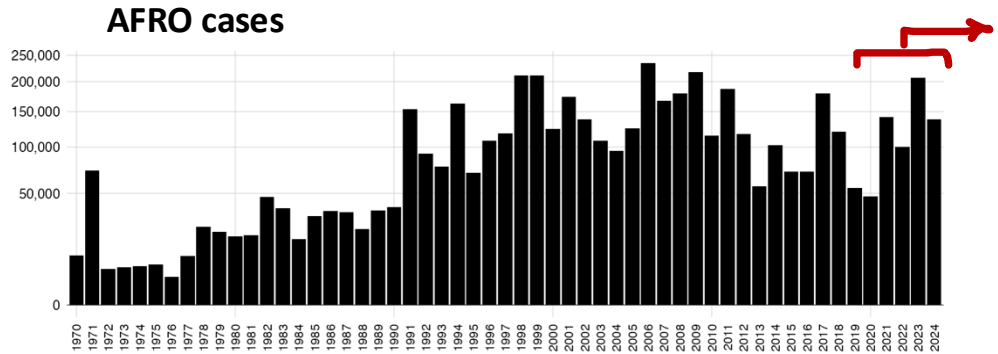
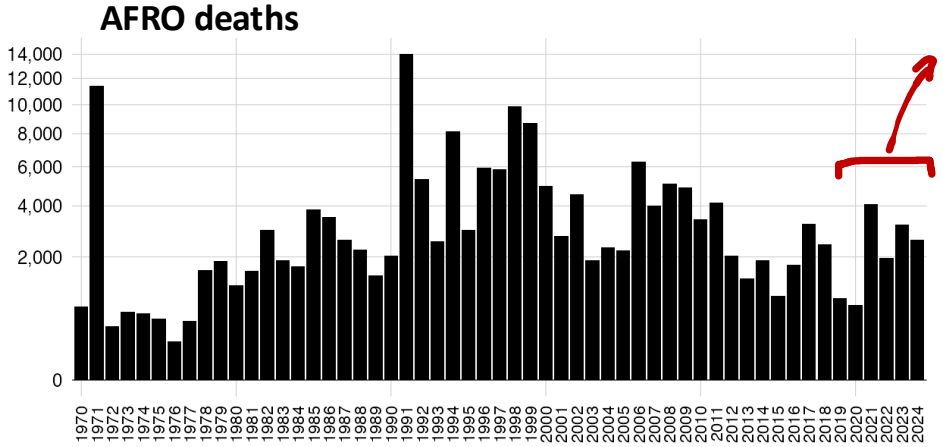
Cholera modeling capacity at IDM:
leveraging diverse data streams for scenarios and forecasting

John R Giles

ASTMH 2024

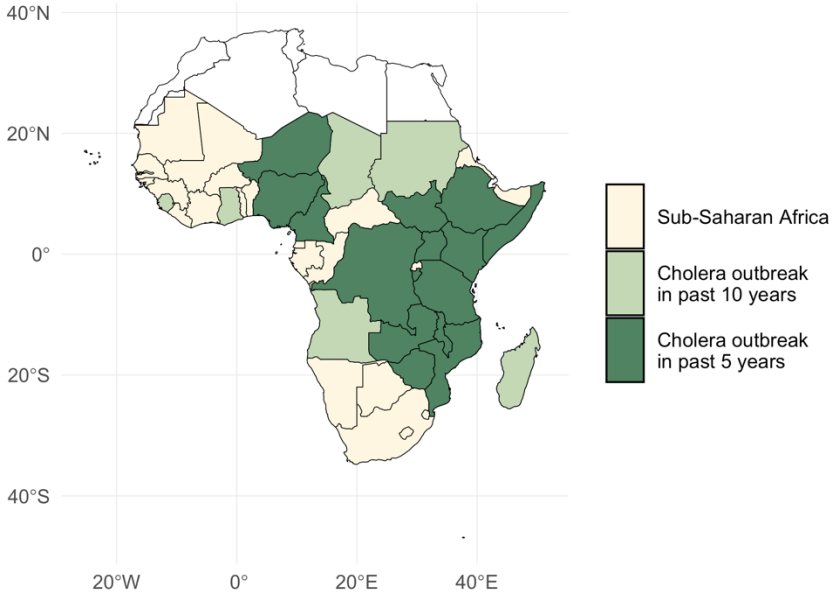
Recent resurgence of cholera in Sub-Saharan Africa



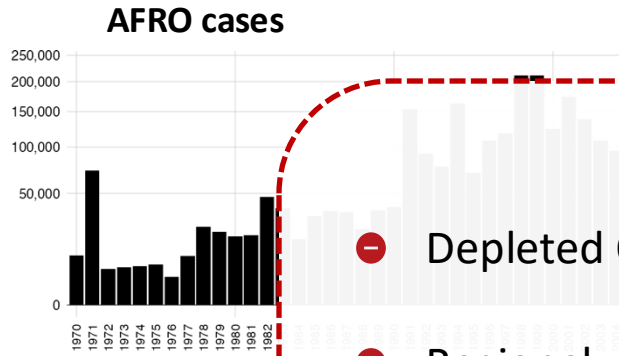


Increase in cases and deaths since 2021

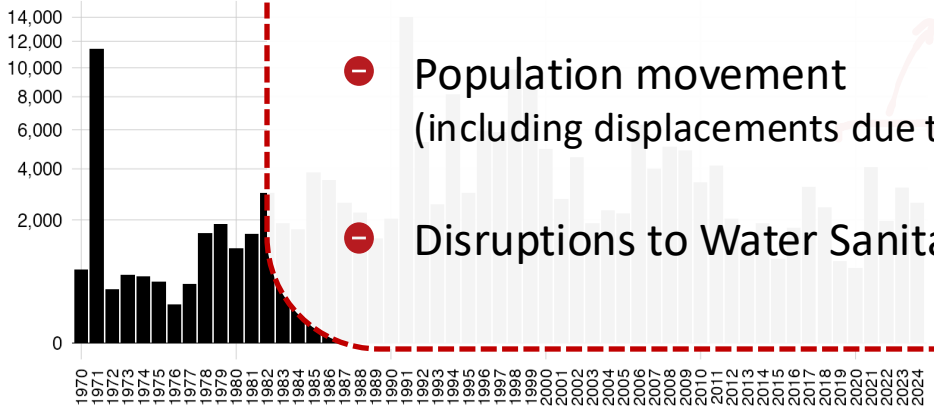
Eastern and Central Africa most affected



Recent resurgence of cholera in Sub-Saharan Africa



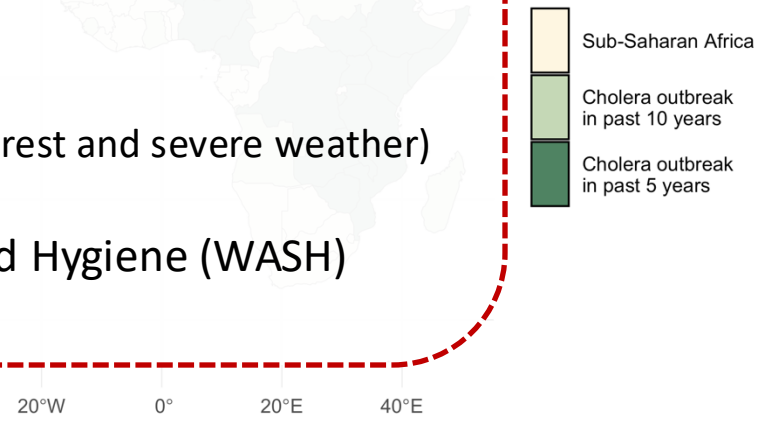
Increase in cases and deaths since 2021



Complex set of drivers:

- Depleted Oral Cholera Vaccine (OCV) stockpile
- Regional weather patterns + climate change
- Population movement
  (including displacements due to civil unrest and severe weather)
- Disruptions to Water Sanitation and Hygiene (WASH)

Eastern and Central Africa most affected



Cholera response set to improve

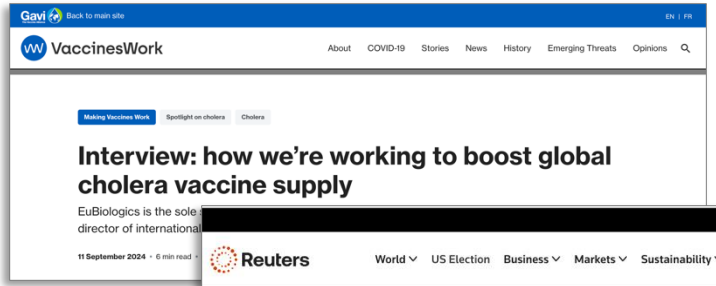
GLOBAL TASK FORCE ON CHOLERA CONTROL

ROADMAP 2030

A global strategy to reduce cholera deaths and stop cholera transmission

- Eliminate from 20 countries
- Reduce deaths by 90%

- OCV stockpile expected to improve in 2025 and 2026



EuBiologics

2025: 70 million

2026: 90 million

Bharat Biotech

2025: 45 million

2026: up to 200 million

IDM INSTITUTE FOR DISEASE MODELING

IDM cholera modeling capacity: key questions for surveillance and modeling

1. Scenarios of OCV administration at regional scale
   - OCV supply required for GTFCC goal?
   - Impacts of reactive and preventative OCV campaigns
   - One- and two-dose strategies

2. What is the relative contribution of the various drivers of transmission?

   OCV + climate change + population movement + WASH

Metapopulation Outbreak Simulation with Agent-based Implementation for Cholera (MOSAIC)

Spatial model of endemic cholera in Sub-Saharan Africa

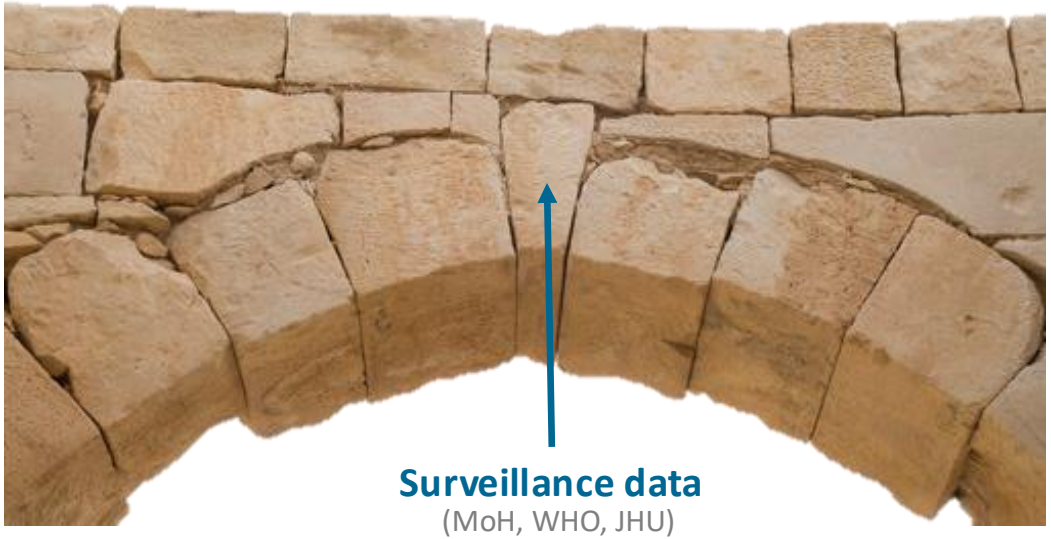
Scenarios and forecasting Q2 2025

MOSAIC

Metapopulation Outbreak Simulation with Agent-based Implementation for Cholera (MOSAIC)

Diverse data streams required

→ **Spatial model of endemic cholera**



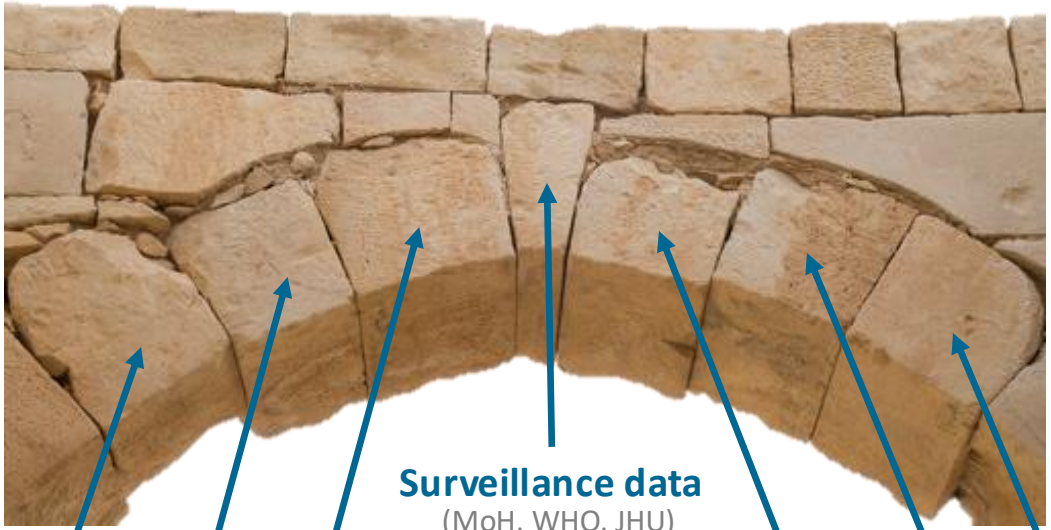
Surveillance data
(MoH, WHO, JHU)

IDM INSTITUTE FOR DISEASE MODELING

→ Spatial model of endemic cholera

Metapopulation Outbreak Simulation with Agent-based Implementation for Cholera (MOSAIC)

Diverse data streams required



Surveillance data
(MoH, WHO, JHU)

OCV
(WHO ICG, GTFCC)

Pop movement
(OAG, CDR, surveys)

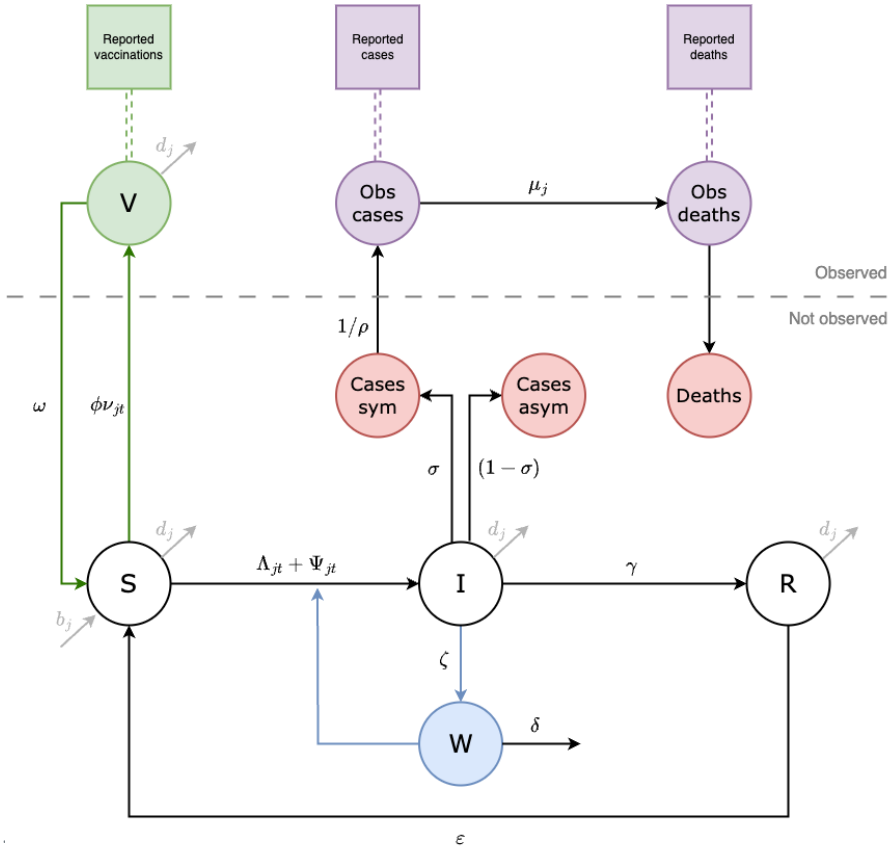
WASH
(WHO UNICEF JMP)

Climate
(Open-Meteo, BOM)

Immune dynamics
(Published studies)

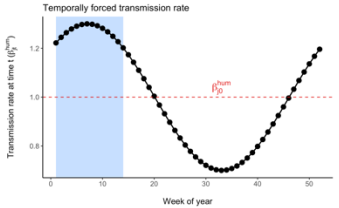
Observational bias
(Published studies)

MOSAIC model structure: complex drivers require diverse data streams and model mechanisms

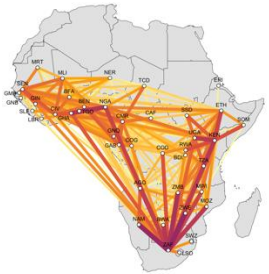


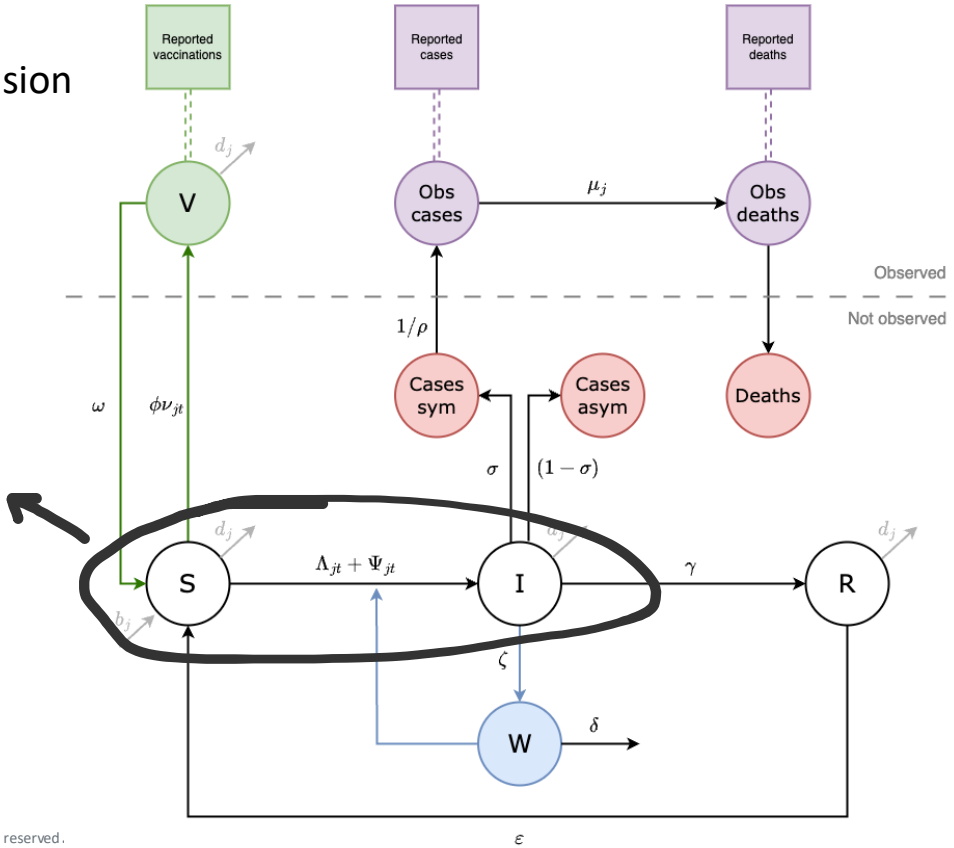
MOSAIC model structure: complex drivers require diverse data streams and model mechanisms

Human-to-human transmission with seasonal forcing



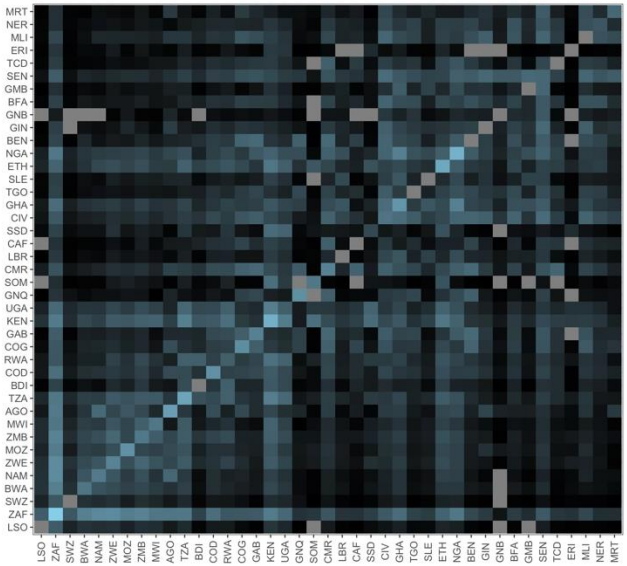
Spatial transmission among countries



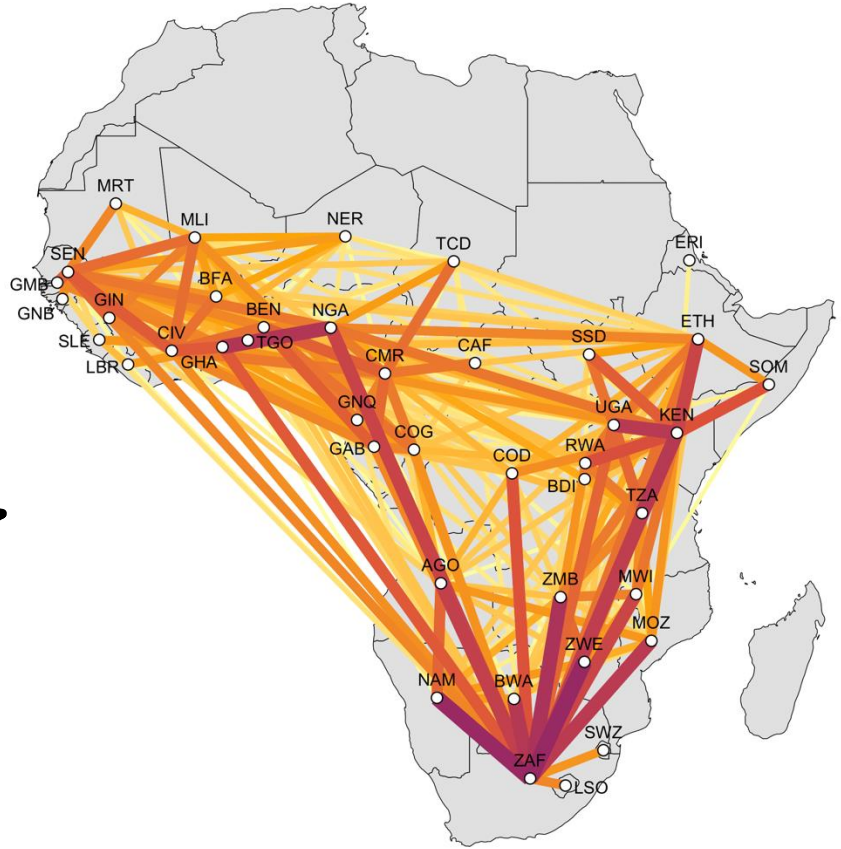


Population movement

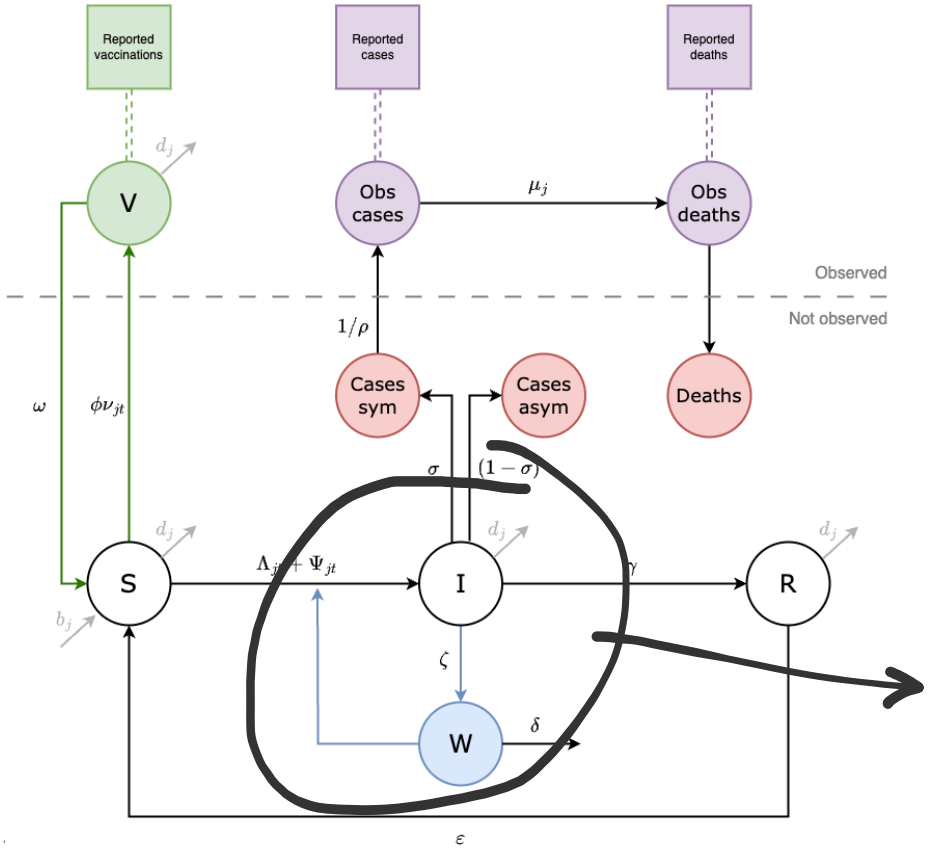
- OAG air passenger data crucial for model spatial structure



Network model of average weekly connectivity within SSA



MOSAIC model structure: complex drivers require diverse data streams and model mechanisms

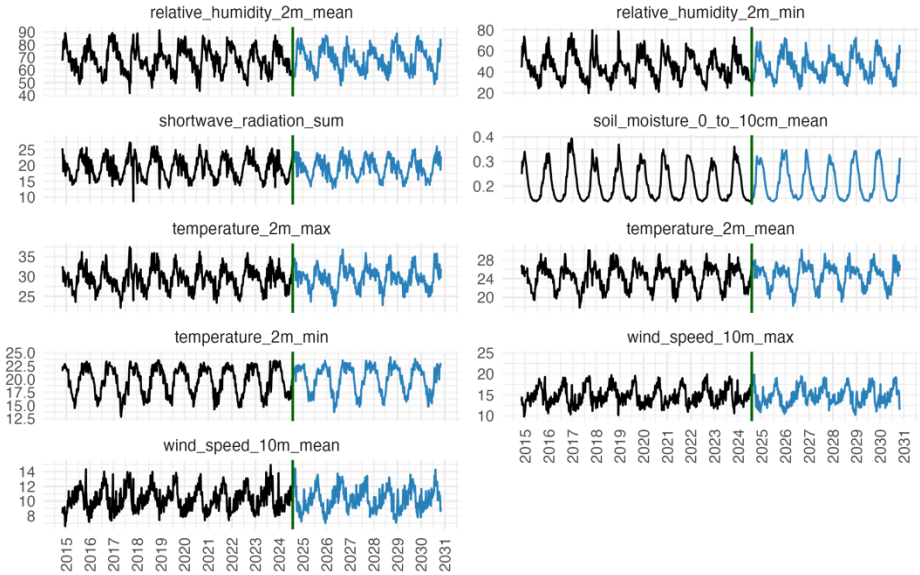


Within patch environment-to-human transmission

Environmental suitability

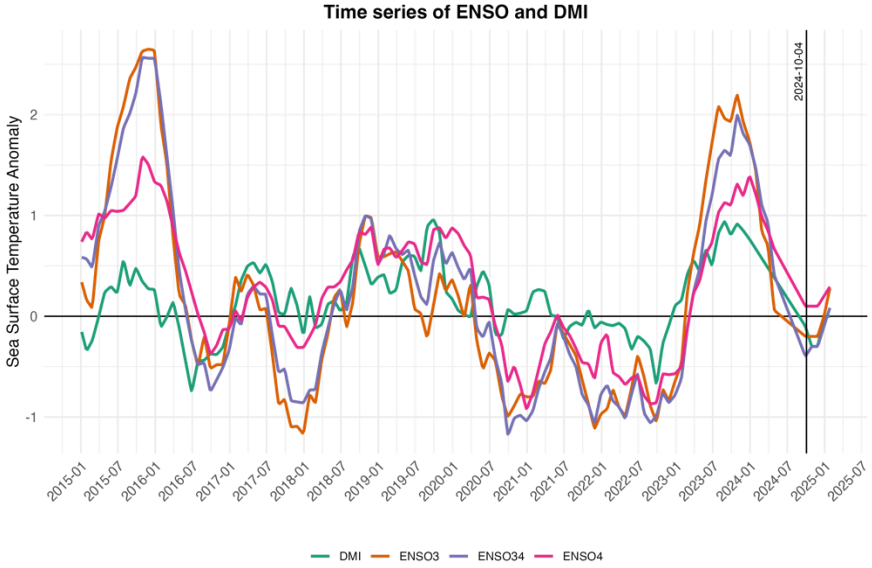
Climate data from Open-Meteo

- 19 weather variables (1970-2030)
- Based on MRI_AGCM3_2_S climate model from Meteorological Research Institute, Japan (MRI)



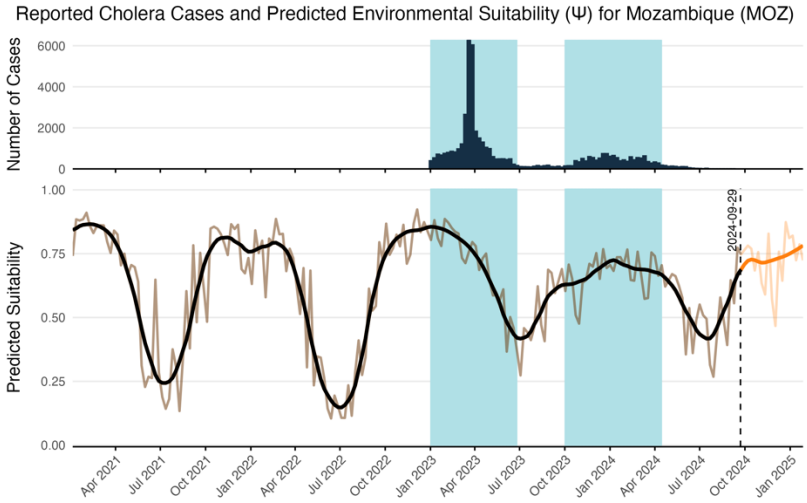
Regional climate drivers

- El Nino Southern Oscillation (3 Pacific Ocean regions)
- Dipole Mode Index (Indian Ocean)
- Projections from Bureau of Meteorology, Australia



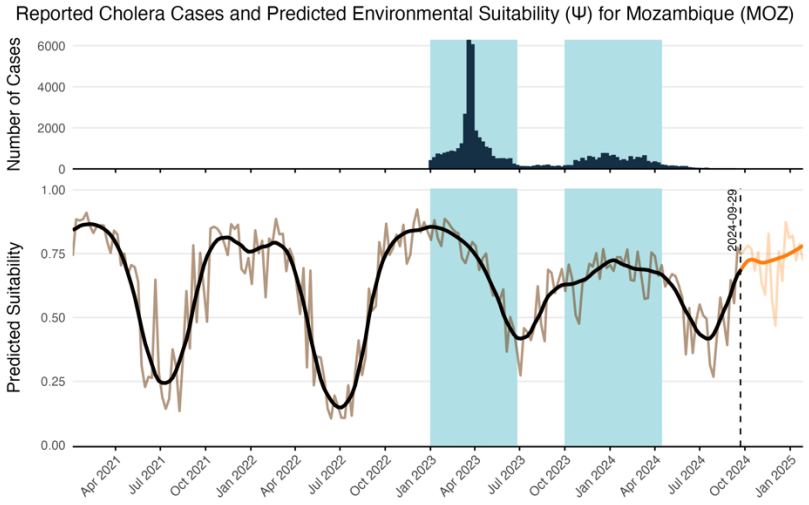
Environmental suitability

- Requires historical incidence data (currently 2023-2024)
- Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) model. Multiple network layers = deep learning AI.
- Predictions up to 5 months for all SSA countries

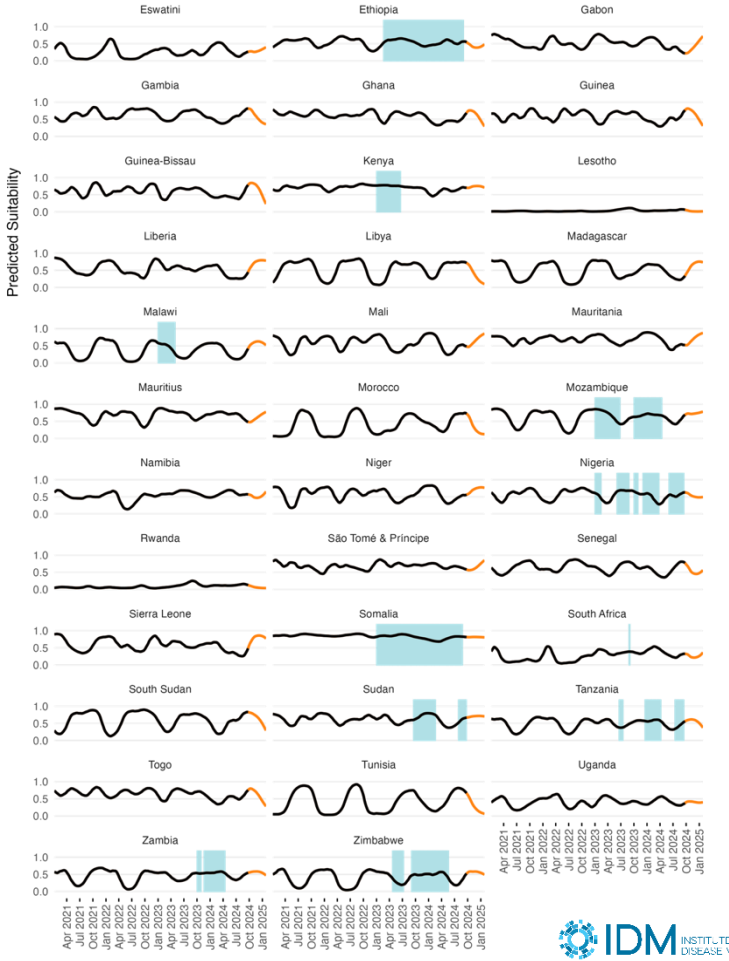


Environmental suitability

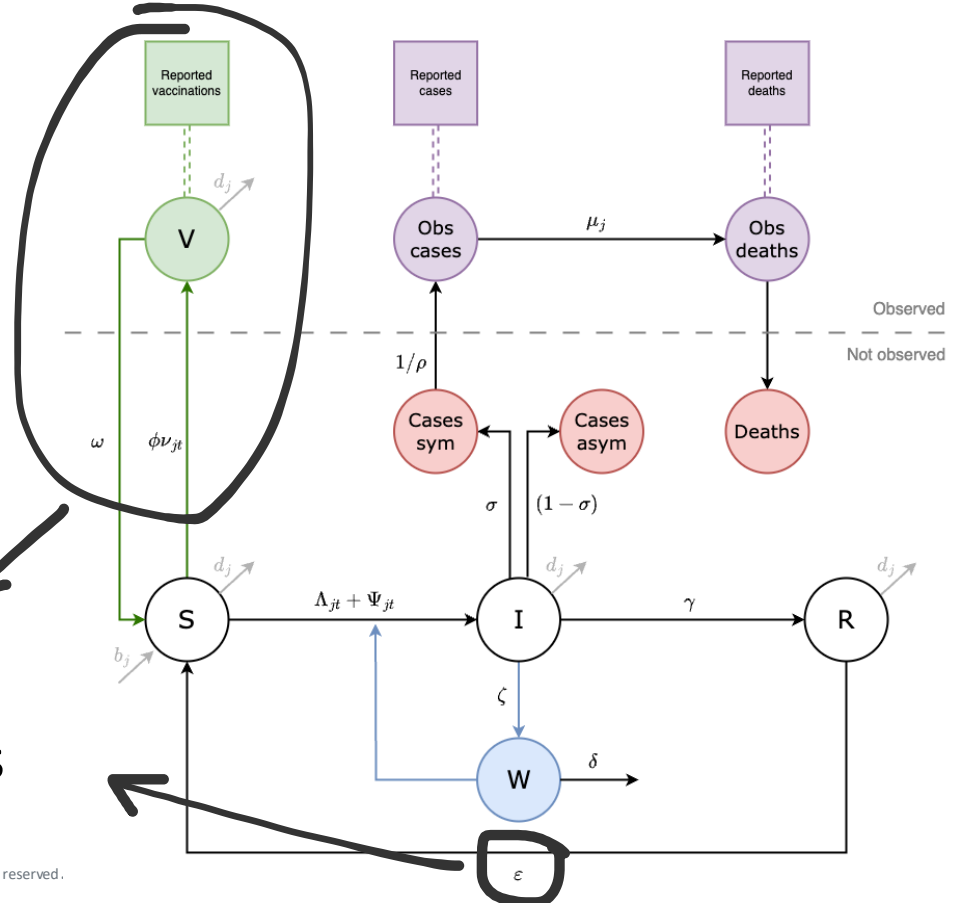
- Requires historical incidence data (currently 2023-2024)
- Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) model. Multiple network layers = deep learning AI.
- Predictions up to 5 months for all SSA countries



Predicted suitability for 40+ SSA countries



MOSAIC model structure: complex drivers require diverse data streams and model mechanisms



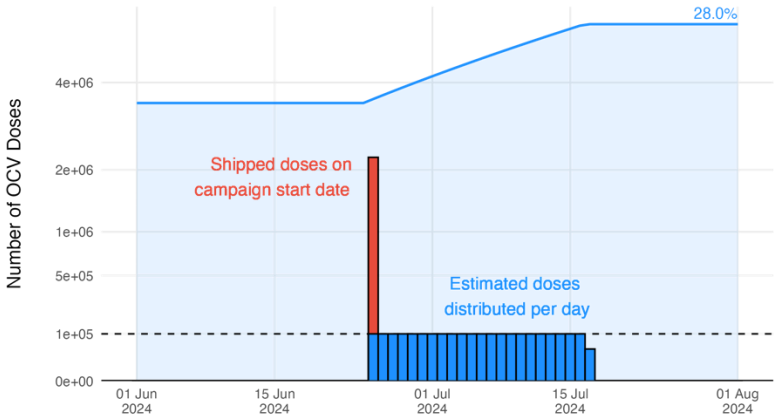
Immune dynamics

Vaccination data: reported OCV campaigns

Model of OCV vaccination rate is highly data dependent

- Data from WHO ICG OCV dashboard (2016-present).
- Single assumption about maximum daily rate of OCV administration during campaigns.

Example of distributional assumption (max = 100k doses/day)

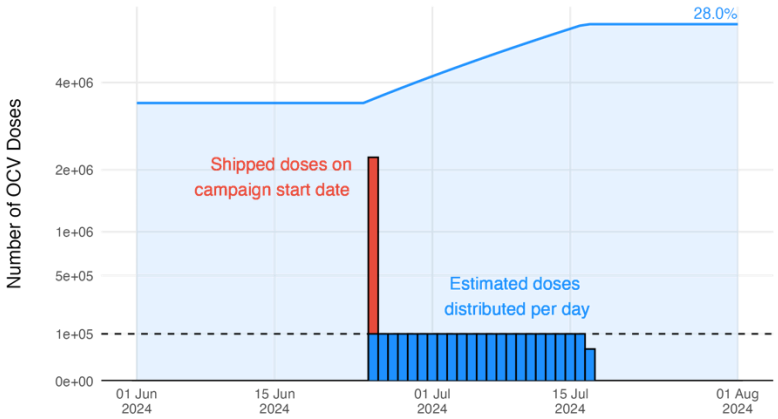


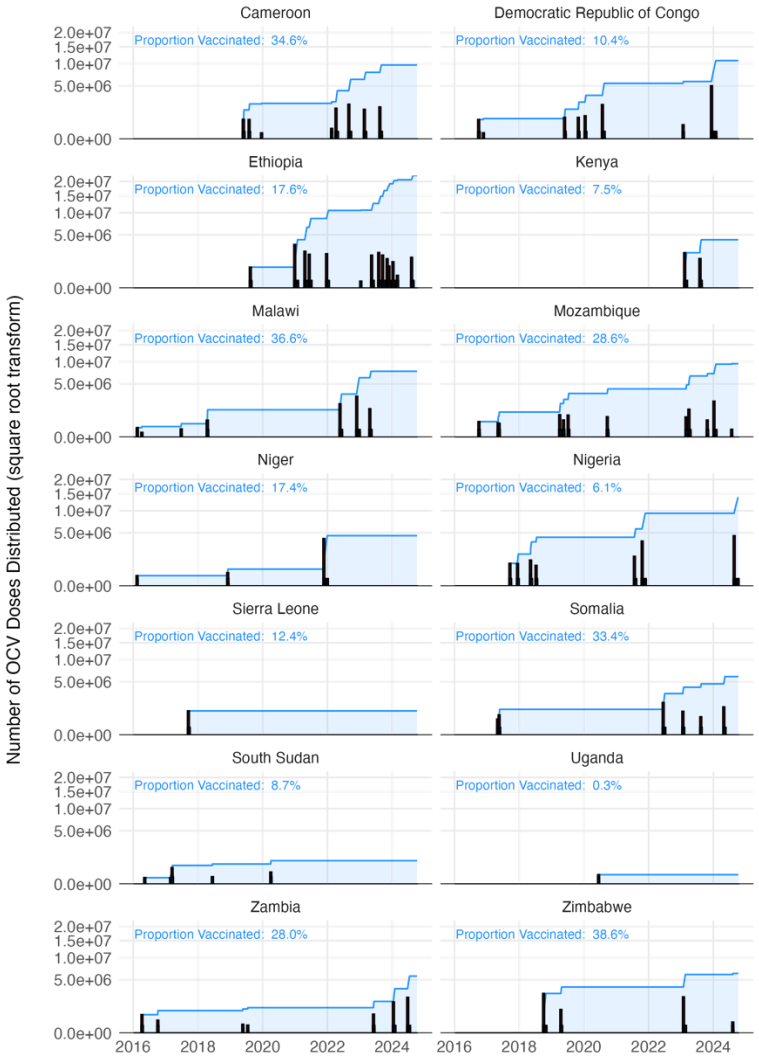
Reported OCV campaigns

Model of OCV vaccination rate is highly data dependent

- Data from WHO ICG OCV dashboard (2016-present).
- Single assumption about maximum daily rate of OCV administration during campaigns.

Example of distributional assumption (max = 100k doses/day)

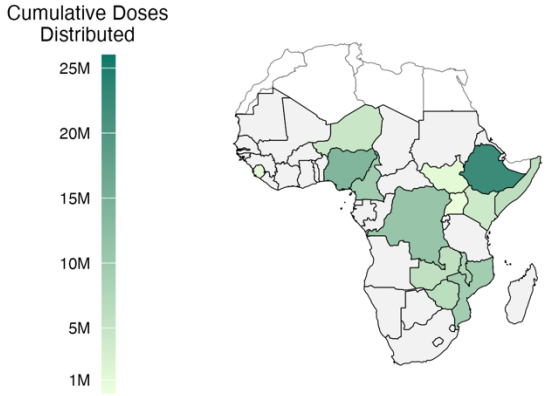


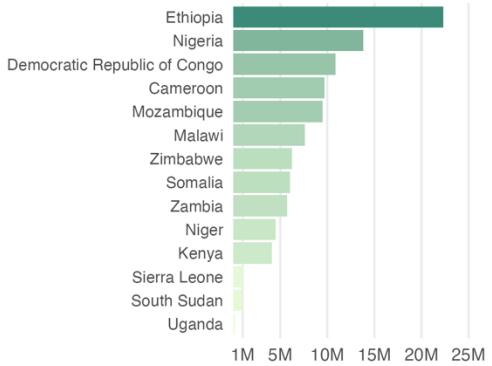


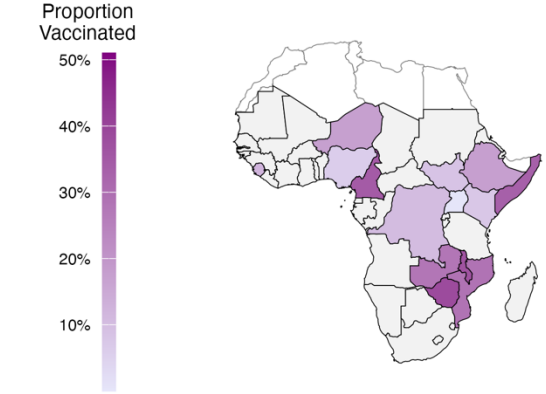
Reported OCV campaigns

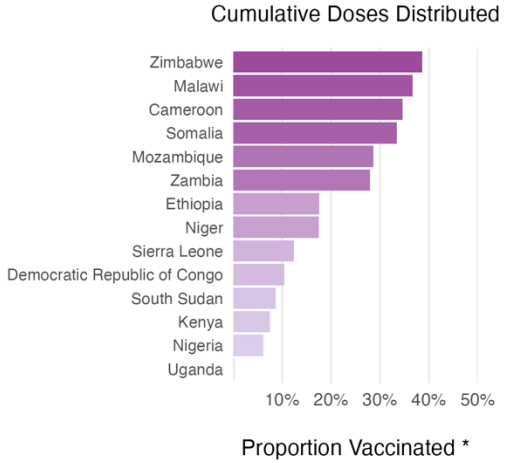
- WHO ICG data contains reactive OCV campaigns only (103M doses)
- OCV coverage underestimated by ~55M doses
- Working on combining with GTFCC vaccination data (reactive and preventative campaigns)

Reported OCV doses distributed by WHO-ICG as of October 09, 2024







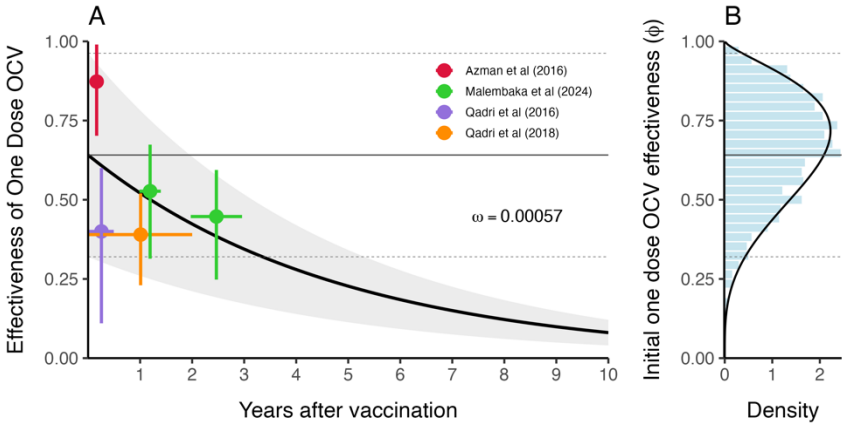


** Assumes single-dose OCV strategy*

Modeling immune dynamics

Immunity from vaccination

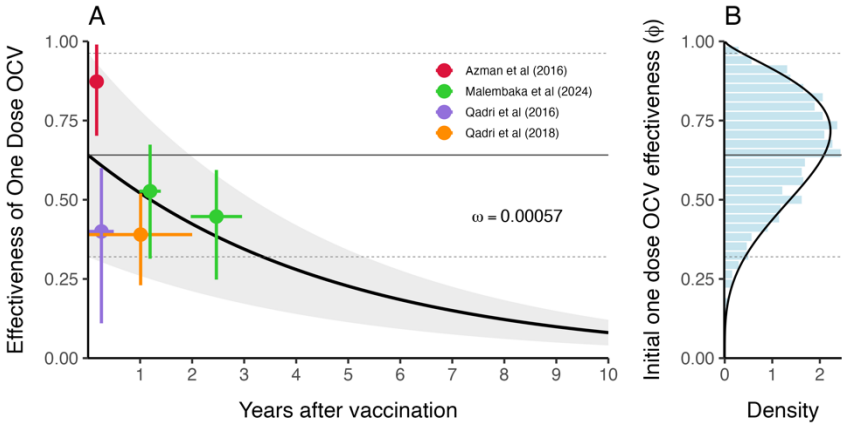
- Initial effectiveness of one-dose OCV
- Vaccine derived immunity decays over time (mean = 4.8 yrs)
- Relying on 5 observations from 4 studies



Modeling immune dynamics

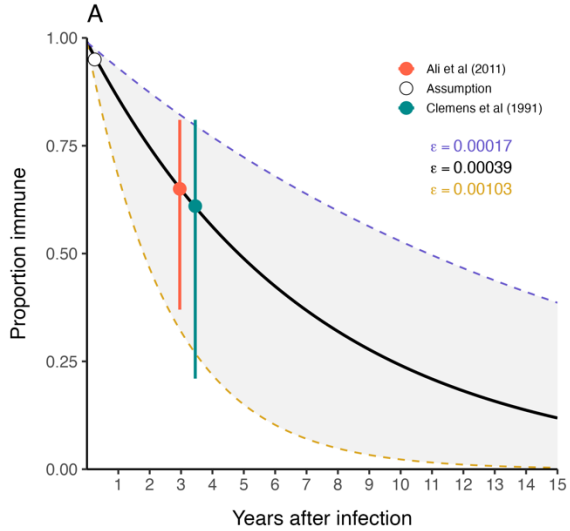
Immunity from vaccination

- Initial effectiveness of one-dose OCV
- Vaccine derived immunity decays over time (mean = 4.8 yrs)
- Relying on 5 observations from 4 studies

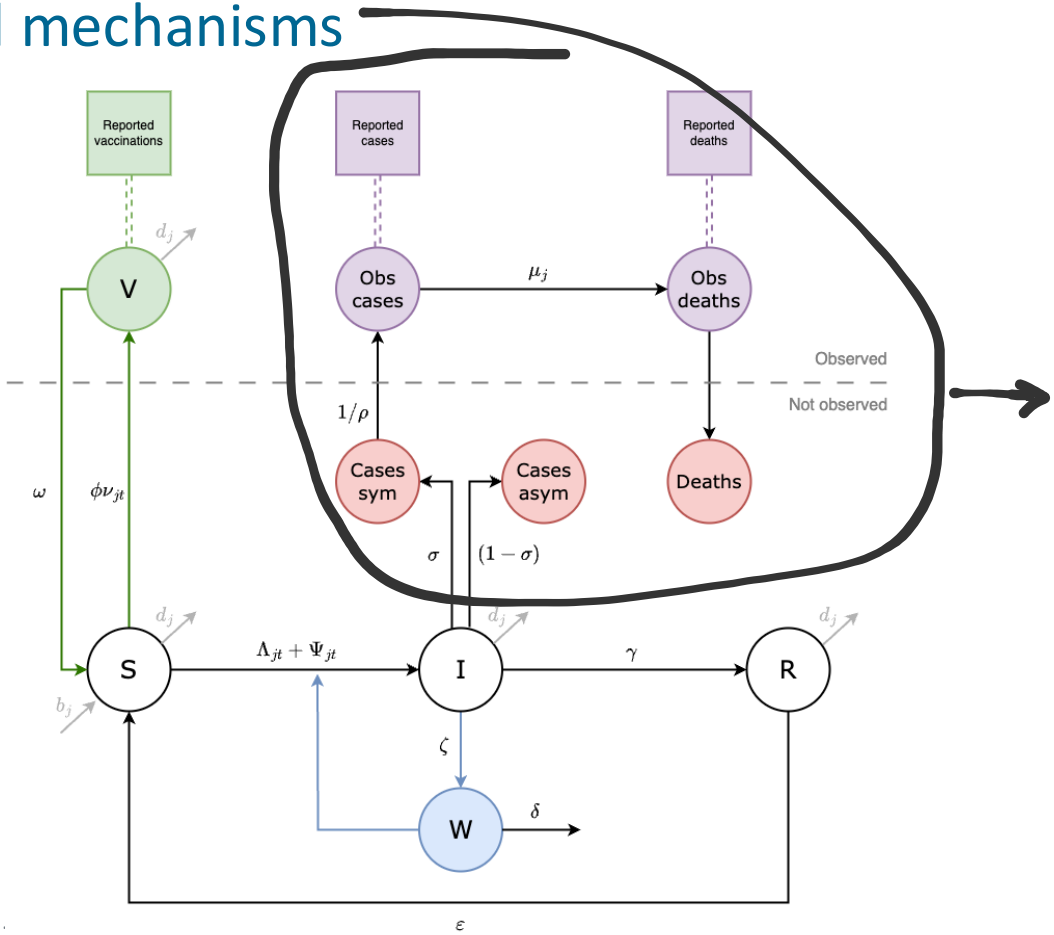


Immunity from infection

- Assumed >95% immune up to 90 days post infection
- Immune decay over time (mean = 7.2 yrs)
- Only 2 studies, seeking more with collabs

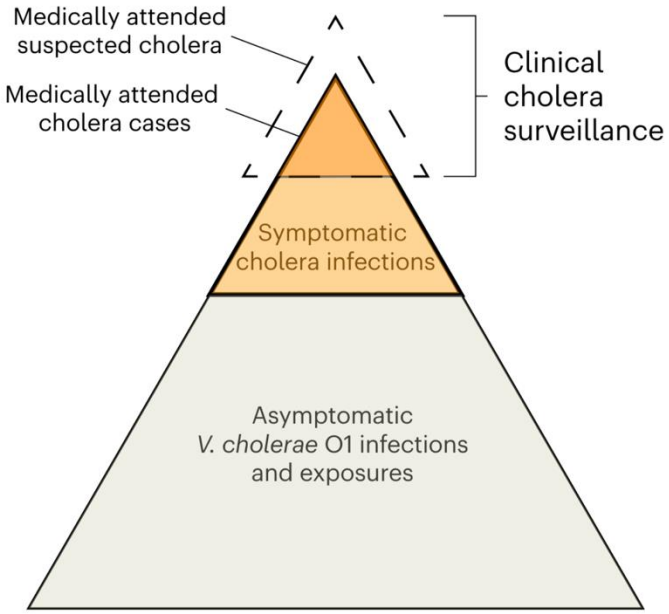


MOSAIC model structure: complex drivers require diverse data streams and model mechanisms



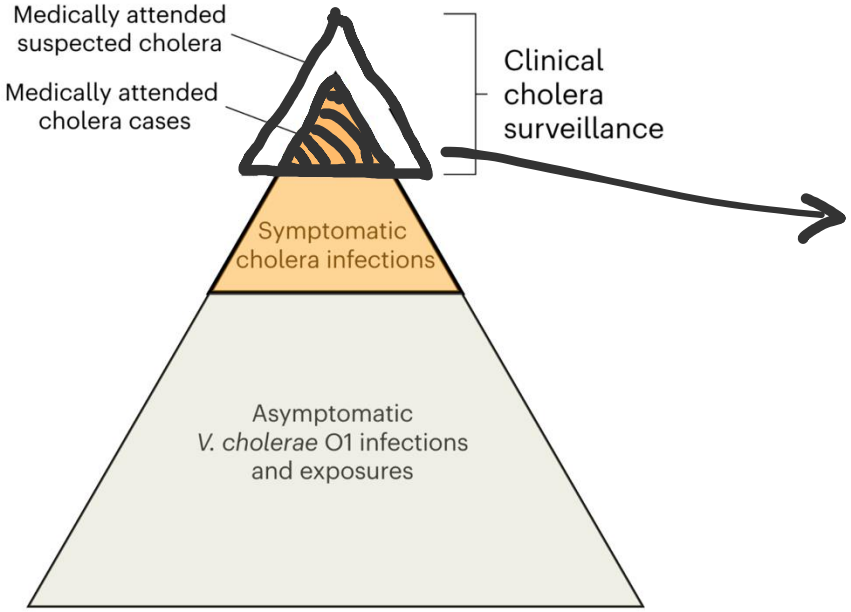
Observation process

Observation process



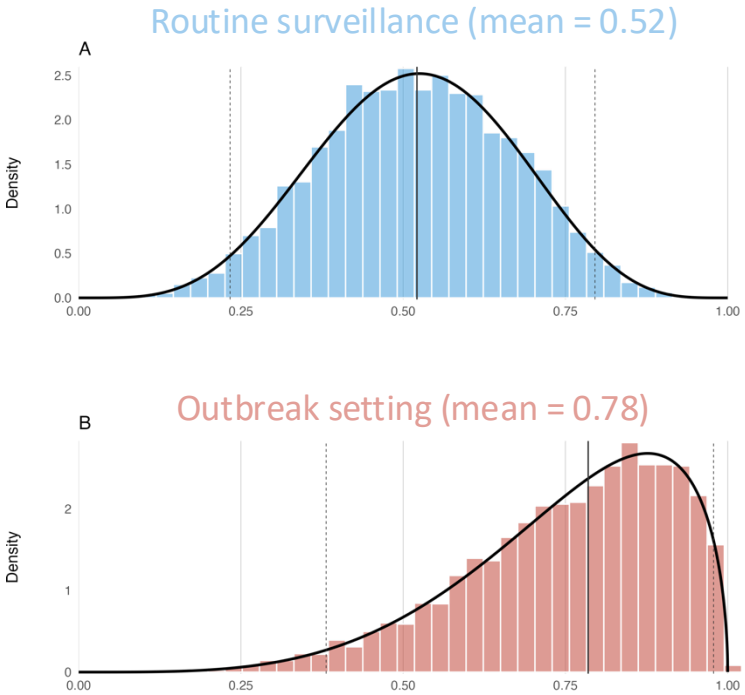
Hegde et al. 2023

Observation process

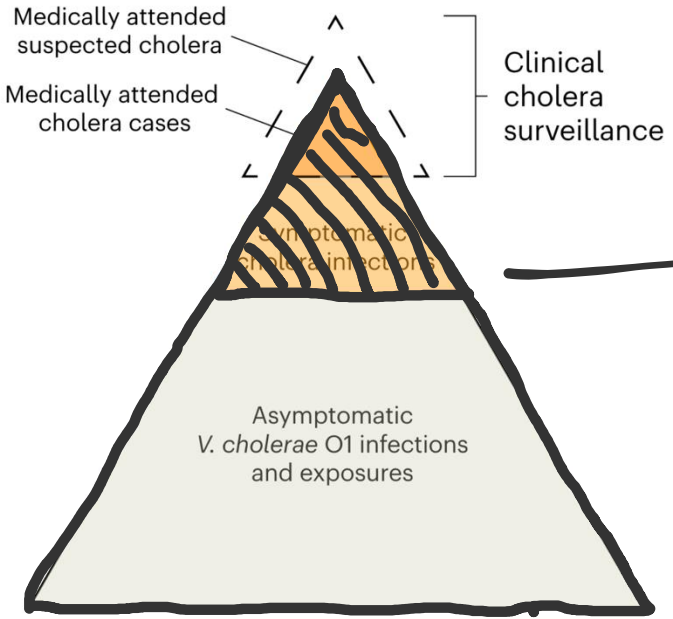


Hegde et al. 2023

Proportion of reported cases that are true infections (calibrated to data in Wiens et al. 2023)



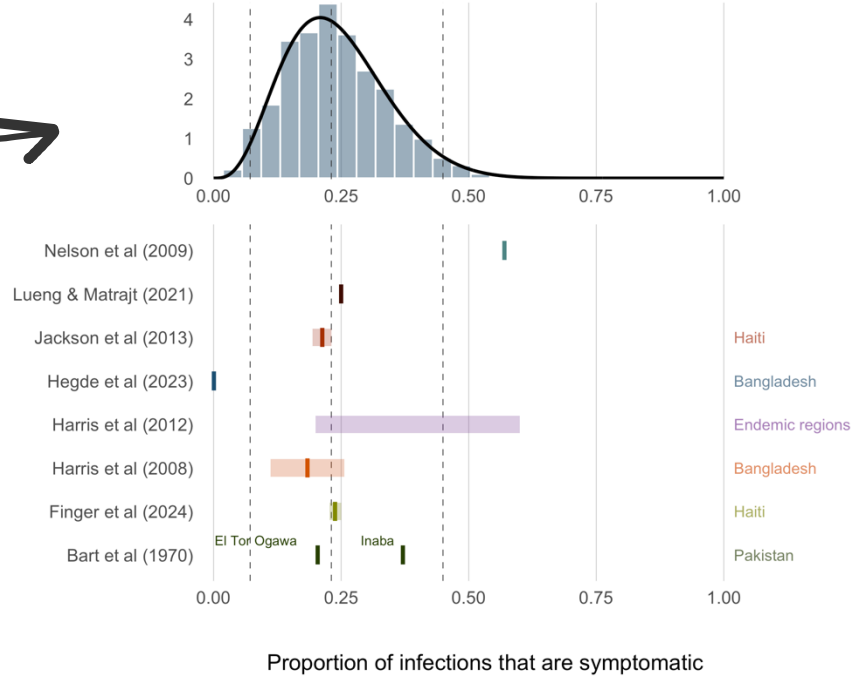
Observation process



Hegde et al. 2023

Proportion of true infections that are symptomatic

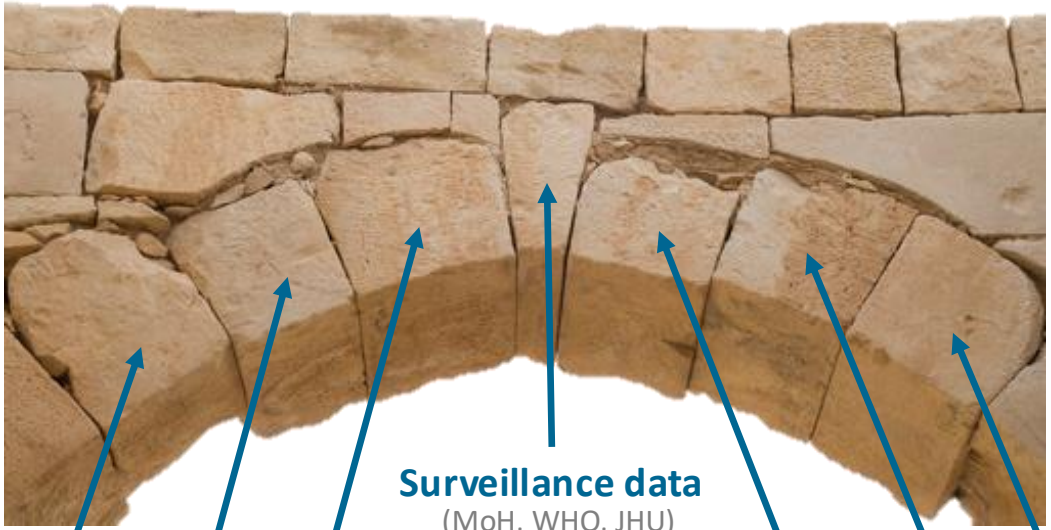
(Meta-analysis of 8 studies suggests mean = 0.24)



→ Spatial model of endemic cholera

Metapopulation Outbreak Simulation with Agent-based Implementation for Cholera (MOSAIC)

Diverse data streams required



Surveillance data
(MoH, WHO, JHU)

OCV
(WHO ICG, GTFCC)

Pop movement
(OAG, CDR, surveys)

WASH
(WHO UNICEF JMP)

Climate
(Open-Meteo, BOM)

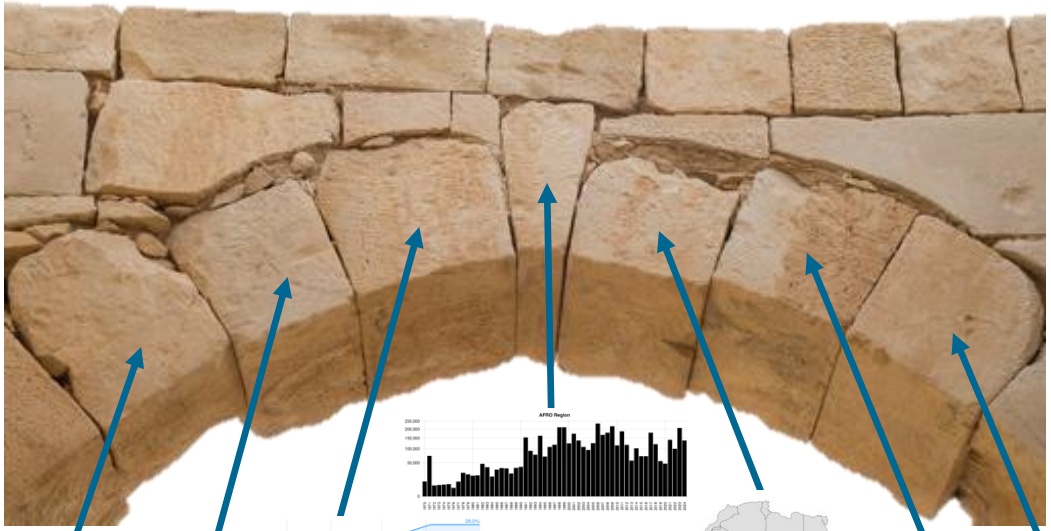
Immune dynamics
(Published studies)

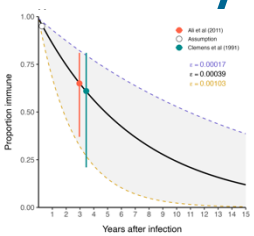
Observational bias
(Published studies)

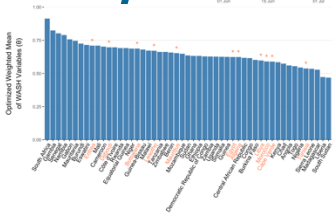
→ Spatial model of endemic cholera

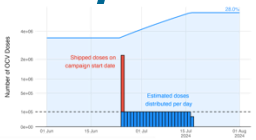
Metapopulation Outbreak Simulation with Agent-based Implementation for Cholera (MOSAIC)

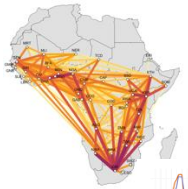
Diverse data streams required

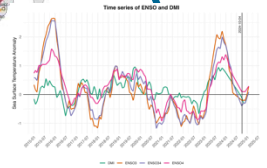


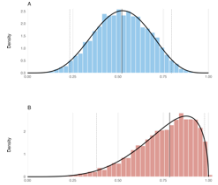












IDM INSTITUTE FOR DISEASE MODELING

Key take-aways: from data collection to forecasting

Reconstructing endemic cholera dynamics is challenging, and useful forecasting requires robust mechanisms for all major drivers.

Forecasting is data hungry: >10 data streams for model already. Real-time surveillance is vital (clinic + MoH + WHO + dashboard).

Modeling frameworks must be flexible and absorb new data streams (wastewater, sero-surveys relevant) or new situations (climate change/ conflict).

Modeling endemic cholera at scale is possible if we cultivate the right data and validate. Aim is to provide additional context around GTFCC goals and OCV campaigns.

Thank you

Collaborators

Jillian Gauld
Alicia Kray
Kyra Grantz
Rajiv Sodi
Jonathon Bloedow
Christopher Lorton
Vanessa Frances
Dejan Lukacevic
Jennifer Jo Simpson
Charles Eliot
Duncan Steele
Gagandeep Kang
Tanya Shewchuk
Kirsten Vannice
Stephanie Grow
Lyou-Fu Ma
Kevin McCarthy
Katherine Rosenfeld
Edward Wenger

Email: john.giles@gatesfoundation.org

MOSAIC

https://www.mosaicmod.org/