



**simprints**  
every person counts

# Biometrics for patient identity management

Alexandra Grigore - CPO

# Duplicates

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



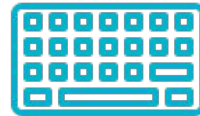
Palm/Vein



Gait



Signature



Keystroke



Voice

# How are biometrics used?

1. Client registration

Medical record

Biometrics + UID

+

Optional:  
Physical credentials

2a. Duplicate prevention

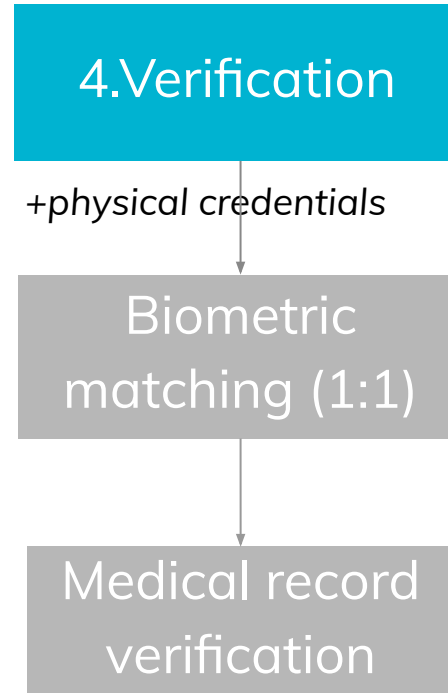
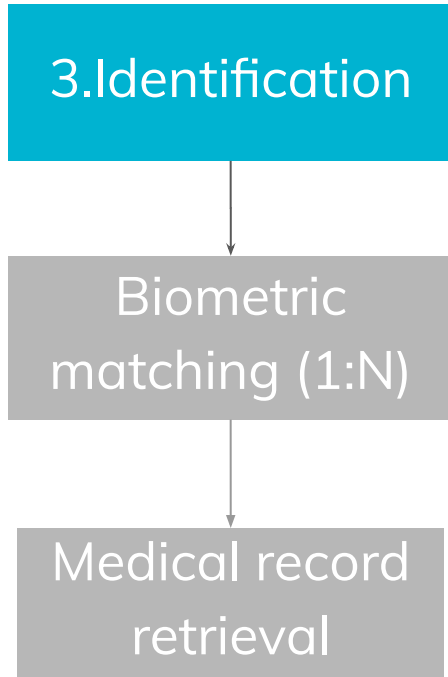
2b. Duplicate removal

Biometric matching (1:N or N:N)

Biographic matching (1:1)

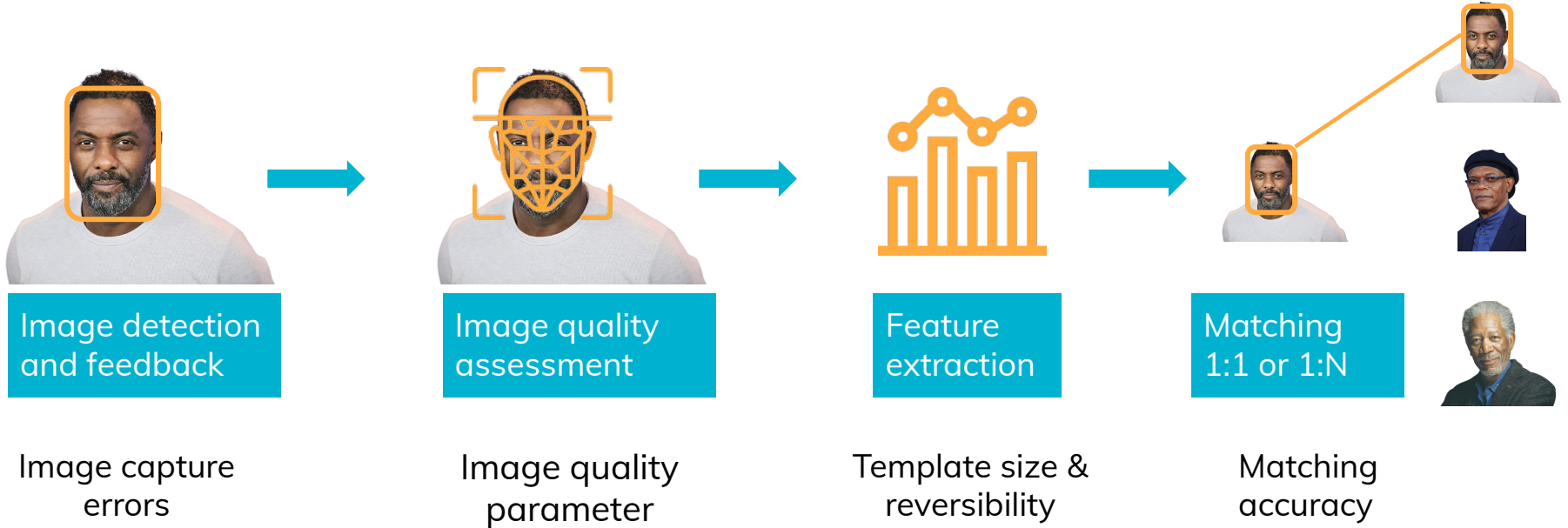
Duplicates removed

# Identification vs. verification

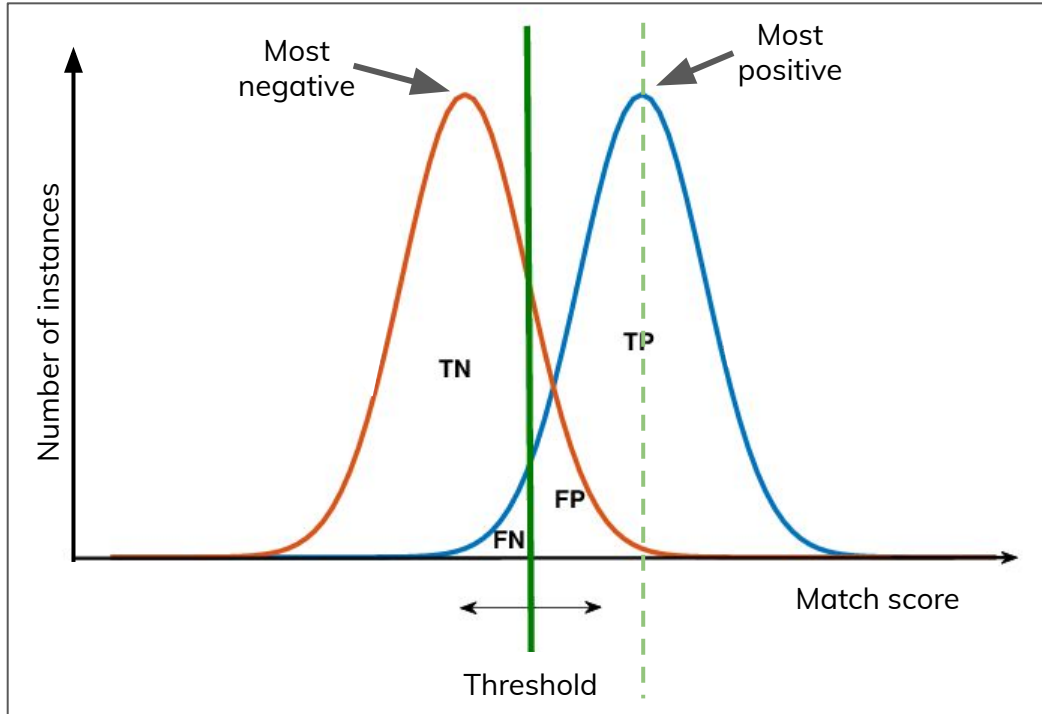




# Biometrics system components



# 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

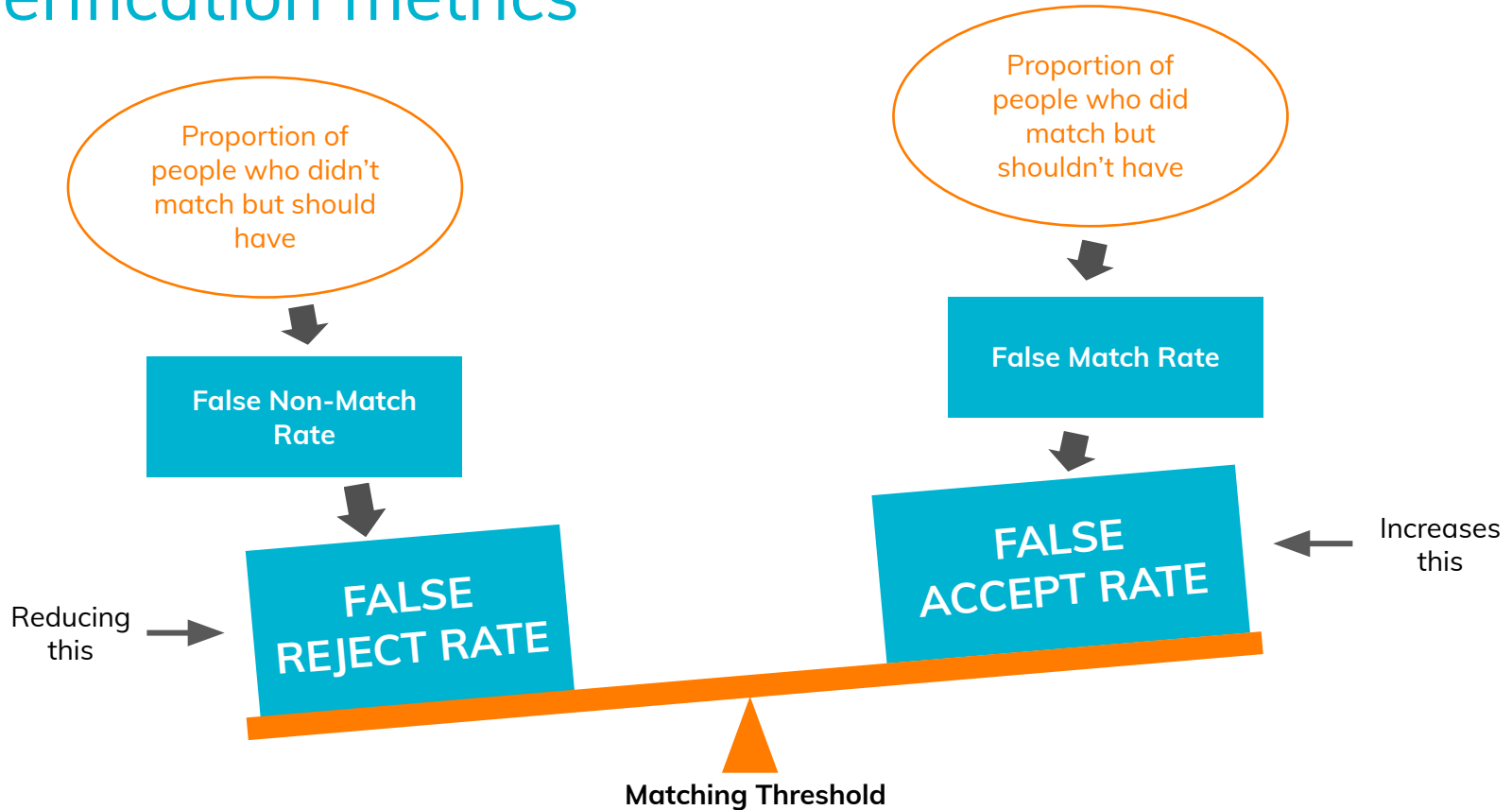


**FTA:** Failure to Acquire

**FRR:** False Reject Rate

**TPIR:** True Positive Identification Rate

# Verification metrics

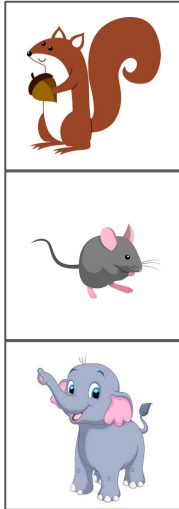


# Identification (search) metrics

Who am I?



Algorithm  
A



Rank 1

Rank 2

Rank 3

Algorithm  
B



## True Positive Identification Rate

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

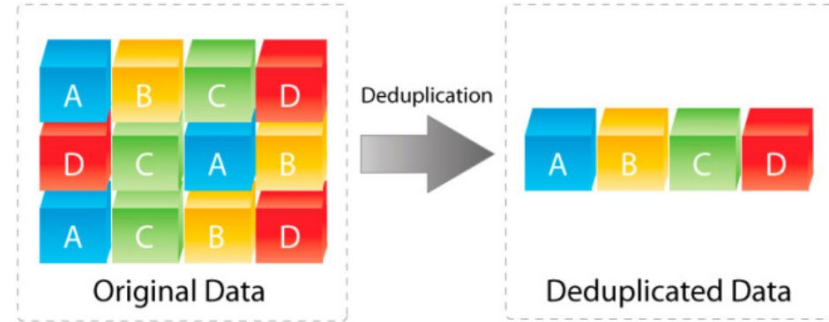
- TPIR 1 - 1st (95%)
- TPIR 2 - 2nd (97%)
- TPIR 3 - 3rd. (100%)
- etc.

# Data validation and deduplication process

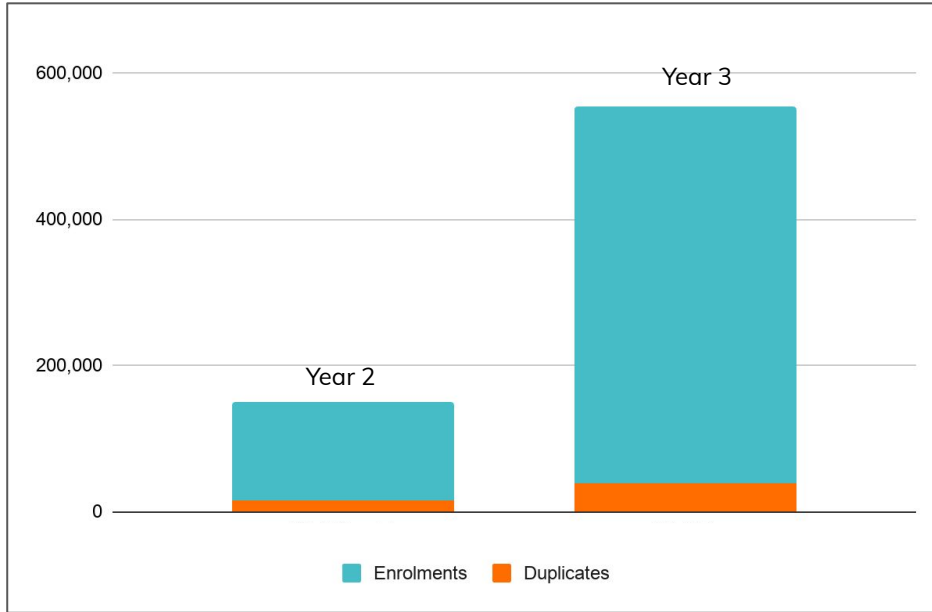
1. \*Establish ground truths through well labeled data collection exercise
2. Biometric template matching
3. Clustering
4. Analysis and threshold setting
5. Outlier analyses with metadata
6. Record adjudication with biographics
7. Record reconciliation and deletion

**Biometric** duplication analysis

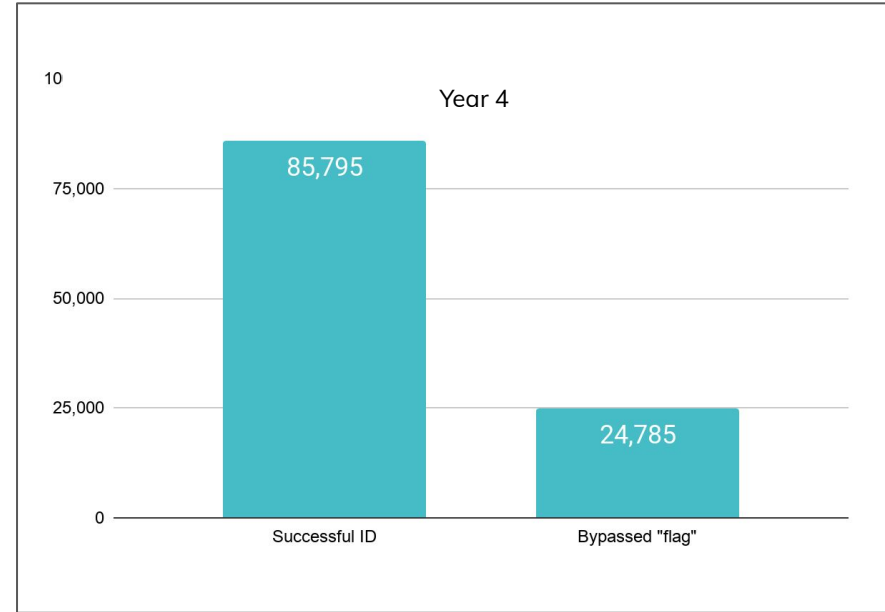
\*Ideal, but not required. Makes it possible to do #4 confidently.



# Case-study 1 : Duplication volumes



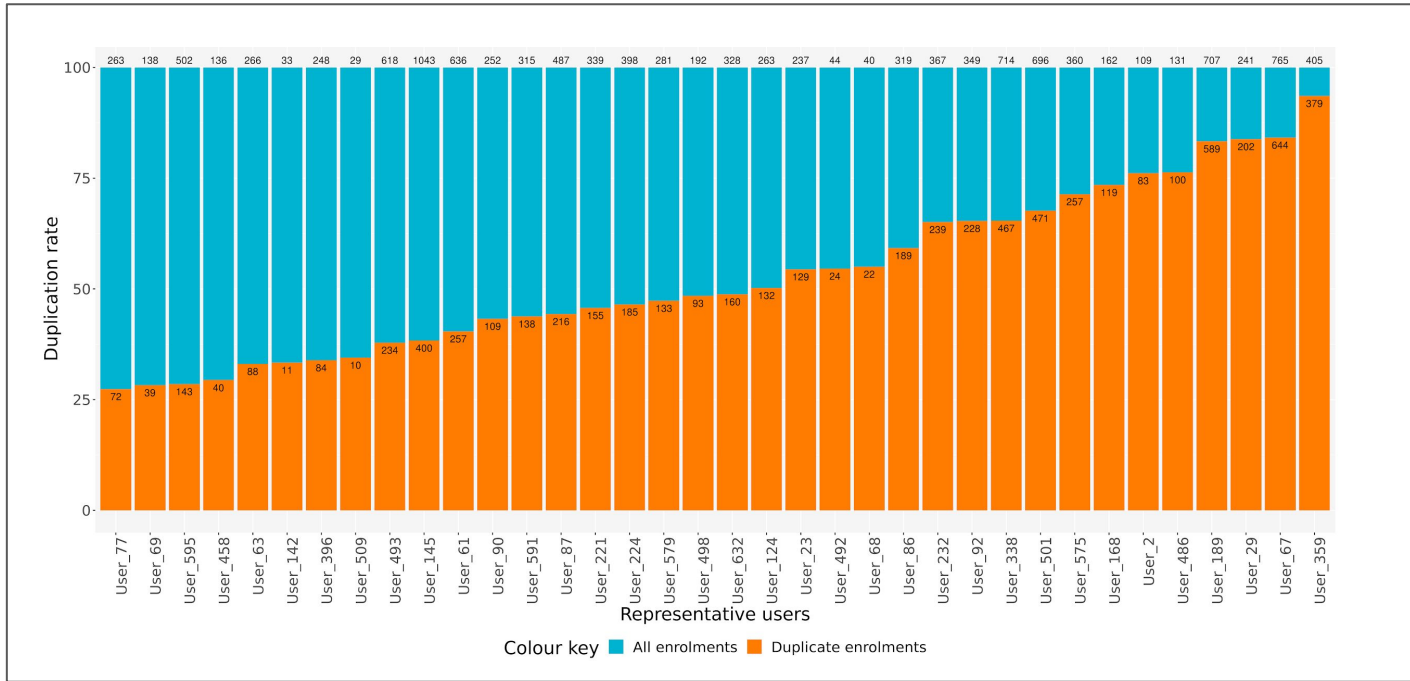
Found 6% duplication, acceptable in your project?



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

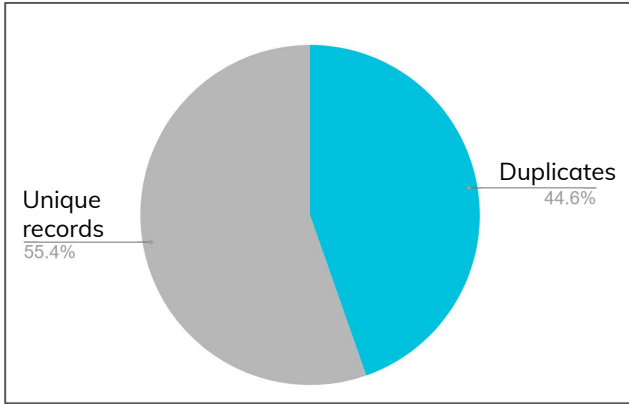


Many duplicates from few front line workers

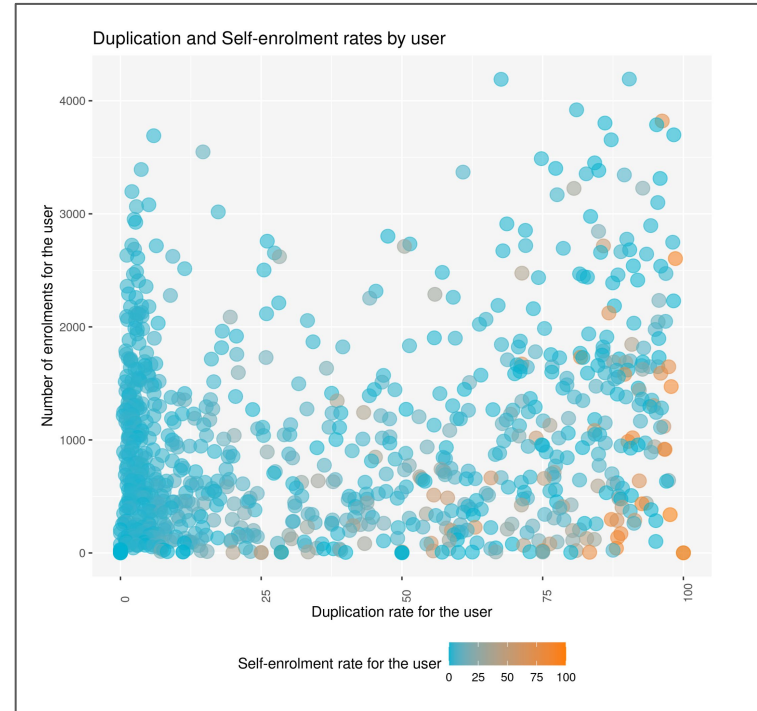




# Case-study 2 : Duplication types



- 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





MINISTRY OF HEALTH  
REPUBLIC OF GHANA



Ethiopia  
Ministry of Health



BROWN

**Ghana:**  
Vaccine Delivery

**Ethiopia:**  
Health System strengthening

**Malawi:**  
HIV tracking



**2.5 mil people**  
enrolled

# Verified data drives impact



BROWN

**Malawi:** 62% increase in women linked to HIV care



Foreign, Commonwealth  
& Development Office



GLOBAL  
INNOVATION  
FUND

**Bangladesh:** 39% increase in maternal health coverage

Imperial College  
London

**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



