Replication Project

Ruian Yang, Sihao Miao, Jiyu Wang, Xiaolu Qian, David Wei
https://github.com/mshhh/miao-qian-wang-wei-yang-replication-project

Quattrochi John, Joshua A Salomon, Kenneth Hill, and Marcia C Castro. 2019. “Measuring and Correcting Bias in Indirect Estimates of Under-5 Mortality in Populations Affected by Hiv/Aids: A Simulation Study.” *BMC Public Health*19 (1). Springer: 1516.

Introduction

- Goal of the project

  Complete a professional, publication-worthy replication of a public health paper: Measuring and correcting bias in indirect estimates of under-5 mortality in populations affected by HIV/AIDS: a simulation study.

- Reasons why we chose this paper
  - Recent published
  - Relevant subject
  - Well written and documented
  - Presented a model that's replicable

Quattrochi John, Joshua A Salomon, Kenneth Hill, and Marcia C Castro. 2019. "Measuring and Correcting Bias in Indirect Estimates of Under-5 Mortality in Populations Affected by Hiv/Aids: A Simulation Study." *BMC Public Health*19 (1). Springer: 1516..

Original Paper

- Research question

  How does indirect methods affect under-5-mortality (U5M) estimation in populations affected by HIV/AIDS?

- Methods

  Individual-level, discrete time-step simulation model

  Generalized linear regression model

- Result

  Indirect methods can underestimate U5M by 0–41% in populations with HIV prevalence of 0–40%.

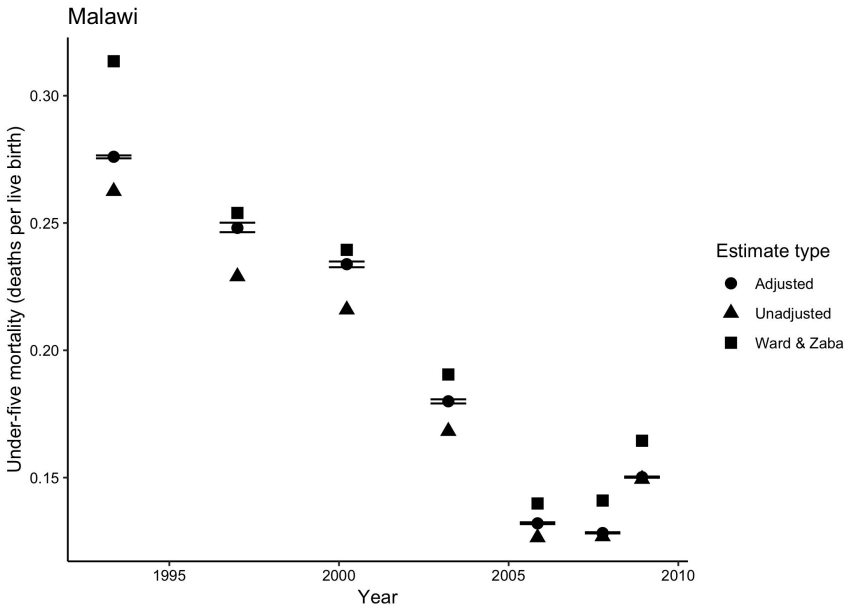
Quattrochi John, Joshua A Salomon, Kenneth Hill, and Marcia C Castro. 2019. “Measuring and Correcting Bias in Indirect Estimates of Under-5 Mortality in Populations Affected by Hiv/Aids: A Simulation Study.” *BMC Public Health*19 (1). Springer: 1516..

Replication

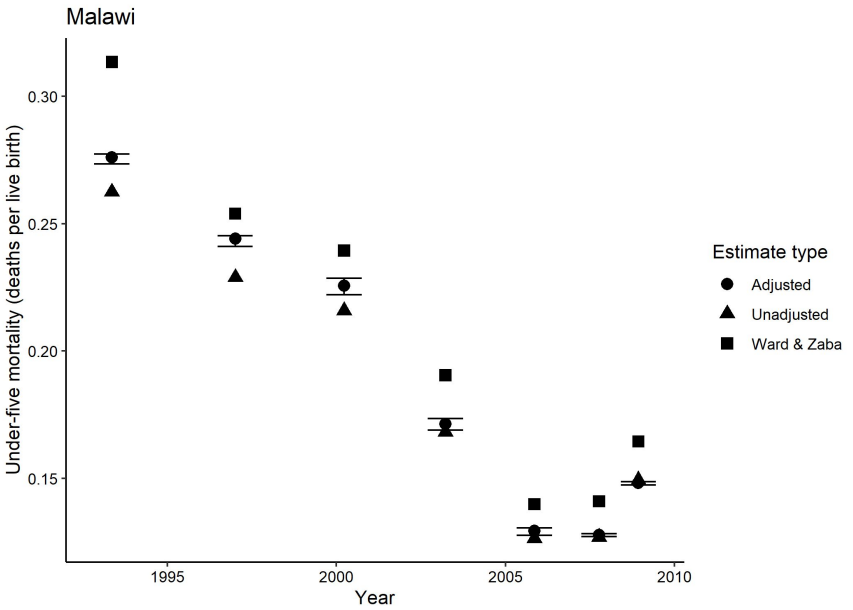
- Adopted the original individual-level, discrete time-step simulation model to generate simulation data for model building.
- Reduced the number of simulated population from 4480 to 1000.
- Selected the best performing model using error metrics
- Focused on replicating the figures showing the difference in estimating U5M using the best performing model for real data from two countries (Malawi and Tanzania).

Replication

Target figure 1 (Figure 4 in original paper)



Original

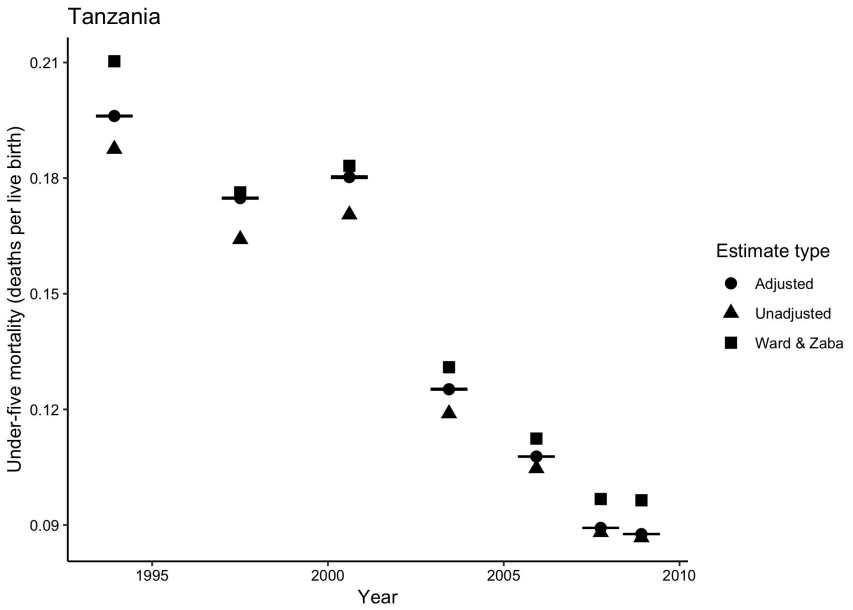


Replicated

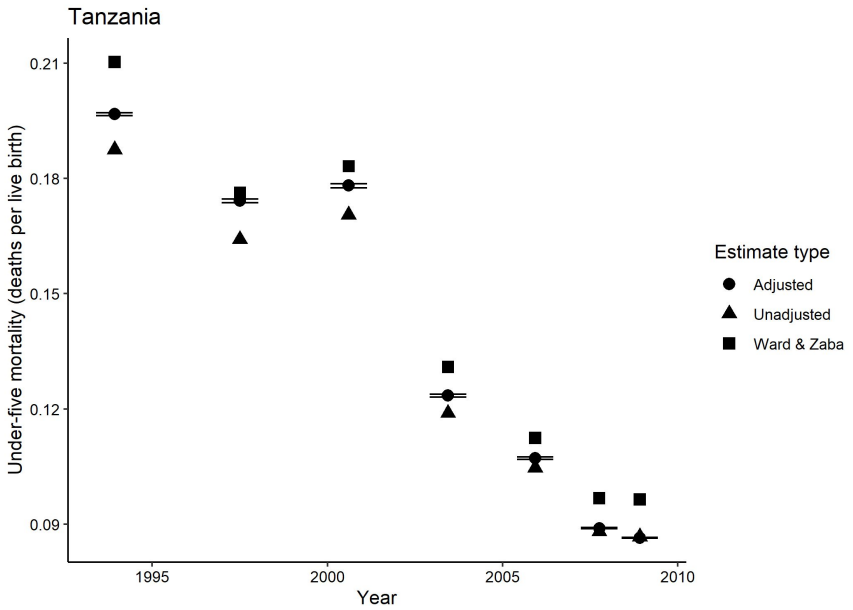
Quattrochi John, Joshua A Salomon, Kenneth Hill, and Marcia C Castro. 2019. "Measuring and Correcting Bias in Indirect Estimates of Under-5 Mortality in Populations Affected by Hiv/Aids: A Simulation Study." *BMC Public Health*19 (1). Springer: 1516..

Replication

Target figure 2 (Figure 5 in original paper)



Original



Replicated

Quattrochi John, Joshua A Salomon, Kenneth Hill, and Marcia C Castro. 2019. "Measuring and Correcting Bias in Indirect Estimates of Under-5 Mortality in Populations Affected by Hiv/Aids: A Simulation Study." *BMC Public Health*19 (1). Springer: 1516..

Replication

Table 2 Outcomes for simulated populations, summary statistics

Variable	Mean	Std dev	Median	Min	Max
HIV prevalence, 1990	0.07	0.07	0.08	0	0.22
HIV prevalence, 2000	0.15	0.14	0.09	0	0.40
HIV prevalence, 2010	0.11	0.09	0.07	0	0.24
ART coverage, 2004	0	0	0	0	0
ART coverage, 2008	0.03	0.07	0	0	0.21
ART coverage, 2010	0.08	0.20	0	0	0.58
ART prevalence, 2005	0.0002	0.0007	0	0	0.0030
ART prevalence, 2007	0.0010	0.0031	0	0	0.0126
ART prevalence, 2009	0.0019	0.0058	0	0	0.0239
Total fertility rate, 2000	4.88	1.67	4.64	2.87	6.93
Total fertility rate, 2010	4.28	1.64	4.16	2.45	6.19

Notes: Based on 1,000 simulated populations. ART coverage is defined as the percent of women with a CD4 count under 200 who are on ART. ART prevalence is defined as the percent of all women who are on ART.

- The highest HIV prevalence in any simulation was 40% in 2000.
- The same trend of ART coverage and prevalence as in the original paper that they both increase with time.
- The highest ART prevalence in any simulation was 2.4% in 2009 and this was 1.7% lower compared to the original model.

Replication

Table 3 Bias in indeitect estimates applied to 1,000 simulated populations

Outcome variable	statistic	15-19	20-24	25-29	30-34	35-39	40-44
Absolute bias	*mean*	0.000	-0.002	-0.007	-0.014	-0.021	-0.021
	std dev	0.001	0.002	0.007	0.014	0.019	0.017
	median	0	-0.001	-0.003	-0.008	-0.018	-0.018
	min	-0.003	-0.009	-0.027	-0.050	-0.071	-0.061
	max	0.005	0.001	0	0	0	0
Relative bias	*mean*	0.1%	-2.3%	-6.1%	-9.9%	-11.9%	-10.8%
	std dev	0.7%	2.6%	6.7%	10.4%	11.2%	9.4%
	median	0.0%	-1.4%	-3.7%	-6.7%	-8.8%	-10.0%
	min	-1.9%	-11.0%	-25.5%	-38.6%	-42.4%	-35.0%
	max	2.2%	0.3%	0.0%	0.0%	0.0%	0.0%
Yrs before survey that estimates pertain to	*mean*	1.2	2.7	4.6	6.9	9.3	12.0
Surviving women	*mean*	1	1	1	1	1	1
Women who died from HIV/AIDS	*mean*	1	1	1	1	1	1
Children ever born, surv women	*mean*	23587.24	81585.48	127981.69	142513.09	132254.01	115805.14
Children ever born, surv women + HIV deaths	*mean*	23626.27	82823.29	136070.38	164047.11	164424.69	151326.32
Dead children, surviving women	*mean*	0.09	0.10	0.12	0.14	0.15	0.18
Dead children, surv women + HIV deaths	*mean*	0.09	0.11	0.13	0.15	0.17	0.19
Ratio of HIV deaths to surviving women	*mean*	0	0	0	0	0	0

- The mean absolute bias was largest for estimates from women aged 35–39 and 40–44 (−0.021) and smallest for estimates from women aged 15–19 (0.000).
- Similar replication result on the largest absolute bias recorded, -0.071 for women 35-39 compared to the original paper (-0.069).
- The differences could be due to the smaller population size we used in our simulation.

Highlights

- Kept the population age (0, 120) and year (1906, 2010) parameters for simulation same as in the original paper.
- Reduced the simulated population size from 4480 to 1000.
- Revised some parts of the regression model and corrected inconsistent variable names in the original code during the debugging process.

Reflections

- R Markdown is handy
- Github collaboration worked pretty well for source code
- Google doc was favored in terms of parallel editing
- Need a better estimation on source data preparation
- A shared and reproducible computing environment (Docker) can be helpful to troubleshoot specific errors across team

FAQ

Questions?

Thank you for your attention!