NHLS Data Strategy

Professor Wendy Stevens
Department of Molecular Medicine and Haematology
University of the Witwatersrand, National Priority Program,
NHLS, Johannesburg, south Africa
Introduction

- NHLS has a unique repository of laboratory information
  - The data is representative of pathology information for +/- 80% of the South African population
- 270+ sites across South Africa
- Single Laboratory System
- Single Data Warehouse
Laboratory Information Systems

- **TrakCare Lab**
  - used in all labs in NHLS, NICD and NIOH
- **Country-wide since Oct 2015**
- **Technical**
  - Application hosted on Citrix Application servers
  - Supported by Caché Database
  - Uses Ensemble Integration engine
- **Network**
  - To all sites
  - MPLS implementation
  - SITA CAP connection

- All instruments are interfaced through a LIS (4 regional centers and connected through wide area network)
- CDW raw data is aggregated
- A disaster recovery site is based within 20km.
- Hospital information systems (HIS) use HL7 protocol and interface to the data center.
- All patient laboratory results are made available to HCW via an internet service provider (ISP) such as [www.labresults.nhls.ac.za](http://www.labresults.nhls.ac.za).
- Virtual private networks (MPLS VPN) and other network traffic are routed for mHealth, point of care, SMS printer devices, etc
Information held in CDW

- Laboratory information:
  - Disa*Lab – legacy laboratory information system
  - TrakCare Lab – new laboratory information system
  - Information from private laboratories

- Business Information
  - Financial information
  - Lab staffing information

- National statistics
  - Population statistics
  - Prevalence surveys (e.g. antenatal HIV)
  - DHIS information
Surveillance and Corporate data warehouse

In future: Algorithm integration into SDW and CDW
Data Capabilities

- Process +/- 300 million raw data records per month
- Perform extensive data cleansing and transformation
- Record linking – Laboratory information systems are specimen-centric and therefore need to build a patient or case view of the data
  - Implemented data linking program
  - Uses fuzzy logic to evaluate demographic attributes and assign a probabilistic % match
  - Automatically allocates a patient identifier to all matches scoring above a given threshold
- Reporting - analytical, formatted, integrated spatial & dashboard reporting is available
Facility coordinate cleaning

- **Process in constant Evolution**
- **Simple exercise of mapping**: clinics, laboratories, drive time, volumes, results, TAT, analyzer volumes
- **Service Mapping Network**: logistics: route selection, physical location, GIS
  - Mapping all this into an LIS/LIMS
  - Mapping **instruments, tests and test codes** to LIS
  - Mapping demographics and result to location (LIMS)
- **Added layer of Sophistication**
  - Mapping to a **Central Data Warehouse (CDW)**
  - “Geospatial mapping”: fixed and now mobile
Across the National Priority Program

80% population, networked, instruments interfaced - 1 LIS

**49 CD4 labs**

**16 VL labs**

**203 GeneXpert labs**

---

### HIV Viral Load

- **Volumes**
  - FY 10/11: 1,219,1
  - FY 11/12: 2,047,7
  - FY 12/13: 3,018,3
  - FY 13/14: 3,782,0
  - FY 14/15: 4,508,6
  - FY 15/16: 5,048,2
  - FY 16/17: 4,018,2
  - FY 17/18: 3,708,2

- **FY Volumes**
  - 2013: 1,219,1
  - 2014: 2,047,7
  - 2015: 3,018,3
  - 2016: 3,782,0
  - 2017: 4,508,6
  - 2018: 5,048,2

- **Total Volumes to date:** 23,879,266

### CD4 Count

- **Volumes**
  - FY 10/11: 3,533,0
  - FY 11/12: 3,914,6
  - FY 12/13: 3,838,5
  - FY 13/14: 3,915,2
  - FY 14/15: 3,904,7
  - FY 15/16: 3,497,0
  - FY 16/17: 3,389,4
  - FY 17/18: 3,058,2

- **FY Volumes Increase**
  - 2013: 10.80%
  - 2014: -1.94%
  - 2015: 2.00%
  - 2016: -10.44%
  - 2017: -3.08%

- **Total Volumes to date:** 29,051,003

---

### HIV PCR

- **Volumes**
  - FY 10/11: 277,092
  - FY 11/12: 314,393
  - FY 12/13: 335,597
  - FY 13/14: 353,976
  - FY 14/15: 378,037
  - FY 15/16: 510,791
  - FY 16/17: 571,812
  - FY 17/18: 590,301

- **FY Volumes Increase**
  - 2013: 13.46%
  - 2014: 6.74%
  - 2015: 5.48%
  - 2016: 6.80%

- **Total volumes to date:** 3,331,999

---

### Xpert MTB/RIF

- **Volumes**
  - FY 10/11: 296,501
  - FY 11/12: 835,164
  - FY 12/13: 2,048,7
  - FY 13/14: 2,463,9
  - FY 14/15: 2,593,5
  - FY 15/16: 2,412,2
  - FY 16/17: 2,188,8

- **FY Volumes Increase**
  - 2013: 181.67%
  - 2014: 145.31%
  - 2015: 20.26%
  - 2016: -6.99%

- **Total Volumes to date:** 12,838,921
Data applications beyond result reporting

- Operational dashboards
- Surveillance
- Gaps in services (coverage)
- GIS mapping for route optimization
- Placement of instruments within the tiered laboratory framework
- Placement of POC
- HIV Longitudinal cohort development (program success)
- Integrate HIV/TB cohort
- Molecular granularity for TB control
Program value
Building the NHLS National HIV Cohort

National HIV Cohort

- # Lab Tests: 98 M
- # Specimens: 71 M
- # Patients: 11.6 M
- # on ART: 4.2 M

Building a National HIV Cohort from Routine Laboratory Data:
Probabilistic Record-Linkage with Graphs

Jacob Bor, William MacLeod, Katia Oleinik, James Potter, Alana T. Brennan, Sue Candy, Mhairi Maskew, Matthew P. Fox, Ian Sanne, Wendy S. Stevens, Sergio Carmona

doi: https://doi.org/10.1101/450304
Research using the cohort (example 1)

Clinical Infectious Diseases

Persistent High Burden of Advanced HIV Disease Among Patients Seeking Care in South Africa’s National HIV Program: Data From a Nationwide Laboratory Cohort

Sergio Carmona, Jacob Bor, Cornelius Nattey, Brendan Maughan-Brown, Mhairi Maskew, Matthew P Fox, Deborah K Glencross, Nathan Ford, William B MacLeod

Clinical Infectious Diseases, Volume 66, Issue suppl_2, 1 April 2018, Pages S111–S117, https://doi.org/10.1093/cid/ciy045
Published: 04 March 2018

Figure shows the share of patients presenting with CD4<100 (red) and CD4 100-199 (blue) over time.
Research using the cohort (example 2)

**PLOS ONE**

**RESEARCH ARTICLE**

CD4 count recovery and associated factors among individuals enrolled in the South African antiretroviral therapy programme: An analysis of national laboratory based data

Tendesayi Kufa, Zara Shubber, William MacLeod, Simbarashe Takuva, Sergio Carmona, Jacob Bor, Marelize Gorgens, Yogan Pillay, Adrian Puren, Jeffrey W. Eaton, Nicole Fraser-Hurt

**Figure** shows CD4 trajectories among men (bottom), women (top), and both (middle) on ART.
Research using the cohort (example 3)

Analysis of Big Data for better targeting of ART Adherence Strategies

Spatial clustering analysis of viral load suppression by South African province, district, sub-district and facility (April 2014 – March 2015)

November 2015

William MacLeod, Jacob Bor, Kathryn Crawford, and Sergio Carmona with NDOH and World Bank collaborators
Research using the cohort (example 4)

Estimating retention in HIV care accounting for patient transfers: A national laboratory cohort study in South Africa

Matthew P. Fox¹,²,³,⁴, Jacob Bor², Alana T. Brennan²,³, William B. MacLeod²,³, Mhairi Maskew⁵, Wendy S. Stevens⁶,⁷, Sergio Carmona⁴

Figure shows effect of patient transfer on retention estimates in South Africa among patients starting ART in 2004-2006. (n=55,836)
Research using the cohort (example 5)

The Adolescent HIV Treatment Bulge in South Africa’s National HIV Program: a Retrospective National Cohort Study

Mhairi Maskew, Jacob Bor, William MacLeod, Sergio Carmona, Gayle G. Sherman, Matthew P. Fox

In press

THE LANCET HIV
Extensions of the HIV cohort linkage

1. How does the algorithm perform for other conditions?
   - Ongoing work with the NPP and NICD TB groups to link and validate for TB (BU/WITS R01 under review)
   - Recently funded work with Jaya George and Nigel Crowther to look at NCDs (Alana Brennan K01)
   - Research opportunities to look at HIV/TB/NCD multi-morbidity

2. Can performance be improved for infant and paediatric cases?
   - Ongoing work with Gayle Sherman

3. Can the algorithm be integrated into real-time linkage in the data warehouse, and how does it perform when scaled to all conditions?
   - Ongoing work with NICD
   - For pragmatic reasons, pilot at SDW, then move to CDW if successful
Potential future opportunities

• Can NHLS National HIV Cohort be linked to TIER and other data sources to build integrated cohort to improve surveillance/research?
  • We have proposed some pilot work
• Can algorithm be leveraged to improve delivery of patient care?
Lots of Potential for Collaborative Research

• Here are some focus areas: (Are these the right ones? Are there others?)

1. VL monitoring and treatment outcomes
2. Paediatric HIV and pregnancy cohort
3. HIV drug resistance
4. HIV advanced disease
5. TB and HIV/TB co-infection
6. Non-communicable diseases
7. Data linkages and technical
• CDW (March 2011 – June 2019)
  • > 12.3 million Xpert MTB/RIF tests
  • > 3.2 million Xpert MTB/RIF Ultra tests

• C360 (June 2016 – July 2019)
  • >1.8 million Xpert MTB/RIF tests
  • >2.4 million Xpert MTB/RIF Ultra Tests
Data Architecture

- Data is transferred in two pathways:
  1. TrakCare, Laboratory Information System (LIS), Central Data Warehouse (CDW)
  2. C360 – Cepheid’s platform.

- The data from CDW is comprehensive

- The data from C360 is at the laboratory level
Size of circles indicative of data size and the data can be further categorised into instrument, assay, user groups.
How is the data used and potential future

- TB Dashboards for the NDoH
In the absence of a unique ID?

The (Potential) Promise of Biometrics

HIV Cohort

TB Cohort (in development)

Biometrics Considerations:
- Database infrastructure (speed of queries)
- Data ownership (security)
- Identification vs Verification Models
- Mobility
How is the data used and potential future

Monitoring of the TB Cascade at a National level: Xpert MTB/RIF 2011-2018

Monitoring of the TB Cascade at a Provincial level: Xpert MTB/RIF 2013-2015 in the North West
How is the data used and potential future

- Operations
Continuous quality monitoring

Dashboards

Molecular granularity

Data curator, storage infrastructure, interfacing, maintenance, hosting, backup
Research concept and design and partnerships
Data access approval and ethics (security) process
Data dictionary, analytics and visualization, tools
Stakeholder engagement and communications, dashboard access (download speed)
Expanding the laboratory data backbone

- The “information chain” – poor communication between referring clinics and central labs, increases TAT and decreases quality of care.
- Barriers must be addressed
  - Quick and efficient test requests
  - Efficient transport of specimens to central testing labs
  - Standardise communication systems to rapidly transmit results back to clinics (HCW and patients)
  - Workflow automation and monitoring

eLABS – backbone to a connected laboratory
Adoption Statistics: South Africa

- 41,500 VL Tests Requested
- 37,650 Results returned to Clinics
- 90% Suppression Rate
- 87% Results read by Clinics
- 4 days AVG Total Turnaround Time

Project Start Date: October 2018

Reduction in TAT at facility by 62%
Registration decrease of 57%
Reduction in LSS by 50%
Specimen rejection reduced by 1% (2%)%

Data Updated: 17 July 2019
“With iThemba, my health is in my hands.”
iThemba pilot data (as of June 24th 2019)  

**Enrollments in April, May and June**

1. **The Power of Mobile**  
   Harnessing growth of smart mobile access to transform healthcare delivery

   **Overall**  
   501 / 706 (70.9%)  

   **Enrolled / Screened**  
   - Hillbrow: 238 / 360 (66.1%)  
   - Yeoville: 263 / 346 (76.0%)  

   *Some patients are not accounted for in the count due to various IE criteria. Only criteria (5 – 10) are counted in this bucket.*

2. **The Power of Simplicity**  
   Simple, useful solution, appreciated by users

   - **Login Rating**  
     - 454: 5%  
     - 43: 17%  
     - 78%  

   - **Barcode Rating**  
     - 127: 27%  
     - 72%  

   - **VI. Result Rating**  
     - 78: 82%  
     - 5%  

   - **Never Rated**  
     - 241: 12%  
     - 13%

3. **The Power of Knowledge**  
   Closing the gap in knowledge of HIV viral load status

   - Blood Draw & Barcode Scan: 500  
     - Median days = 5 (range: 1-31d)

   - Result Released & Notification Sent: 458  
     - Median days = 0.5 (12.5hr) (range: 1-55d)

   - Result Viewed in App: 347  
     - (75.7%)

   - iThemba median days = 6 (range: 1-56d)

   - Blood Draw  
   - Results Sent to Clinic  
   - Patient Travels to Clinic to Get Results  
   - Patient Receives Results

   - SoC median days = 56 (range: 10-430d)

4. **The Power of Community**  
   Engaging users with virtual communities to provide support

   - **Total Joined**  
     - 208 (41.5%)

   - **Hillbrow**  
     - 151 (63.4%)

   - **Yeoville**  
     - 57 (21.6%)

   *Denominator for percent calculation is based on each site’s virtual community enrollment*
iThemba
Empowering patients to remain adherent to treatment and engaged in care
Laboratory as the command centre

- Bi-directional Push and Pull of information
- Continuous real-time monitoring
- Clinical relevancy and footprint expanded by clinical partners
- Priorities to be determined by stakeholders and partner consultations
- Complexity will increase beyond the proof of concepts
- Platform for innovation and social entrepreneurship
- Significant investment required for maintenance
Considerations

• The CDW storing all public sector laboratory results is indeed a national treasure and needs to be resourced.
• The value of the laboratory for individual patient management has always been relatively clear.
• New roles:
  • The value of aggregated laboratory data has been demonstrated in the projects presented today: programmatic value.
  • The role of the laboratory in proactive enhancement of linkage to care via solutions such as mHealth, webview etc is in its infancy requires investment.
  • Facilitates a platform to support POCT initiatives.
  • Monitoring of analyzer performance from a distance.
Acknowledgements

• National Department of Health (HIV and TB cluster)
• NHLS National Priority Program
• Department of Molecular Medicine and Haematology, Wits University R&D team
• NHLS, National Priority Program, TB/HIV Expert working groups
• Funders (specifically CDC, Bill and Melinda Gates foundation, NIH/ACTG, USAID, MRC/Newton)
• Clinical partners Right to care, Aurum, TB/HIV care, WRHI, PHRU
• Research collaborators: HERO (BU),
• Commercial collaborators (Cepheid, Abbott, Roche, Hain, BD)
• Innovators