Institute for Health
Metrics and Evaluation

League tables based on cost-effectiveness estimates of 132 interventions in 204 countries: a meta-regression approach

Marcia Weaver, PhD July 21, 2025

UNIVERSITY of WASHINGTON

Introduction: League tables

Country-specific incremental cost-effectiveness ratio (ICERs) for multiple interventions are necessary to construct a league table, but rarely available.

$$\text{ICER} = \frac{\textit{Change in the cost of health care}}{\textit{Change in health outcome}}$$

Country-specific league tables rank health interventions from the lowest ICER to the highest, i.e. best to worst option.

Funding could be allocated to the intervention with the lowest ICER first, and interventions added according to their rank and available funds.

We used meta-regression analysis to synthesize cost-effectiveness results, predict country-specific ICERs, and create league tables.

Methods: Data and categories

Data are from the Tufts University's cost-effectiveness registries through 2020 with health improvements measured in disability-adjusted life-years (DALYs) or quality-adjusted life-years.

We included health care interventions with at least two articles and three ICERs.

We grouped interventions into 17 broad categories such as food and diet, and estimated 17 meta-regression models.

For each model, the dependent variable is ICER, and independent variables are country, methods, and intervention characteristics.

	Prevention	Diagnosis	Treatment
Antibiotics			X
Antivirals for Hepatitis B & C			X
Cancer		X	X
Cataract surgery			X
Cochlear implants			X
Food and diet	X		
Hypertension drugs including ACE inhibitor	X		
HIV	X		X
Malaria	X		
Psychological services			X
Syphilis	X		
Tuberculosis	X	X	X
Vaccines	X		

Estimating equation

Estimate $\log ICER_i = \sum_j \alpha_j I_{Gij} + \sum_k \beta_{Gk} X_{ik} + \sum_l \gamma_{Gl} Y_{il} + \sum_m \delta_{Gm} Z_{im} + \mu_s + \varepsilon_i$

i refers to a ratio, j refers to intervention, and s refers to article for random effect

I_{Gij} = Matrix of indicator variables for each intervention j in group G

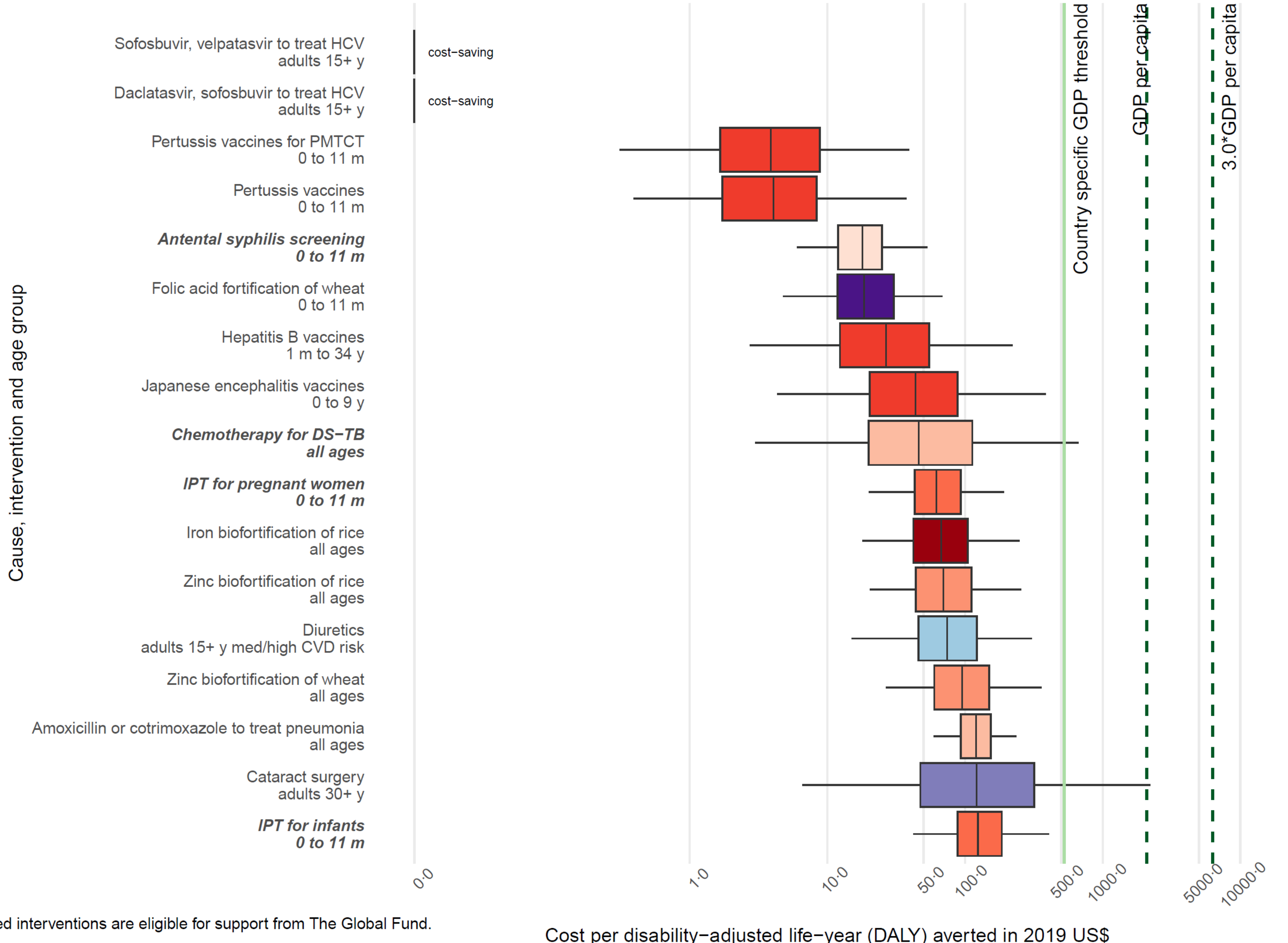
Subscript G denotes that coefficients are specific to a group, and coefficient for α_j are specific to an intervention

X_k= GBD variables explaining true variation, e.g. GDP and DALYs per capita

Y_{il}= bias variables, e.g. discount rate, perspective, outcome variable

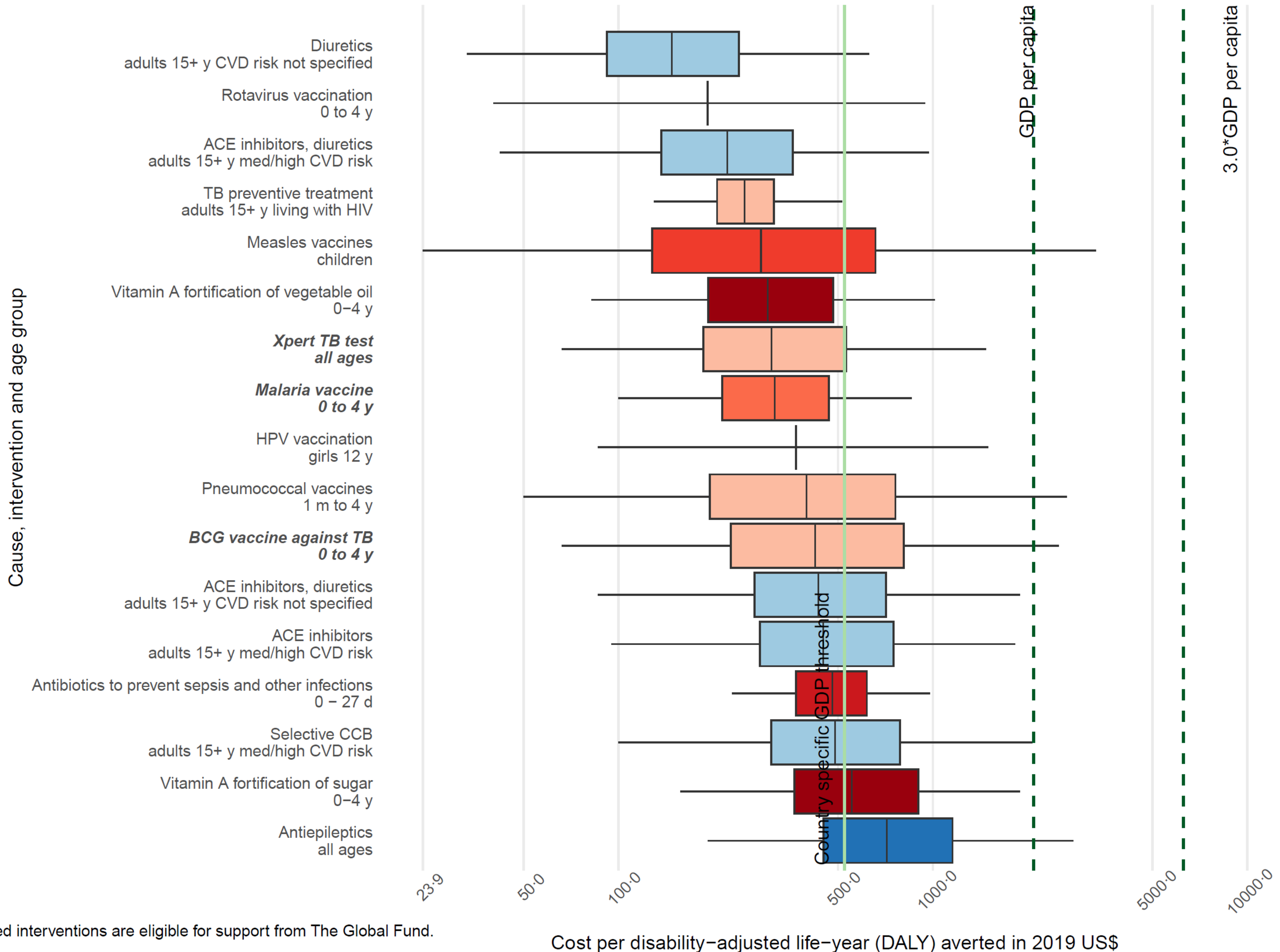
Z_{im} = intervention characteristics, e.g. cost, efficacy, risk group

Preliminary results of interventions ranked by incremental cost−effectiveness ratio (ICER) in India in 2019



Bolded interventions are eligible for support from The Global Fund.

Preliminary results of interventions ranked by incremental cost−effectiveness ratio (ICER) in India in 2019



Bolded interventions are eligible for support from The Global Fund.

Conclusion

League tables show that many interventions for non-communicable diseases are among the best options.

Meta-regression is a promising method to produce country-specific ICERs and league tables.

Countries with health technology assessment (HTA) agencies could use our estimates for interventions with many published ICERs, and focus their original research on those without published CEA.

Countries without HTA could use our estimates for ranking and selecting interventions.

Acknowledgments

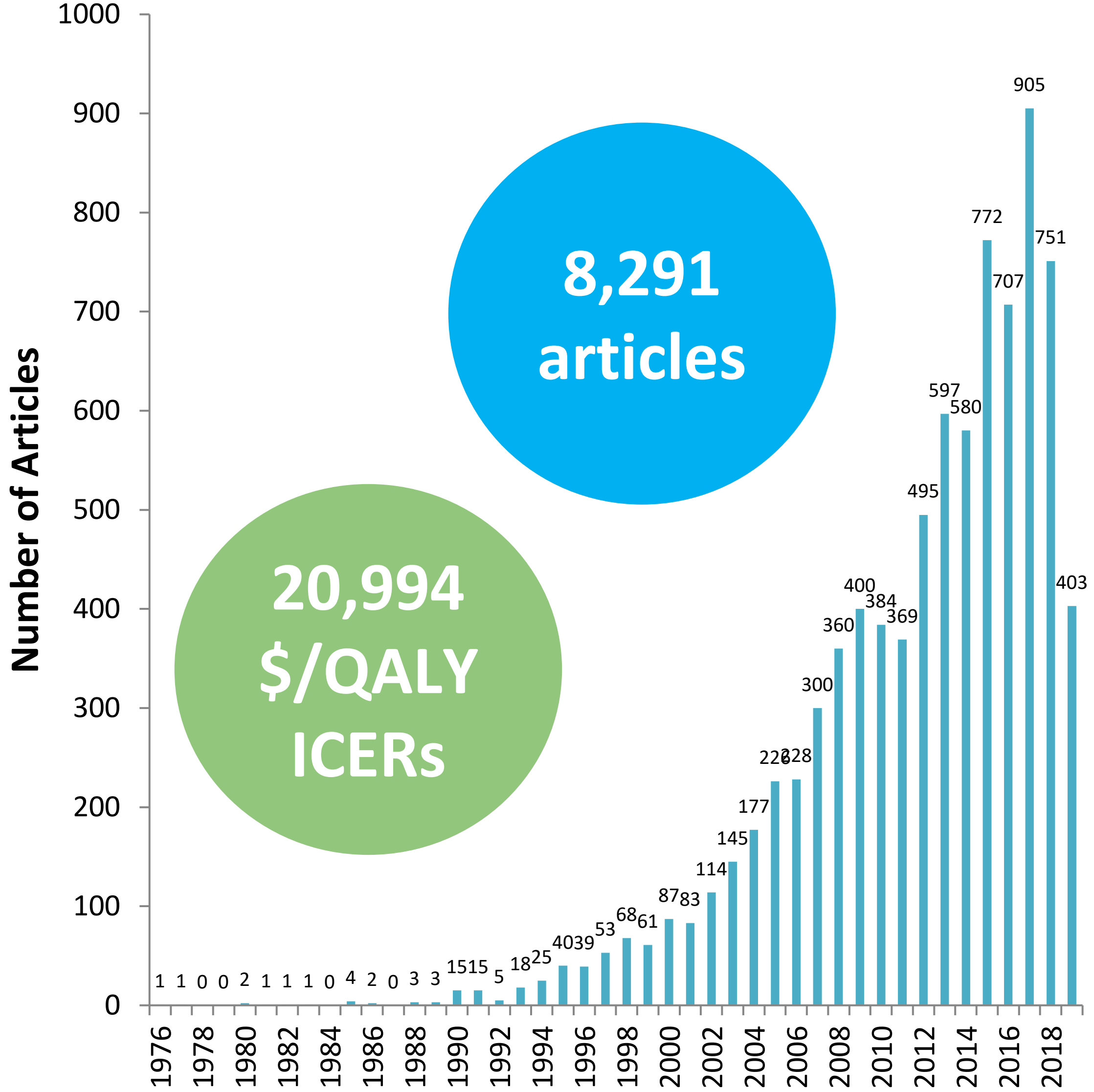
Co-authors: Fiona Silke, Esha Dwibedi, Lauren Earl, Johnathan Hsu, Jessica Klusty, Aishe Memetova, Peng Zheng, Aleksandr Aravkin, Christopher JL Murray

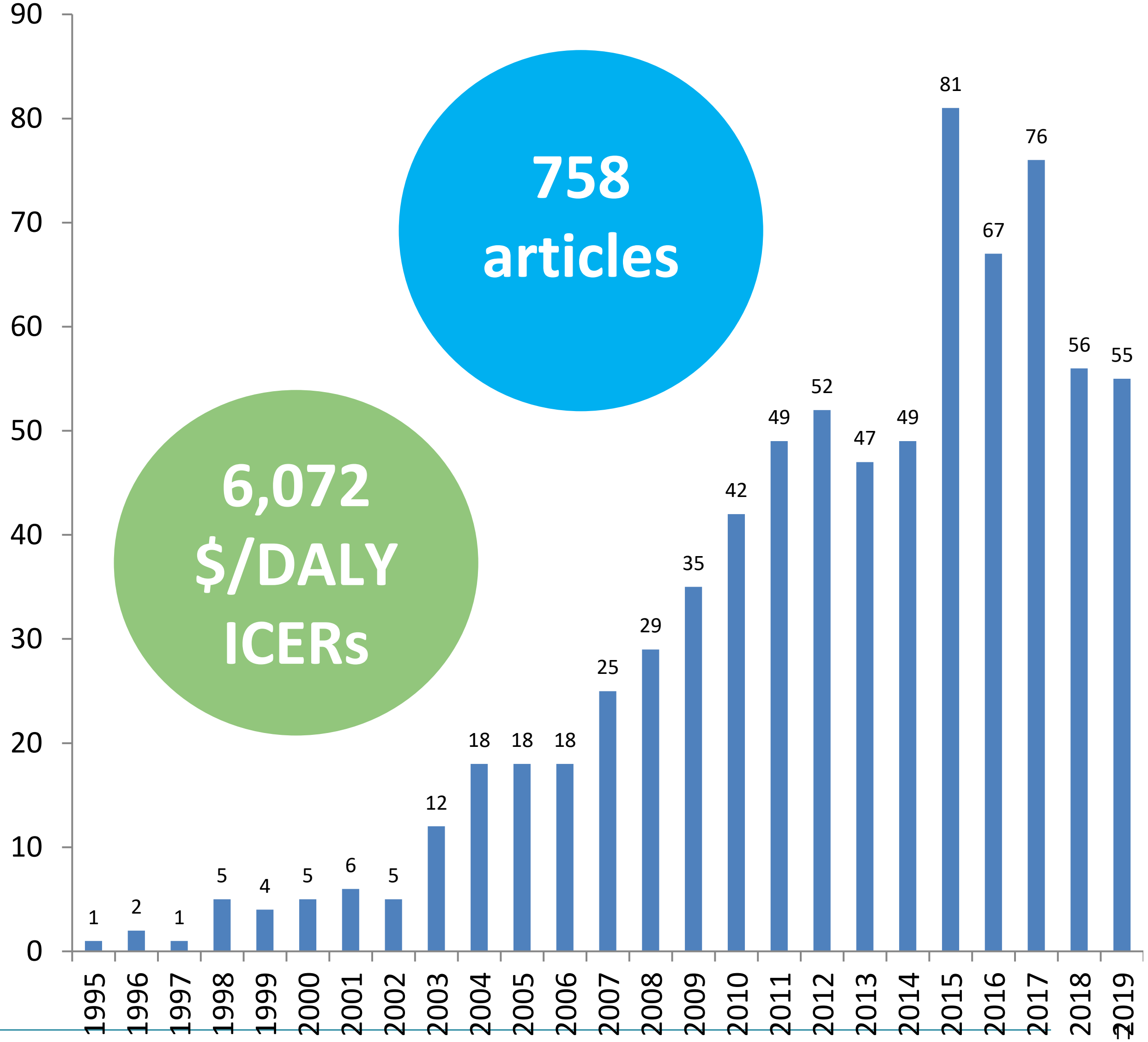
Funding: Bill and Melinda Gates Foundation

Extra slides

NUMBER OF $/QALY & $/DALY ARTICLES (ALL COUNTRIES)

GLOBAL HEALTH CEA REGISTRY





Source: David Wilson, BMGF

Methods: Meta-analysis

- In clinical research, statistical methods have been developed (fixed, random and mixed effect models) and improved over time to meta-analyze and meta-regress results from multiple studies.
- Cochrane Collaboration has made these methods widely used in evidence-based medicine.
- Campbell Collaboration has expanded this for policy interventions.
- In economics research, synthesis of evidence is rare, even though conflicting results are common.
- We are applying meta-regression methods to the published CE literature with the goal of producing intervention-location-time specific ICERs and their uncertainty intervals.

Methods: Advances in meta-regression tools

Sasha Aravkin, Peng Zheng and others have developed more flexible meta-regression tools including:

- crosswalk estimation methods
- non-linear splines on key covariates such as GDP per capita

The tool MR-BRT now used for relative risk function estimation in the GBD.

ARTICLES
https://doi.org/10.1038/s41591-022-01973-2
nature medicine

Check for updates

OPEN

The Burden of Proof studies: assessing the evidence of risk

Peng Zheng[1,2], Ashkan Afshin[1,2], Stan Biryukov[1], Catherine Bisignano[1], Michael Brauer[1,2,3], Dana Bryazka[1], Katrin Burkart[1,2], Kelly M. Cercy[1], Leslie Cornaby[1], Xiaochen Dai[1,2], M. Ashworth Dirac[1,2], Kara Estep[1], Kairsten A. Fay[1], Rachel Feldman[1], Alize J. Ferrari[1,2,4,5], Emmanuela Gakidou[1,2], Gabriela Fernanda Gil[1], Max Griswold[1], Simon I. Hay[1,2], Jiawei He[1], Caleb M. S. Irvine[1], Nicholas J. Kassebaum[1,2,6], Kate E. LeGrand[1], Haley Lescinsky[1], Stephen S. Lim[1,2], Justin Lo[1], Erin C. Mullany[1], Kanyin Liane Ong[1], Puja C. Rao[1], Christian Razo[1], Marissa B. Reitsma[1], Gregory A. Roth[1,2,7], Damian F. Santomauro[1,2,4,5], Reed J. D. Sorensen[1], Vinay Srinivasan[1], Jeffrey D. Stanaway[1,2], Stein Emil Vollset[1,2], Theo Vos[1,2], Nelson Wang[8], Catherine A. Welgan[1], Sarah S. Wozniak[1], Aleksandr Y. Aravkin[1,2,9,10] and Christopher J. L. Murray[1,2,10] ✉

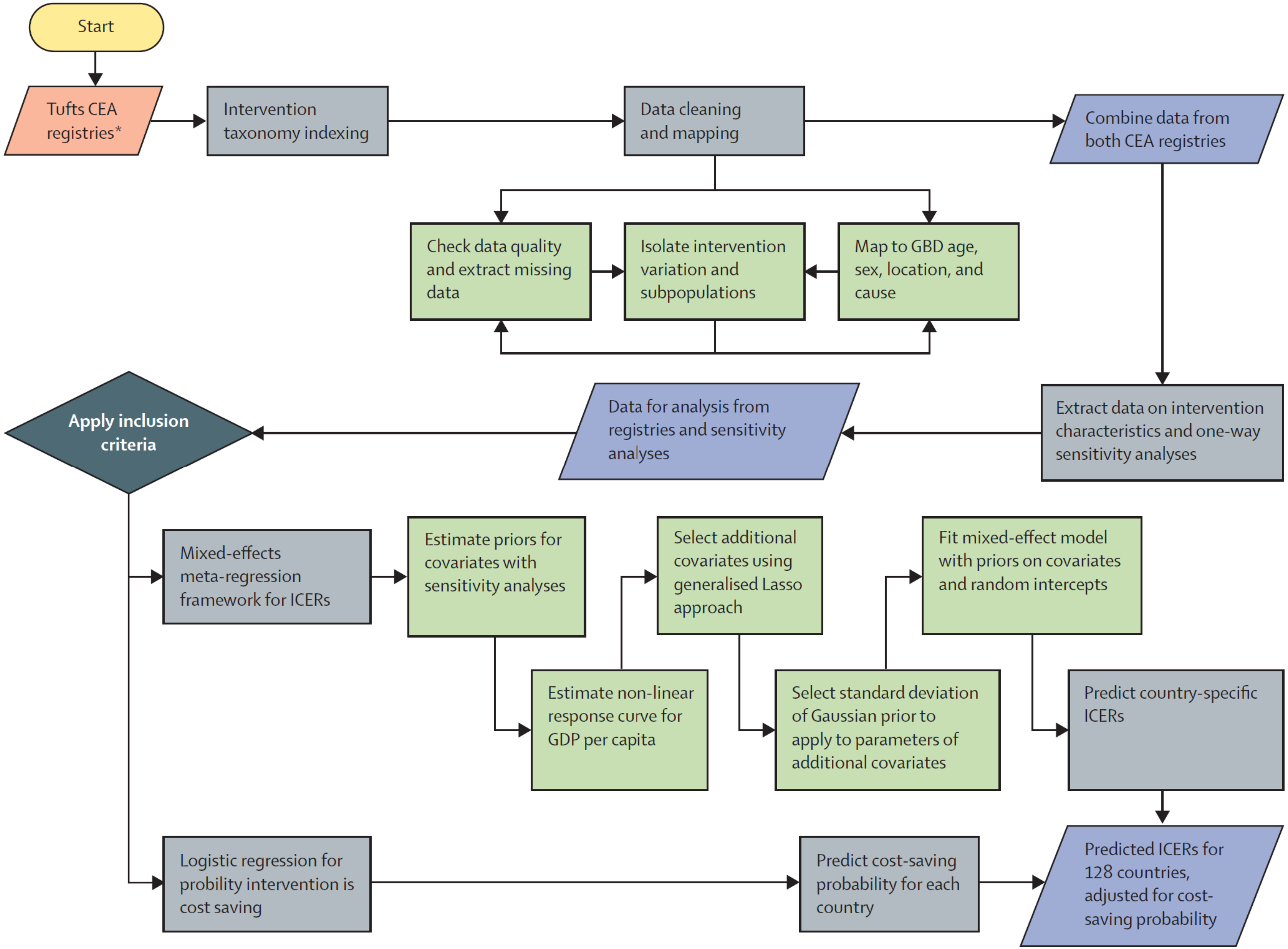
Exposure to risks throughout life results in a wide variety of outcomes. Objectively judging the relative impact of these risks on personal and population health is fundamental to individual survival and societal prosperity. Existing mechanisms to quantify and rank the magnitude of these myriad effects and the uncertainty in their estimation are largely subjective, leaving room for interpretation that can fuel academic controversy and add to confusion when communicating risk. We present a new suite of meta-analyses—termed the Burden of Proof studies—designed specifically to help evaluate these methodological issues objectively and quantitatively. Through this data-driven approach that complements existing systems, including GRADE and Cochrane Reviews, we aim to aggregate evidence across multiple studies and enable a quantitative comparison of risk-outcome pairs. We introduce the burden of proof risk function (BPRF), which estimates the level of risk closest to the null hypothesis that is consistent with available data. Here we illustrate the BPRF methodology for the evaluation of four exemplar risk-outcome pairs: smoking and lung cancer, systolic blood pressure and ischemic heart disease, vegetable consumption and ischemic heart disease, and unprocessed red meat consumption and ischemic heart disease. The strength of evidence for each relationship is assessed by computing and summarizing the BPRF, and then translating the summary to a simple star rating. The Burden of Proof methodology provides a consistent way to understand, evaluate and summarize evidence of risk across different risk-outcome pairs, and informs risk analysis conducted as part of the Global Burden of Diseases, Injuries, and Risk Factors Study.

Exposure to different risk factors plays an important role in the likelihood of an individual developing or experiencing more severe outcomes from certain diseases, such as high blood pressure increasing the risk of heart disease or not having access to a safe water source increasing the risk of diarrheal diseases[1]. Understanding and quantifying the relationship between risk factor exposure and the risk of a subsequent outcome is therefore essential to set priorities for public policy, to guide public health practices, to help clinicians advise their patients and to inform personal health choices. Consequently, information on risk–outcome relationships can be used in the formulation of many types of public policies, including national recommendations on diet, occupational health rules, regulations on behavior such as smoking in public places, and guidance on appropriate levels of taxes and subsidies. As new evidence is continuously being produced and published, the systematic and comparable assessment of risk functions is a dynamic challenge. Up-to-date assessments of risk–outcome relationships are essential to, and a core component of, the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) comparative risk assessment (CRA)[1–3], which aims to help decision-makers understand the magnitude of different health problems.

Evidence on risk–outcome relationships comes from many types of studies, including randomized controlled trials (RCTs), cohort studies, case-control studies, cross-sectional analyses, ecological studies and animal studies. Each study type has characteristic strengths and weaknesses. For example, RCTs are the most robust method for dealing with confounding but are often conducted with strict inclusion and exclusion criteria, meaning that trial participants

[1]Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA. [2]Department of Health Metrics Sciences, School of Medicine, University of Washington, Seattle, WA, USA. [3]School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada. [4]School of Public Health, The University of Queensland, Brisbane, Queensland, Australia. [5]Queensland Centre for Mental Health Research, Wacol, Queensland, Australia. [6]Department of Anesthesiology & Pain Medicine, University of Washington, Seattle, WA, USA. [7]Division of Cardiology, University of Washington, Seattle, WA, USA. [8]The George Institute for Global Health, The University of New South Wales, Sydney, New South Wales, Australia. [9]Department of Applied Mathematics, University of Washington, Seattle, WA, USA. [10]These authors contributed equally: Aleksandr Y. Aravkin, Christopher J. L. Murray. ✉e-mail: cjlm@uw.edu

2038 NATURE MEDICINE | VOL 28 | OCTOBER 2022 | 2038–2044 | www.nature.com/naturemedicine


Start
Tufts CEA registries*
Intervention taxonomy indexing
Data cleaning and mapping
Combine data from both CEA registries
Check data quality and extract missing data
Isolate intervention variation and subpopulations
Map to GBD age, sex, location, and cause
Apply inclusion criteria
Data for analysis from registries and sensitivity analyses
Extract data on intervention characteristics and one-way sensitivity analyses
Mixed-effects meta-regression framework for ICERs
Estimate priors for covariates with sensitivity analyses
Select additional covariates using generalised Lasso approach
Fit mixed-effect model with priors on covariates and random intercepts
Estimate non-linear response curve for GDP per capita
Select standard deviation of Gaussian prior to apply to parameters of additional covariates
Predict country-specific ICERs
Logistic regression for probility intervention is cost saving
Predict cost-saving probability for each country
Predicted ICERs for 128 countries, adjusted for cost-saving probability

IHME

Improving CEA data for meta-regression estimates

- Update the data from Tufts CEA registries from 2020 to 2023
- Add published WHO-CHOICE results
- Add published estimates for surgical interventions by mapping them to GBD results injuries by body part (N-codes) as opposed to cause (E-codes)
- Extract data from published CEA with disease-specific outcomes such as HIV infections averted and convert those outcomes to DALYs
- Collaborate with international agencies and other research groups to obtain unpublished CEA results.