Biometrics for patient identity management

Alexandra Grigore - CPO
According to a study by UNAIDS of Brazil, of an estimated 140 million records in the national user database only 100 million unique patients were represented.

Source: It's a puzzle, it's an algorithm, it's deduplication
Unreliable data

In vaccine coverage, the gaps between administrative data and WHO surveys in Chad is over 50%.

Source: Gavi Chad Audit 2019
Biometrics can ensure unique patient records across digital health systems
Introduction to Biometrics
What are some examples of biometrics?

- Face
- Fingerprint
- Iris
- Palm/Vein
- Gait
- Signature
- Keystroke
- Voice
How are biometrics used?

1. Client registration
   - Medical record
   - Biometrics + UID
   - Optional: Physical credentials

2a. Duplicate prevention
   - Biometric matching (1:N or N:N)

2b. Duplicate removal
   - Biographic matching (1:1)
   - Duplicates removed
Identification vs. verification

3. Identification

- Biometric matching (1:N)
- Medical record retrieval

4. Verification

+ Physical credentials

- Biometric matching (1:1)
- Medical record verification
Biometrics system components

Image capture and feedback

Image quality errors

Image quality parameter

Feature extraction

Template size & reversibility

Matching 1:1 or 1:N

Matching accuracy
How do we measure accuracy?

- **Match score**
  - similarity
- **Threshold**
  - match?

- TP = True Positive
  - should match
- TN = True Negative
  - should reject
- FP = False Positive
  - unintended match
- FN = False Negative
  - unintended reject
Five metrics to assess accuracy

FTE: Failure to Enroll
FAR: False Accept Rate

FTA: Failure to Acquire
FRR: False Reject Rate
TPIR: True Positive Identification Rate
Verification metrics

Proportion of people who didn't match but should have

False Non-Match Rate

Proportion of people who did match but shouldn't have

False Match Rate

Matching Threshold

Reducing this

Increases this

FALSE REJECT RATE

FALSE ACCEPT RATE
Identification (search) metrics

True Positive Identification Rate

How often (%) an individual ranks on a particular level in their search results

- TPIR 1 - 1st (95%)
- TPIR 2 - 2nd (97%)
- TPIR 3 - 3rd. (100%)
- etc.
Data validation and deduplication process

1. *Establish ground truths through well labeled data collection exercise
2. Biometric template matching
3. Clustering
4. Analysis and threshold setting
5. Outlier analyses with metadata
6. Record adjudication with biographics
7. Record reconciliation and deletion

*Ideal, but not required. Makes it possible to do #4 confidently.

Biometric duplication analysis
Case-study 1: Duplication volumes

Found 6% duplication, acceptable in your project?

- Year 2:
  - Enrolments: 200,000
  - Duplicates: 30,000

- Year 3:
  - Enrolments: 600,000
  - Duplicates: 36,000

- Year 4:
  - Successful ID: 85,795
  - Bypassed "flag": 24,785

85,795 prevented through real time deduplication
24,785 preventable but not selected (mandatory?)
12,107 identified through offline deduplication
Case-study 1: Duplication analysis

Many duplicates from few front line workers
Case-study 2: Duplication types

- Total enrollments: 869,296
- Total duplicates: 387,945
  - Self-enrollments: 61,538
- Systemic programme issues
- Work with management to understand why

As duplication rate increases so does self-enrolment rate
Ghana: Vaccine Delivery

Ethiopia: Health System strengthening

Malawi: HIV tracking

2.5 mil people enrolled
Verified data drives impact

Malawi: 62% increase in women linked to HIV care

Bangladesh: 39% increase in maternal health coverage

Ethiopia: 98% recorded patients received deworming pills, vs 69% in control districts
Implementation challenges

- Community acceptability
- Step-down training quality
- Operational vs. technical accuracy
- Supportive supervision
- Self-enrollments
- Duplicate flag by-pass
- Time-consuming record adjudication process and data reconciliation
Thank you!
Simprints face demo
Biometrics to improve immunisations

- Gavi-Simprints-NEC collaboration to develop under 5 fingerprint biometrics
- 5,000 children 0-5 in Bangladesh
  - 3 collections, 3 months apart
- Contact-based, commercialised scanner with 800-1000 ppi resolution
- Promising early results