Applied data science for global health and COVID-19

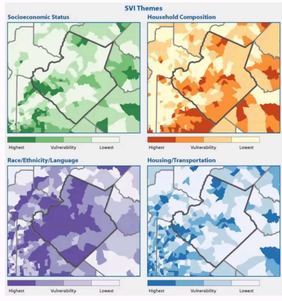
3/11/21 11:00 AM
Speaker: Peter Smittenaar (Surgo Ventures)
Date: 03/11/2021

In Mar 2020 we switched to develop tools and insights to support COVID-19 response
- US & Africa community vulnerability indices
- Toolkit to guide US policy makers and raise awareness of vulnerable pops
- Symptom-checker in Africa compatible with all phones
- Extensive COVID-19 vaccines portfolio (US & soon UK)

US COVID-19 Community Vulnerability Index (CCVI)

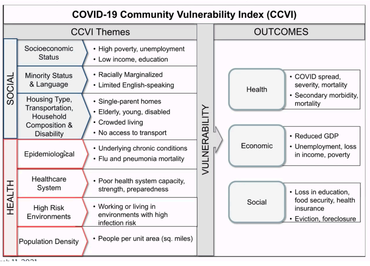
Why index?
- Forecasting models (SEIR) provided dire projections
- If COVID is going to hit everywhere, where do we expect greatest damage and what can we do about it?

Building on the CDC's social vulnerability index

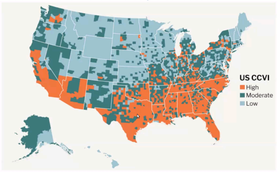


Rapid index construction: trade-offs
- Credible: extend validated and well-known SVI vs. start from scratch
- Practical: group indicators based on themes vs. relying on statistical properties
- Static: map of underlying drivers of vulnerability vs. ever-changing numbers (e.g. SEIR)
  - E.g. if two regions are all with high impact from COVID, but one is from social-economics reason, other is for poor health system. The strategies to control two areas should be implemented differently

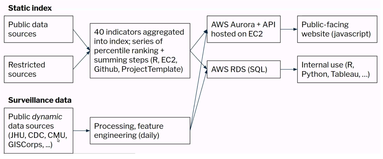
Overview of CCVI (40 indicators across 7 themes)



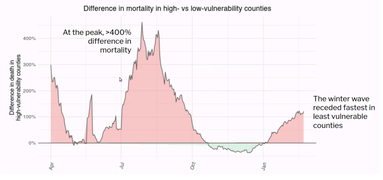
Overall vulnerability is majorly on the South
But drivers are more diversely distributed



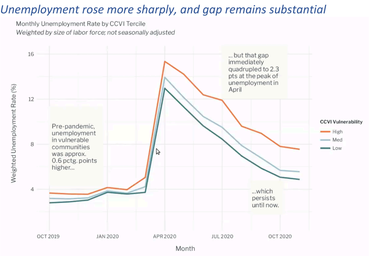
Index Construction



Inequity is not set in stone - varies over time
Difference in death in high-vulnerability counties are not stable



Unemployment rose more sharply, and gap remains substantial



Research scientists play a diverse role
- Do all the regular data science work, but also
  - Write documentation
  - Write reports that focus on the implications
  - Support press releases to tight deadlines
  - Field requests from stakeholders and journalists to 3-hour deadlines
  - Present non-technical decks to the stakeholders (over and over), understanding their needs

COVID-19 in Africa

There's a lot we don't know about COVID-19 in Africa

Entire pandemic:	Nigeria	US	
Tests per 1000	7.5	1020	**136x higher in US**
Cases per 100k	76	8652	**114x higher in US**
Deaths per 10pk	0.93	156	**168x higher in US**

Differences can arise from testing, age structure, comorbidities, policy, culture, travel, etc.

U.S vaccine roll-out

Our US COVID portfolio is designed to drive equitable allocation and sufficient uptake with precision at scale

Federal and state officials define priority groups, but:
- Where are these people?
- How much overlap exists between these groups?

We collected data on county level, for priority group defined by CDC.
But there's significant overlap (double-counting) in the groups

3142 counties x priority groups

Category	Data Source	Granularity	Complicatio n
Demographics	CDC, ACS 5-year	County	None
Health status	CDC, BRFSS	County	Overlap
Occupation	BLS QCEW, CA OES	County	Suppressed data
Institutionalizatio n	HIFLD, CMS, CDC, ACS 5-year, Vera Institute of Justice	County	None
Household Structure	CDC, ACS 5-year	County	None