

An error framework for longitudinal administrative sources; its use for understanding the statistical properties of data on international migration

**UNECE Migration Statistics Work Session
Geneva 2019**

Dr Louisa Blackwell and Nicky Rogers
Office for National Statistics United Kingdom


30 October 2019

1

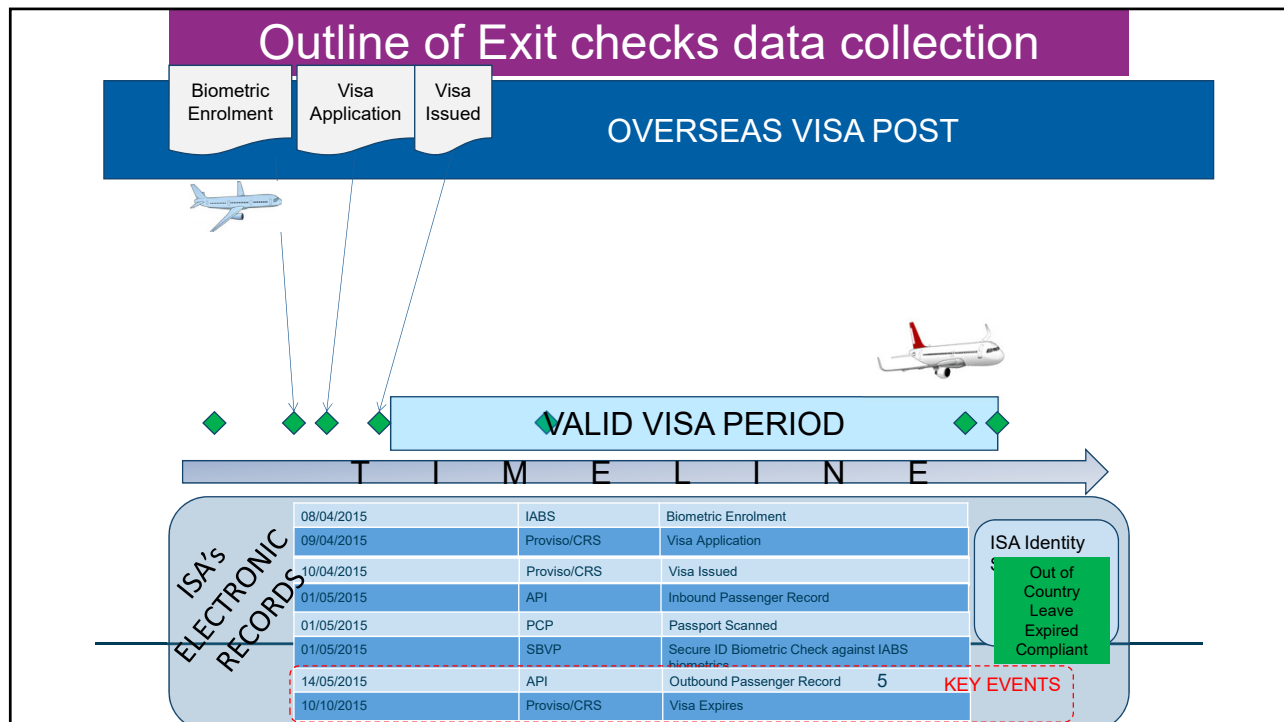
Presentation outline

- Context: ONS UK Population & Migration Statistics Transformation Programme
- International migration and the use of Home Office administrative data
- An error framework for longitudinally linked administrative data
- Quality indicators

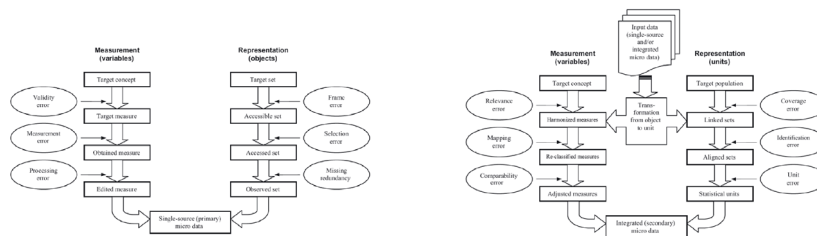
**We are transforming the way we
produce population and
migration statistics,
to better meet the needs of our
users**



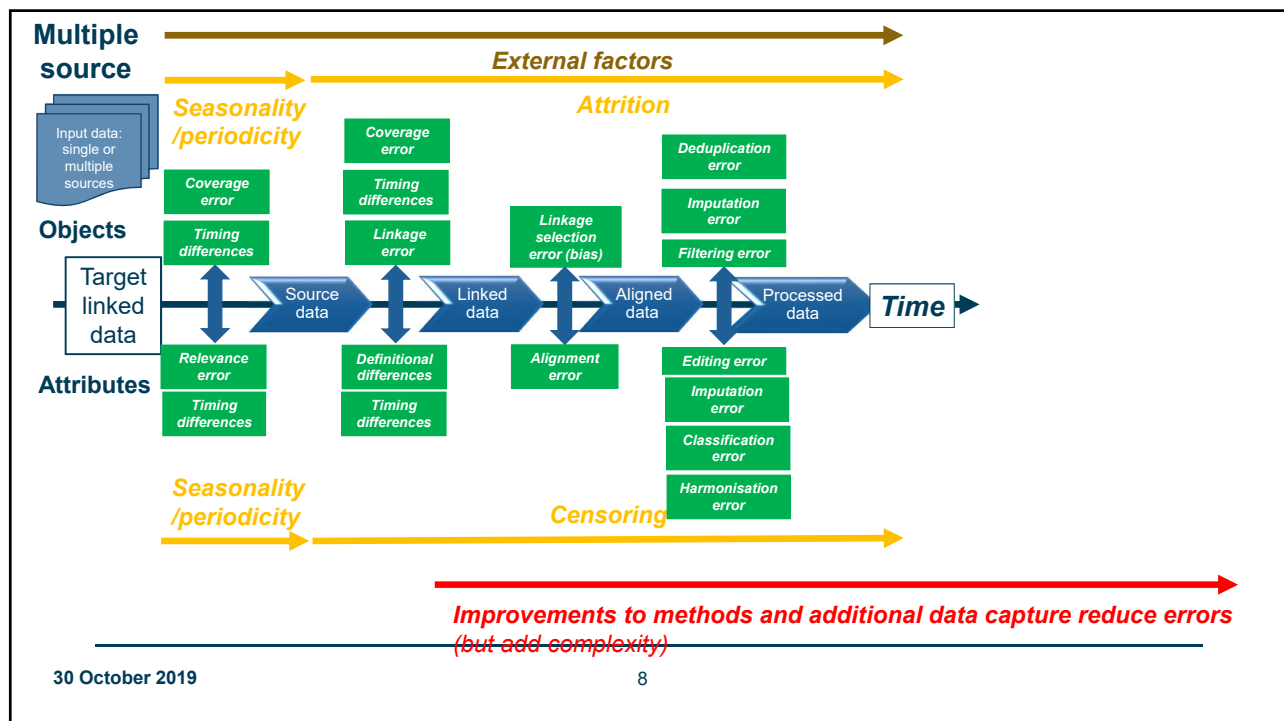
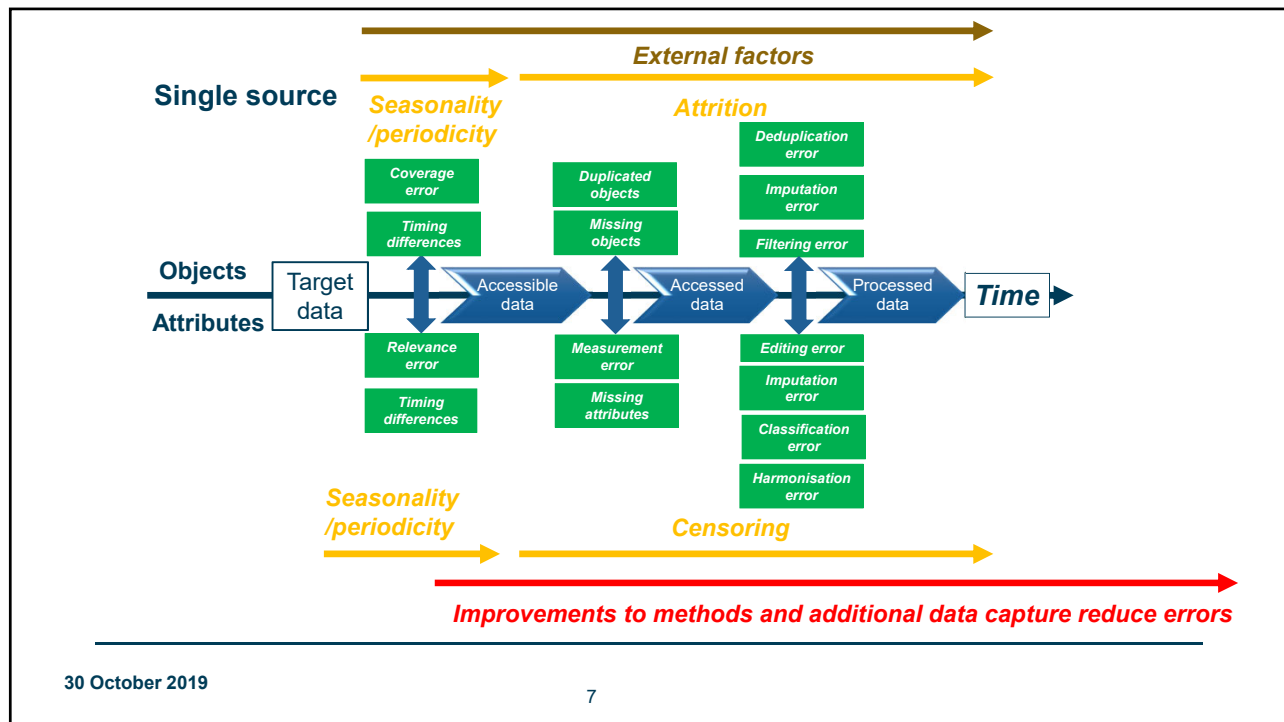
**Our mission is to make
the best possible
population statistics using
all available data sources**



SNZ quality framework: Phases 1 and 2



- Extremely valuable taxonomy of error to help us to understand quality issues
- Taken it a step further to explicitly acknowledge error arising from the longitudinal dimension



Exit checks – Common Travel Area (CTA)

Travel within CTA is **out of scope**

- i.e. between UK and Ireland
- between UK and the Channel Islands
- not in scope of exit checks programme

Someone flying into UK then departing via Ireland would likely be a 'not initially identified departure'

Home Office has built **operational feedback loops** to fill in missing data

- eg from Beijing for Chinese tour groups travelling via CTA



30 October 2019

9

9

Exit checks – measuring & improving quality

- **Coverage:** the extent to which systems are in place to collect travel data for all routes Latest quarter = **100% outbound** , **89% inbound** (to 30 June 2019)
- Inbound supplemented by information from passport scans, to help fill the gap
- **Voyage level data receipt:** Aviation carriers supply Advance Passenger Information for voyages= **99.5% inbound**, **99.0%** outbound (June 2019)
- Further work included **improving data collection at source** (to address data entry errors due to User Interface (UI) design (drop down menus used by passengers) and discussions with some carriers on improving quality)
- *As individuals travel more often, data systems are more likely to be able to identify them correctly, so **data about individuals and their journeys improves over time***

30 October 2019

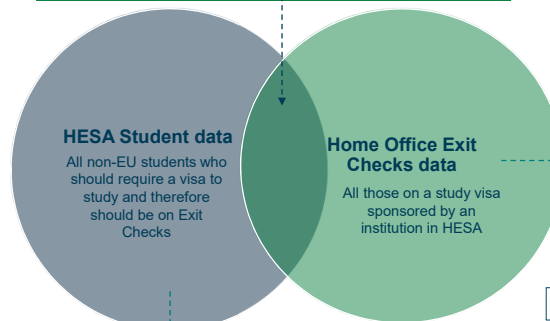
10

10

Linking HESA and Exit Checks data

Linked data – HESA graduating non-EU students linked to Exit Checks data.

We are still developing our data linkage methods but we wanted to show early progress. This analysis is based on 1-1 matches, therefore we will be missing potential links.



Unlinked Exit Checks records

Exit checks records that could not be linked to HESA using 1-1 matches.

Unlinked HESA students

HESA records that could not be linked to Exit Checks using 1-1 matches

HESA & Exit Checks match rates

Match rate	2015/16 (%)	2016/17 (%)
HESA records linked to Exit Checks data	80	74

30 October 2019

11

Understanding linked and unlinked records: Age and sex distributions



Source: ONS analysis of linked HESA and Exit Checks data and unlinked HESA data

30 October 2019

12

Next steps



Ambitious programme of secondments of data scientists, analysts and methodologists from ONS to UK Home Office

Prototyping new ways of rapidly transferring knowledge, developing resources and building data assets

International peer review and feedback and publication in a book and statistical journals

30 October 2019

13

Acknowledgements



With special thanks to contributing authors:

Sarah Cummins and Elzemies Scott-Kortlever, Office for National Statistics

and Professor Peter W Smith, Southampton University

Also special thanks to contributing teams at the UK Home Office including statisticians, operational researchers, immigration system extracts and data suppliers

30 October 2019

14

Any questions?

