

Small-area Estimates of U5MR Disaggregation by Space and Time

Samuel Clark

ALAP, Puebla, Mexico
October, 2018

Outline

- 1 Disaggregation & the SDGs
- 2 Small-area Estimates of the Under-five Mortality Rate – U5MR

Disaggregation and the SDGs

- 'Disaggregation' is a key principle of the SDGs
 - Social dimensions → *inclusion*
 - Time
 - Space
 - Sex
 - Age
 - Others ...

Number of cells

Potential dimensions and categories for a typical demographic indicator for a single country

- Admin-2 areas: ~ 200
- Annual: 15 years
- Sex: 2+
- Age: standard 5-year plus a few more, say ~ 30
- SES: ~ 5
- A few other social dimensions: ~ 5

Total cells for which a measurement is required for the duration of SDGs is \sim **6,750,000**

This is daunting

What to do?

- **Better and more data collection**
- **Data amalgamation**
 - Traditional sources of data
 - 'Big' data
 - Routine and administrative data
 - Requires appropriate models
- **New methods, models**
- **Training**

Models I

At present

- Models used to substitute for data
- Many models are complex 'black boxes' constructed by well-resourced groups in developed countries – **not transparent or transferrable**, hence not very useful
- Many of those same models focus on **national level** estimates, not very useful for local policy making and monitoring
- Many models provide unrealistically detailed estimates and consequently over-interpreted, e.g. 5km grid of U5MR

Models II

Need new methods and models for

- Data amalgamation
- Disaggregation

Must be transparent, open-source, and as simple as possible

1 Disaggregation & the SDGs

2 Small-area Estimates of the Under-five Mortality Rate –
U5MR

Small-area estimates of the U5MR

Small-area Analysis of 122 DHS Surveys in 262 Subregions of 35 Countries in Africa

Collaborators

- Jon Wakefield, UW
- Zehang (Richard) Li, Yale
- Yuan Hsiao, UW
- Jessica Godwin, UW
- Bryan Martin, UW

U5MR

- U5MR is the probability that an infant dies before its 5th birthday
- U5MR is an indicator for SDG 3: target $< 25/1,000$ live births
- Potential data sources
 - CRVS: birth and death records
 - **Census**: summary birth history or Brass questions
 - **Household surveys**: full birth histories
 - Health facilities: birth and death data
 - **Population surveillance systems**: health and demographic surveillance (HDSS)
 - Variety of predictive covariates available **remote sensing**: moisture, elevation, vegetation – e.g. malaria
 - Other predictive covariates from **digital exhaust**: SES from call record data

Small-area estimates of U5MR in Africa I

Aim: annual estimates of U5MR at admin-1 or admin-2 level

- Data
 - Full birth history from all DHS surveys in a country
 - HDSS surveillance data
 - Summary birth history from censuses – *working on new methods for this now*
- Approach
 - For each data source, estimate U5MR with variance (uncertainty) for all possible place and times – account for design and coverage of each data source
 - Combine estimates of U5MR from all data sources into one space-time smoothing model across all places and time
 - Accounts for auto correlation in space and times

Small-area estimates of U5MR in Africa II

- Result
 - **Disaggregated by space and time** using
 - Data Amalgamation
 - Minimal modeling
 - Developed new modeling approaches as necessary
 - Stack of estimated U5MR surfaces, one for each year, with holes filled in
 - Precision in both space and time – areas and times with certain estimates of U5MR have narrow credible intervals
 - Allows identification of spatial heterogeneity in both levels and changes in U5MR
 - Can **utilize all available data**

Small-area estimates of U5MR in Africa III

- **Comparatively simple and transparent**
open-source R package for the whole method: SUMMER
<https://cran.r-project.org/package=SUMMER>
- Manuscript in review

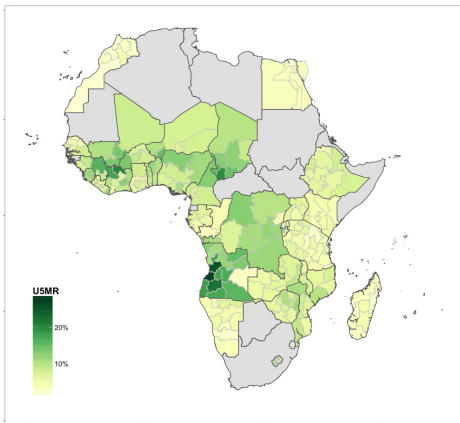
and

Laina D. Mercer, Jon Wakefield, Athena Pantazis, Angelina M Lutambi, Honorati Masanja, and Samuel J. Clark. “Small Area Estimation of Child Mortality in the Absence of Vital Registration”. *The Annals of Applied Statistics*, 9(4):1889-1905, 2015.

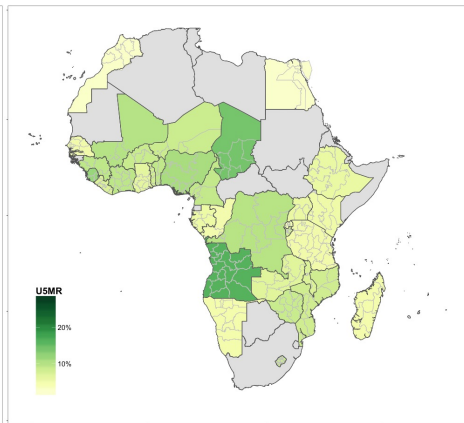
Challenges

- All first-stage estimates of U5MR must have estimate with variance – many traditional demographic techniques do not include variance estimate, e.g. Brass methods → new methods
- Most data sources do not include data for all places and times – model is estimated on a data set that in space and time resembles a block of Swiss cheese
- Data must be acquired and prepared from many sources – logistically complex and tedious
- Second-stage smoothing model is not simple to create, estimate, or validate
- ‘Fit for purpose’ scaling – what level of space and time resolution is useful and doable
- Communicating results is complex

U5MR in 2015

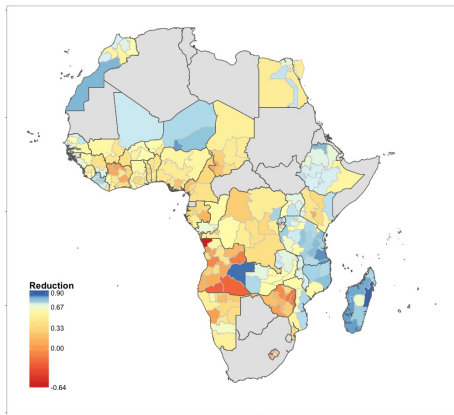


(1) Subnational

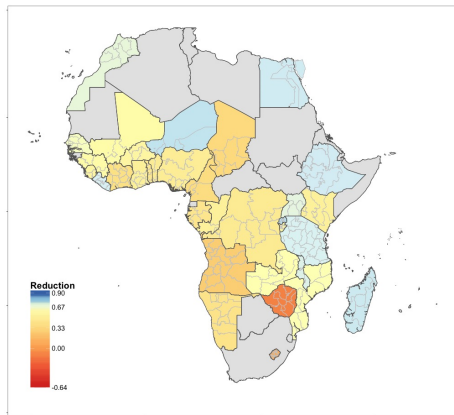


(2) National

Reduction in U5MR 1990–2015

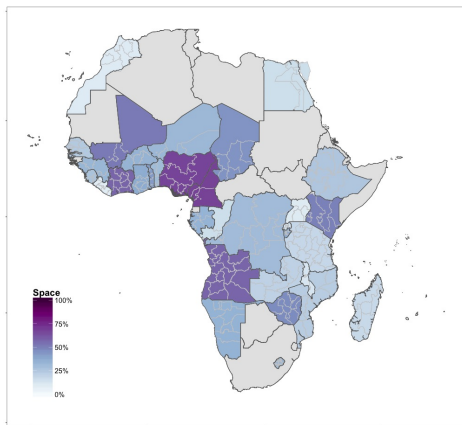


(3) Subnational

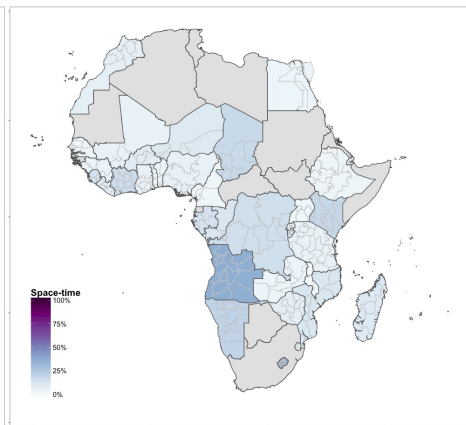


(4) National

Variance explained by space and time



(5) Space: subregions different



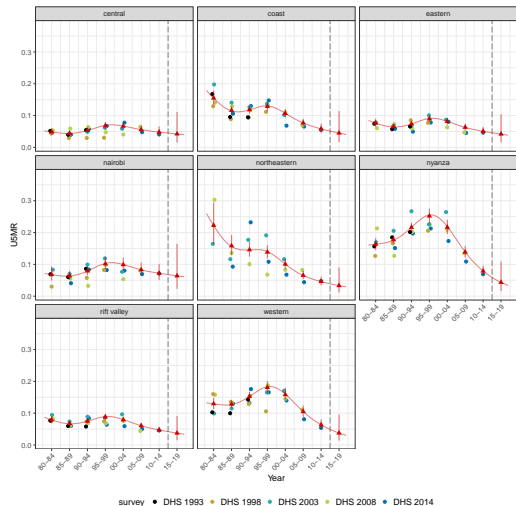
(6) Space-time: subregions changing differently

Africa: Kenya



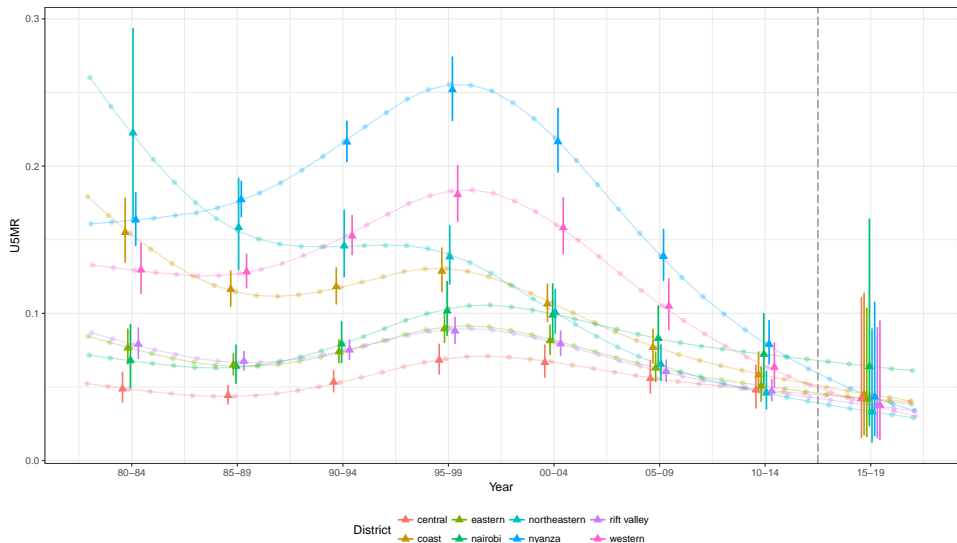
(7) Africa with countries

U5MR Kenya by admin-1 area and year with data source



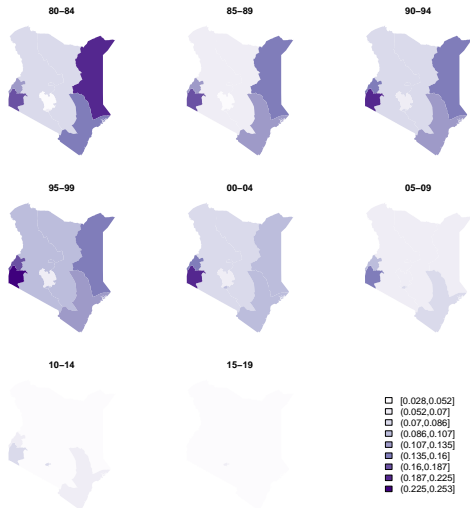
(8) U5MR by admin-1 and time with DHS survey estimates

U5MR Kenya by admin-1 area and year



(9) U5MR by admin-1 and time

Median U5MR Kenya by admin-1 area and time period



(10) U5MR by admin-1 and time period

We need to grow

- All this requires new methods
- New methods require **new training**
- Need other fields
 - Statistics
 - Computer Science
 - GIS/Remote Sensing
 - ...
- New approaches to data management
- New thinking about sampling and representativity
- Data Amalgamation
- A lot more attention to precision/uncertainty in everything

