



Who uses health services? Record linkage between health facility and demographic surveillance data in rural South Africa.

ChodziwadziwaKabudula, F. Xavier Gómez-Olivé

Georges Reniers, Benjamin Clark

Samuel clark, Stephen Tollman



INDEPTH Network



Background

- Health and Demographic Surveillance Systems (HDSS) have progressed enormously in documenting the burden of diseases in communities with weak information systems.
- Several types of socioeconomic data are used as explanatory variables for observed differentials.
- HDSS, however, have provided limited information on utilisation of health services where they have existed in isolation from health facility data.



INDEPTH Network



Question

- How feasible it is to produce representative datasets on patterns of health care utilisation in communities with weak registration systems in Primary Health Care by linking health facility and HDSS data?



INDEPTH Network



Overview of record linkage

- Record linkage between two data sources **A** and **B** attempts to identify records that belong to the same individual
- Two types of approaches:
 - ✓ Deterministic → Record pairs linked based on agreement on specific identifiers or unique identifier.
 - ✓ Probabilistic → Record pairs linked based on probabilities of agreement on a set of identifiers.



INDEPTH Network



Use of record linkage for population-based research

- Common in developed countries (both approaches)
- Rare in developing countries.
 - Lack of unique identification system e.g social security number
 - Inaccurate reporting of identifying information due to high illiteracy



INDEPTH Network



Objectives

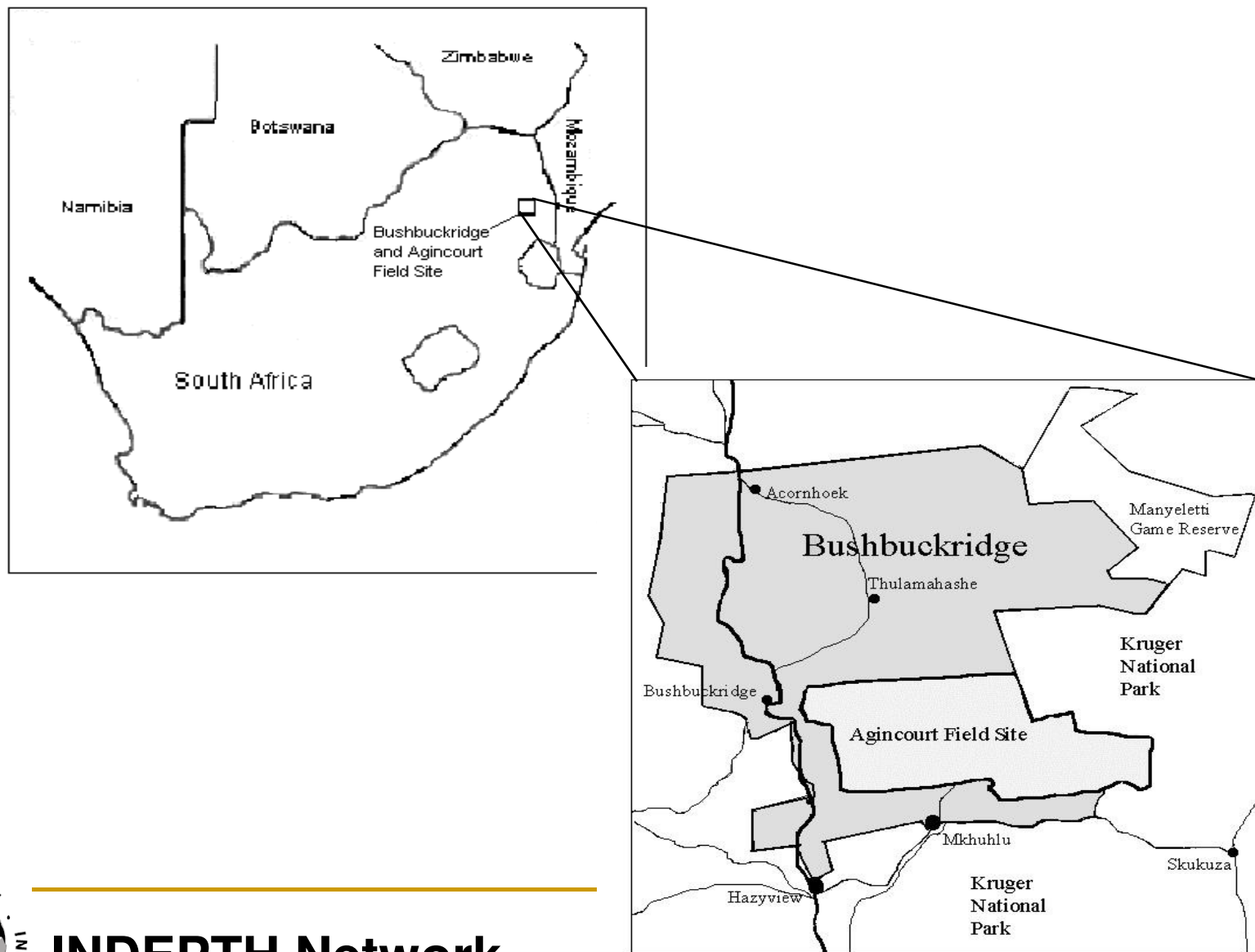
- ❑ To assess the representativeness of datasets generated by record linkage between health facility data and health demographic surveillance system in rural northeast South Africa.



INDEPTH Network



Study area: Agincourt HDSS



INDEPTH Network



Methods

- Construction of gold standard dataset

Fingerprints collection from males and females aged 18+ years from 2 villages in AHDSS

Fingerprints collection from all patients at Agincourt Health centre

Individuals linked by matching fingerprints



Methods (Cont'd...)

■ Deterministic record linkage

All individuals in AHDSS database

Individuals in gold standard dataset from Agincourt Health Centre

Individuals linked by matching identifiers

Mobile phone number
&JaroWinkler(First names) ≥ 0.8
&Gender

OR

Mobile phone number
&DOB
&Gender

JaroWinkler(First Name) ≥ 0.8
&JaroWinkler(Surname) ≥ 0.8
&Gender
& Age difference ≤ 1
&Village
& Clinic visit date before
end of residence episode in
AHDSS



INDEPTH Network



Methods (Cont'd...)

- Comparison of fingerprint matched versus identifier matched:
 - ✓ Gender distribution
 - ✓ Age distribution
 - ✓ Ethnicity
 - ✓ Education status
 - ✓ Union status



INDEPTH Network



Results: Fingerprint linkage

Fingerprints collected in AHDSS villages

- ✓ 2895 Individuals
- ✓ 1930 female
- ✓ 965 male

Fingerprints collected at Agincourt Health centre

- ✓ 4190 individuals
- ✓ 2734 female
- ✓ 1456 male

Linked by matching fingerprints

- ✓ 623 individuals
- ✓ 511 female
- ✓ 112 male

Results: Deterministic record linkage

Identifiers	Finger prints		Total
	Match	Non-match	
Match	399	19	418
Non-match	224	192653	192877
Total	623	192672	193295

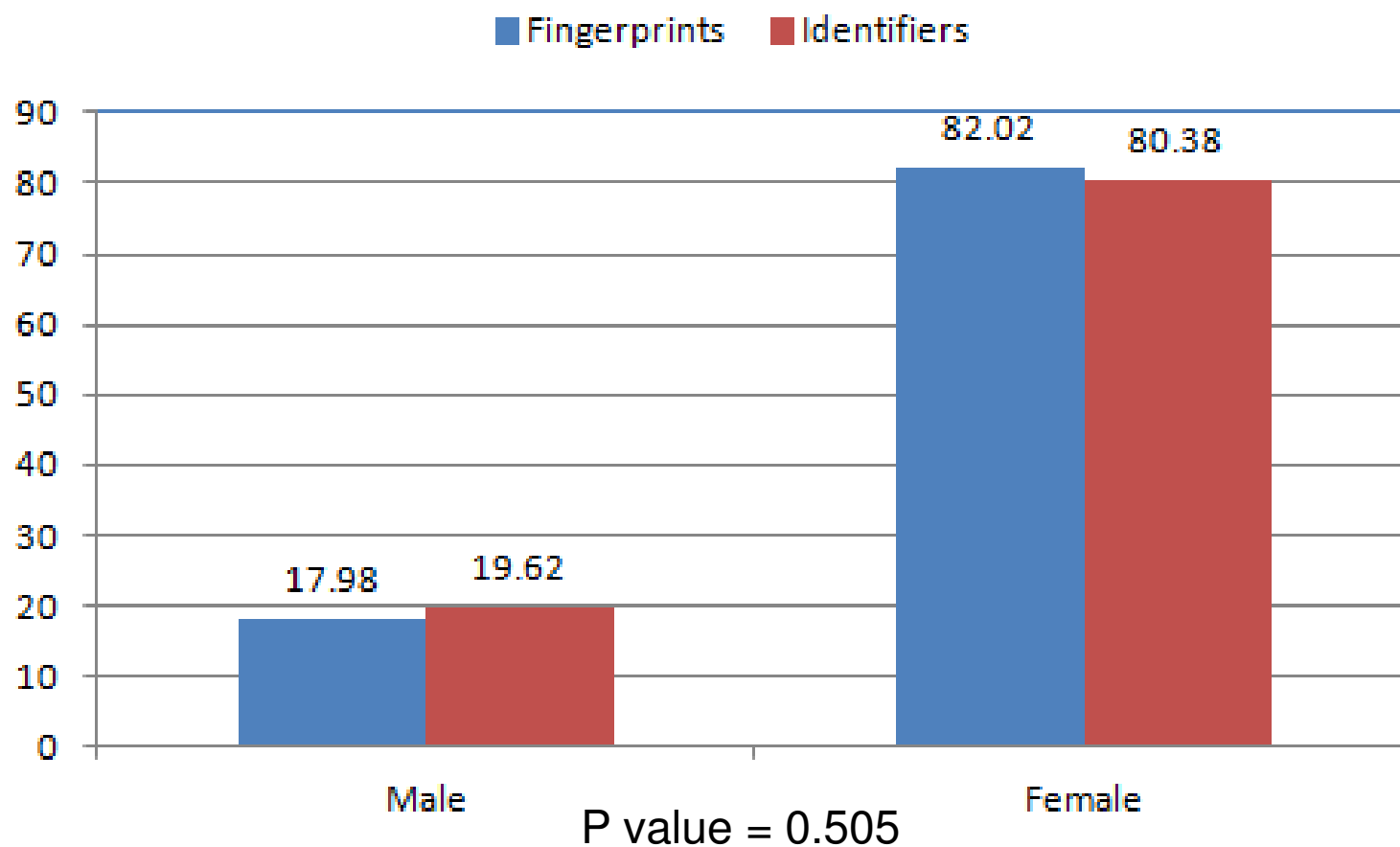
Sensitivity = 64.04%

Specificity = 99.99%

Positive Predictive Value = 95.45%

INDEPTH Network

Results: Fingerprint vs Identifier linked by gender



INDEPTH Network

Results: Fingerprint vs Identifier linked

	Females (Chi sq p-values)	Males (Chi sq p-values)
Age distribution	0.098	0.692
Ethnic group	0.59	0.448
Marital status	0.23	0.982
Education status	0.057	0.846



Conclusion

- Results suggest that record linkage between health facilities and AHDSS data can yield representative information on the characteristics of users of health services despite low linkage sensitivity.
- Further validation need on a larger dataset with known match status



INDEPTH Network



Thank You!



INDEPTH Network

