How to incorporate research into routine data systems

The experience of the Western Cape Government Health Department

Dr Keith Cloete, Head of Department

COVID-19 Global research & innovation forum, 25 February 2022
Outline

Background to the WC

Key contributors to integrated service, surveillance and research platform
• Critical threshold of digital health systems
• Consolidated data environment – data competence
• Relationship with academia

Examples of surveillance and research outputs
• Associations with severity
• Variant analyses
• Vaccine effectiveness
Western Cape Province (1): Population context

7 million population

550,000 people living with HIV

50,000 people diagnosed with TB each year
Western Cape Province (1): Population context

7 million population

550,000 people living with HIV

50,000 people diagnosed with TB each year
Western Cape Province (2): COVID-19 burden
Western Cape Province (3): COVID-19 burden assessed by seroprevalence

Rapid assessment of seroprevalence after each wave by antibody testing of routine residual specimens from people attending health facilities for non-COVID-19 reasons e.g. diabetics and people with HIV.

Overall ~90% of adults had immune protection from prior infection and/or vaccination before wave 4.
Western Cape Province (4): Digital health context

Integrated data systems are not unique globally, but unusual in a context where most record keeping is still paper-based, and administrative systems, network bandwidth and stability are often suboptimal.

Critical digital enablers include

- **Unique identifier**, implemented imperfectly and incrementally over 20 years but covers all fixed health services at point of patient registration with a barcoded numeric identifier affixed to all stationery
- Key **high-throughput administrative systems** digitized and consolidated
  - National Health Laboratory Service and NICD are incredible national assets
  - Electronic Vaccine Data System digitized all vaccinations from program inception
  - Public and private sector hospital admissions/discharges/transfers in the province were all digitized
  - Many other administrative data sources such as pharmacy dispensing data
- **Consolidated data environment**
Disease monitoring systems (eg HIV / TB)

Laboratory and pharmacy data

Hospital and primary care registration systems

Population register

Many other systems

**Consolidated data environment**

- **Health information exchange or Data Centre**
  - Care cascades and operational reports
  - Alerting engine (eg. NMC’s)
  - Management reporting
  - Epi analyses
  - Business intelligence
  - Public data accessibility and accountability
  - Data governance, research support and stewardship

- **Direct clinical use**

- **Consolidated data environment**
  - the Provincial Health Data Centre
    - Focused on clinical care first and foremost, leveraged through a Single Patient Viewer
    - Entirely government managed and housed
    - Privacy concerns dealt with by care remit as a service provider, with all data linkage done on the operational service data
    - Only de-identified linked data released for analysis in line with standard research governance procedures
    - Data competence trumps specific technology – having a team in place enabled rapid pivoting to COVID-19

Boulle, IJPDS, 2019
PHDC “Cascades” or virtual cohorts as the basis for action, surveillance and research

Cascade for a health condition

Epidemiological analyses

Single patient viewer for clinical care

Public-facing dashboards

Service Mx report/internal dashboards

Line list/action list

Alerts
Relationship with academic institutions & other partners a key enabler

• Clinical academics in SA are mostly appointed on government service contracts with university co-appointments. While often there are tensions between service and academic responsibilities, the model can be highly synergistic.

• Senior clinical investigators who have worked for decades on HIV, TB and other health conditions, helped lead the service response to COVID-19, and also embedded cohort and vaccine effectiveness studies into the service response.

• Platform clinicians, pathologists and public health specialists co-appointed with local universities drew in academic resources directly into the surveillance effort, including mathematical modellers, biostatisticians, molecular virologists, and bioinformaticists.

• Phenomenal national collaborations working across institutions and sectors, in a previously unprecedented manner
COVID-19 Dashboard


2nd wave (Beta) 3rd wave (Delta)

Number of daily dashboard views
Disease Epidemiology – HIV/TB & COVID-19

• Early on in the pandemic risk factors of male sex, increasing age and NCDs were established.

• No data on those risk factors in an African setting – rapidly confirm these risk factors with clinical interventions e.g. enhanced telemedicine cases management for diabetics.

• First study to demonstrate increased risk of severe COVID-19 in PLHIV and those with tuberculosis – subsequently confirmed in UK and US studies.
Variant disease severity

- Wave comparisons as proxy for variants
- Link clinical and laboratory data to compare disease severity of different variants

→ Using proxy markers based on PCR results e.g. S gene target failure for Omicron

→ Linking with Whole Genome Sequence data from Network for Genomic Surveillance in South Africa
Vaccine effectiveness

• Linkage to data on vaccination from the national Electronic Vaccine Data System (EVDS) enables rapid assessment of vaccine effectiveness in the context of different variants and comorbidities.

Risk reduction of severe disease outcomes in Omicron period with full vaccination vs none
Summary

• Western Cape COVID-19 surveillance and epidemiological outputs are not unique in the global context, but have been achieved in a context where many of the digital health assets of wealthier countries are not yet in place.

• There are a few critical digital health elements which enable functional integrated data systems in advance of pervasive electronic medical records, including unique identifiers, digitized laboratory data, and a consolidated data environment characterized by staff with the data competence to work efficiently and flexibly with large and imperfect individuated datasets.

• Strong relationships with academic institutions, other partners and networks are a key enabler, often mediated through joint service/academic staff.
Thank you