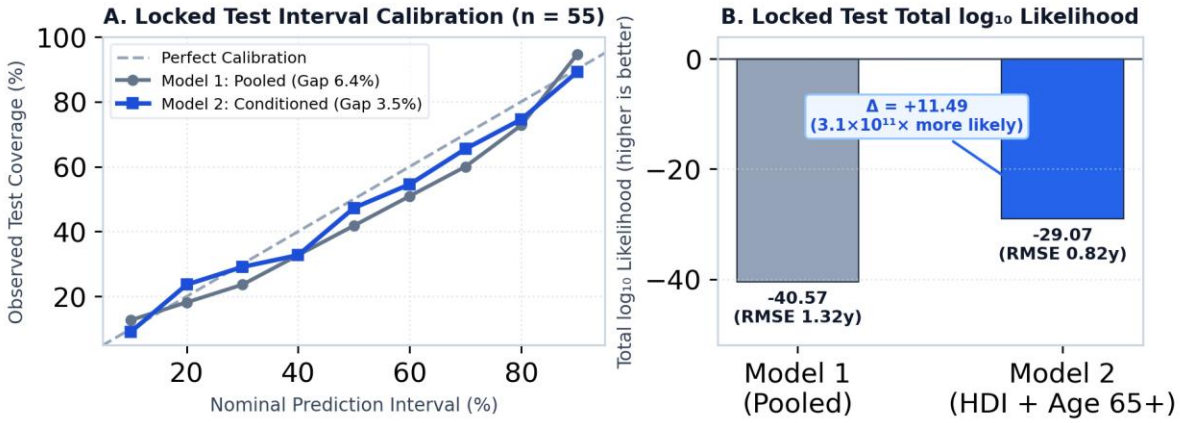

Figure 2. Locked test evaluation (n = 55, 2021). Panel A: Model 2 tracks the ideal calibration line closely (3.5% vs. 6.4% gap). Panel B: Model 2 gains +11.49 total $\log_{10}$ likelihood and cuts RMSE from 1.32 to 0.82 yrs.

5. Key Findings & What Could Still Be Wrong

- **Longevity Paradox & Outliers:** Higher-HDI nations experience more disability years (11–12.5 yrs vs. 6.8–8 yrs) because longer survival extends late-life exposure to chronic impairments. Top positive test residuals are Iraq (+1.21 y, conflict legacy), Portugal (11.66 y vs. 10.55 y predicted, +1.11 y, 23.3% aged 65+), and Italy (+0.88 y); largest negative residuals are Indonesia (−1.79 y), Kiribati (−1.67 y), and Lesotho (−1.42 y).
- **Collinearity & Mathematical Coupling:** HDI and Age 65+ correlate strongly (r = 0.796). Because HDI embeds life expectancy, HDI absorbs almost all linear signal ($t = 8.38$, $p < 10^{-13}$) while Age 65+ adds negligible marginal slope ($t = 0.35$, $p = 0.726$). Since LE appears in both HDI and LE − HALE, journal-grade follow-ups must replace HDI with strictly non-health covariates (GNI per capita and schooling).
- **Composite Metric & Ecological Fallacy:** LE − HALE pools all non-fatal health loss, depends on uneven diagnostic surveillance across health systems, and masks within-country inequality and sex gaps (~2.4 more years in poor health for women globally; Garmany & Terzic, 2024).

6. Sources, References & AI Usage Disclosure

Open Data (CC BY via Our World in Data): WHO GHO (Life Expectancy & HALE, 2000–2021); UNDP HDR (HDI & regions, 1990–2023); UN WPP (age groups, 1950–2023).

Reference: Garmany A, Terzic A. Global Healthspan-Lifespan Gaps Among 183 World Health Organization Member States. *JAMA Network Open*. 2024;7(12):e2450241. doi:10.1001/jamanetworkopen.2024.50241.

AI Disclosure: I selected the global disability question, open CC BY datasets, and locked pre-data guess (80% sure), and verified all splits and results. Gemini assisted with Python analysis, the offline React/Recharts app, charts, and document layout.