

Harnessing data from health and other public services

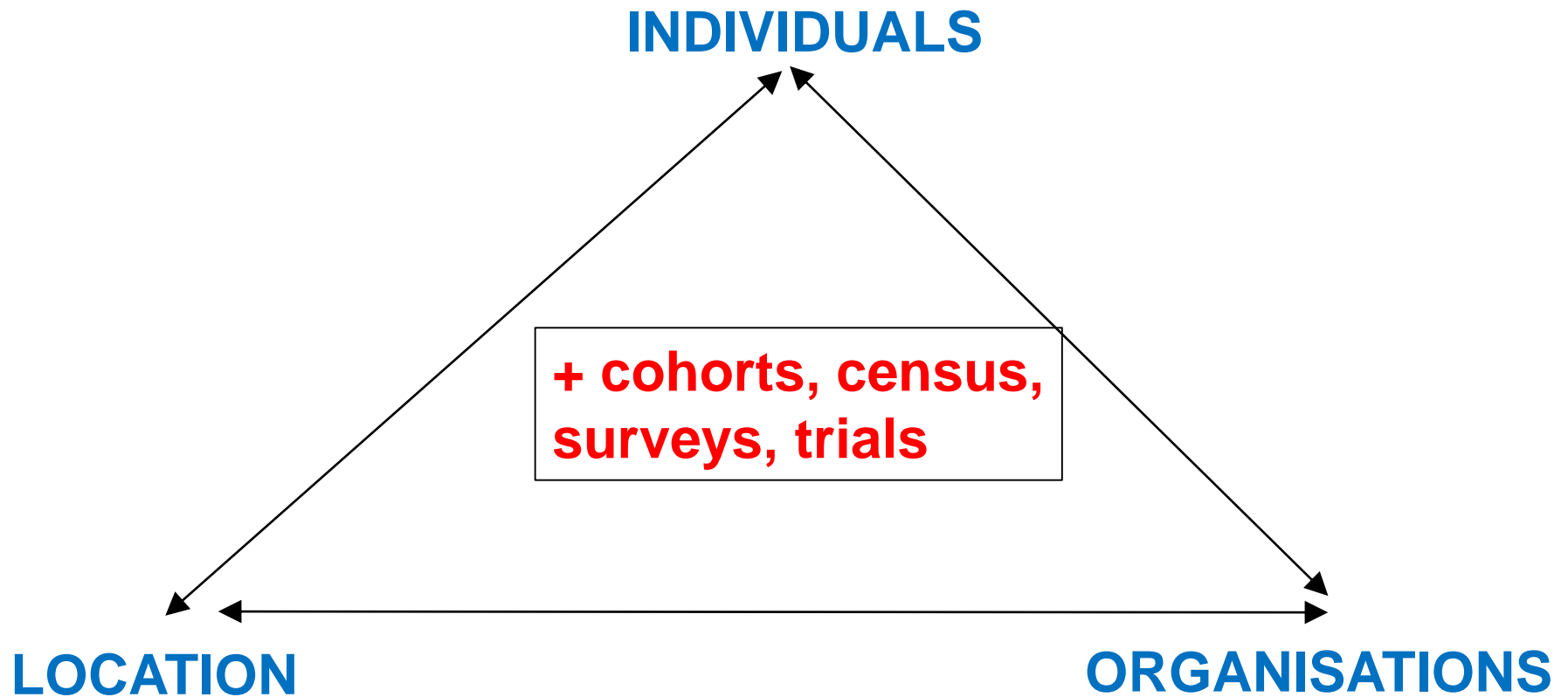
Ruth Gilbert

Professor of Clinical Epidemiology
University College London Great Ormond Street Institute of Child Health



Administrative Data
Research Centre
England

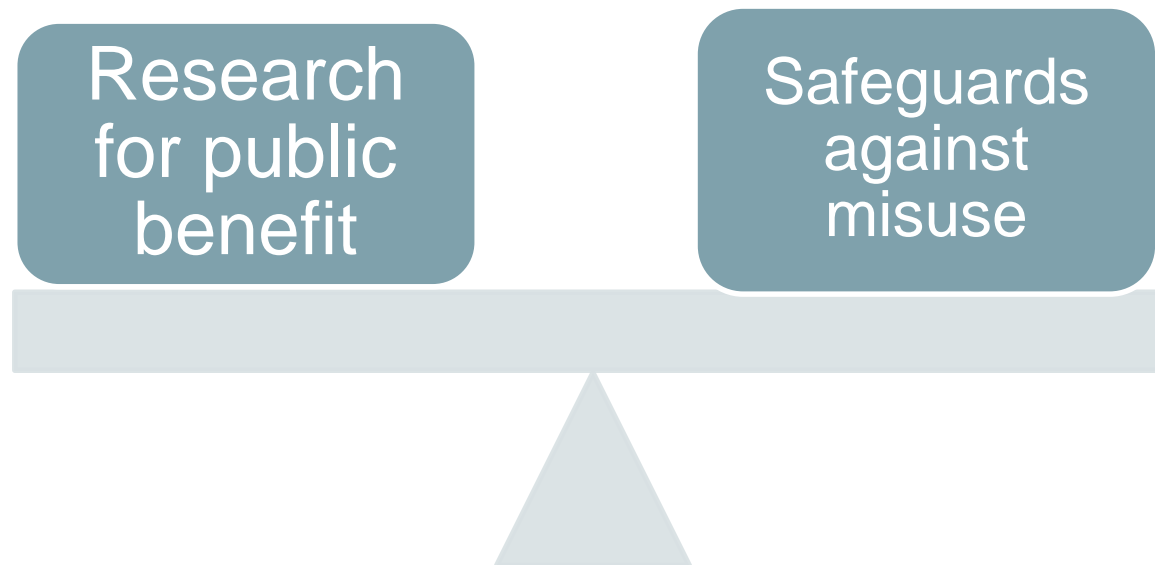
Administrative data



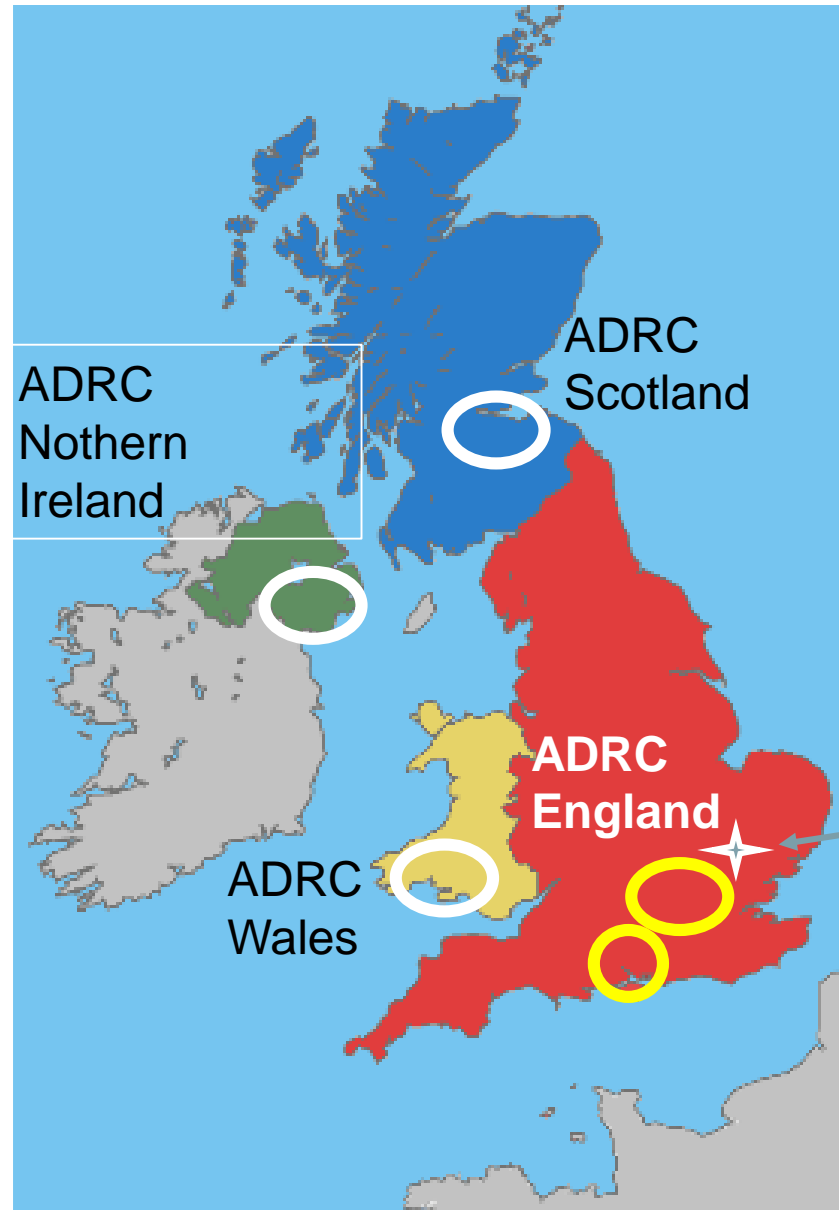
Use of data for research requires a balance of two public goods

Administrative Data Research Network

....'to widen the use of linked administrative data for research'



Administrative Data Research Network



ADRC
Nothern
Ireland

ADRC
Scotland

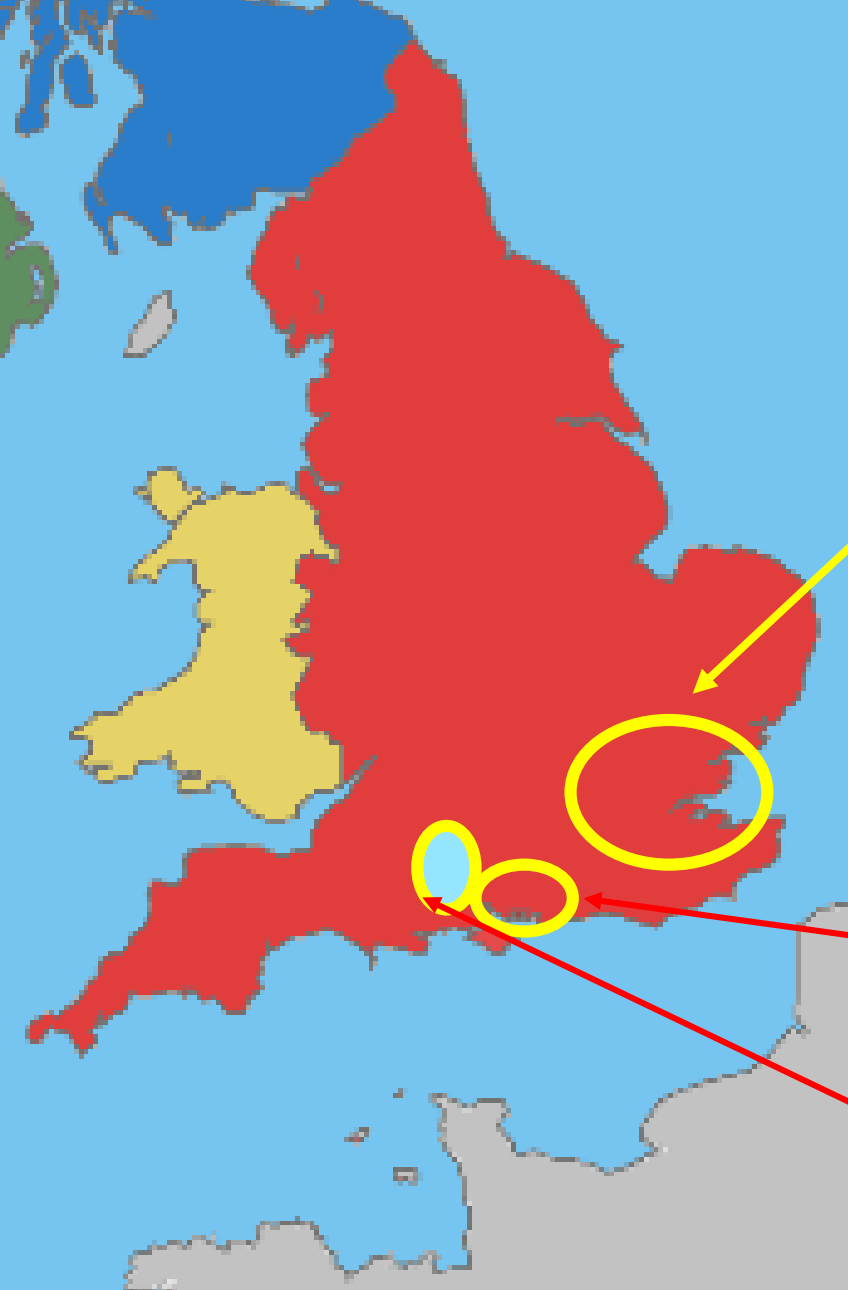
ADRC
Wales

ADRC
England

Administrative
Data Service
Essex Univ



Administrative Data
Research Centre
England



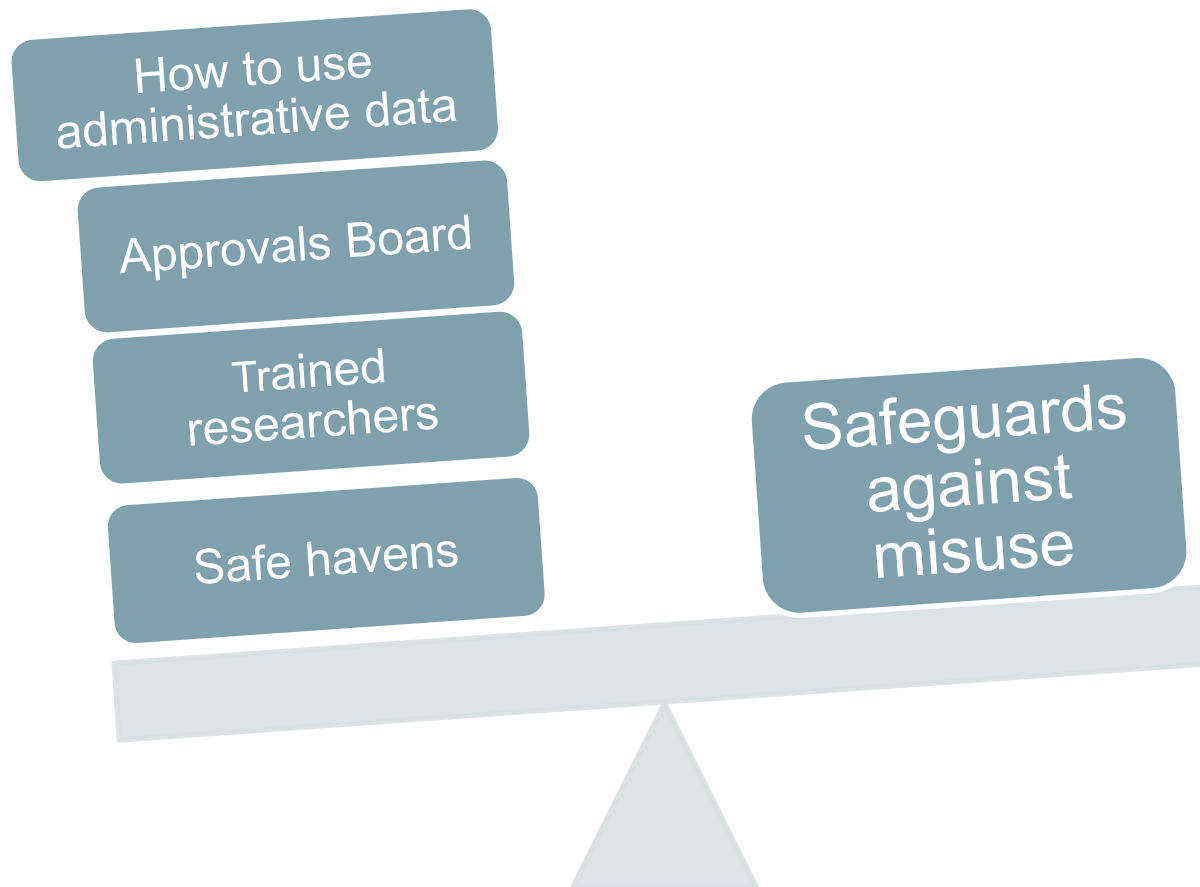
Bloomsbury

- UCL
- LSHTM
- Institute Fiscal Studies

University of
Southampton

Office of National Statistics

Administrative Data Research Network...how does it help?



'Big' health data-some examples (in 2015)

ONS vital registration data (England & Wales)

- 700,000 births
- 500,000 deaths
- 3300 stillbirths
- 185,000 abortions (DH)



Hospital Episode Statistics (England)

- 15 million inpatient admissions
- 18 million A&E attendances
- 108 million outpatient bookings



Children Looked After data (England)

- 69,500 children in care



CPRD/THIN (UK)

- Complete primary care records
- 6-8% of the UK population
- 4-6 million active patients



Paediatric Intensive Care Audit Network (UK)

- 59,500 PICU admissions



Data linkage: need to evaluate within data provider/linker environment

Work with data linkers

- NHS Digital
- Public Health England
- Ministry of Justice

Hagger-Johnson et al. **Probabilistic linking to enhance deterministic algorithms and reduce linkage errors in hospital administrative data.** J Innovation in Health Informatics (in press)

BMJ Open Data linkage errors in hospital administrative data when applying a pseudonymisation algorithm to paediatric intensive care records

Gareth Hagger-Johnson,^{1,2} Katie Harron,¹ Tom Fleming,³ Ruth Gilbert,¹ Harvey Goldstein,^{1,4} Rebecca Landy,⁵ Roger C Parslow³

HSR Health Services Research

Identifying Possible False Matches in Anonymized Hospital Administrative Data without Patient Identifiers

Gareth Hagger-Johnson, Katie Harron, Arturo Gonzalez-Izquierdo, Mario Cortina-Borja, Nirupa Dattani, Berit Muller-Pebody, Roger Parslow, Ruth Gilbert, and Harvey Goldstein

Understanding the Impact of Data Quality on Data Linkage

Version number: v4.0

First published: June 2016 (Final)

NHS
England

Research Article

Statistics
in Medicine

Journal of Public Health | pp. 1–8 | doi:10.1093/pubmed/fdx037

Received: 4 August 2016. Accepted: 2 March 2017. Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.7287

A scaling approach to record linkage

Harvey Goldstein,^{a,c,*†} Katie Harron^b and Mario Cortina-Borja^c

Perspectives

GUILD: GUIDance for Information about Linking Data sets[†]

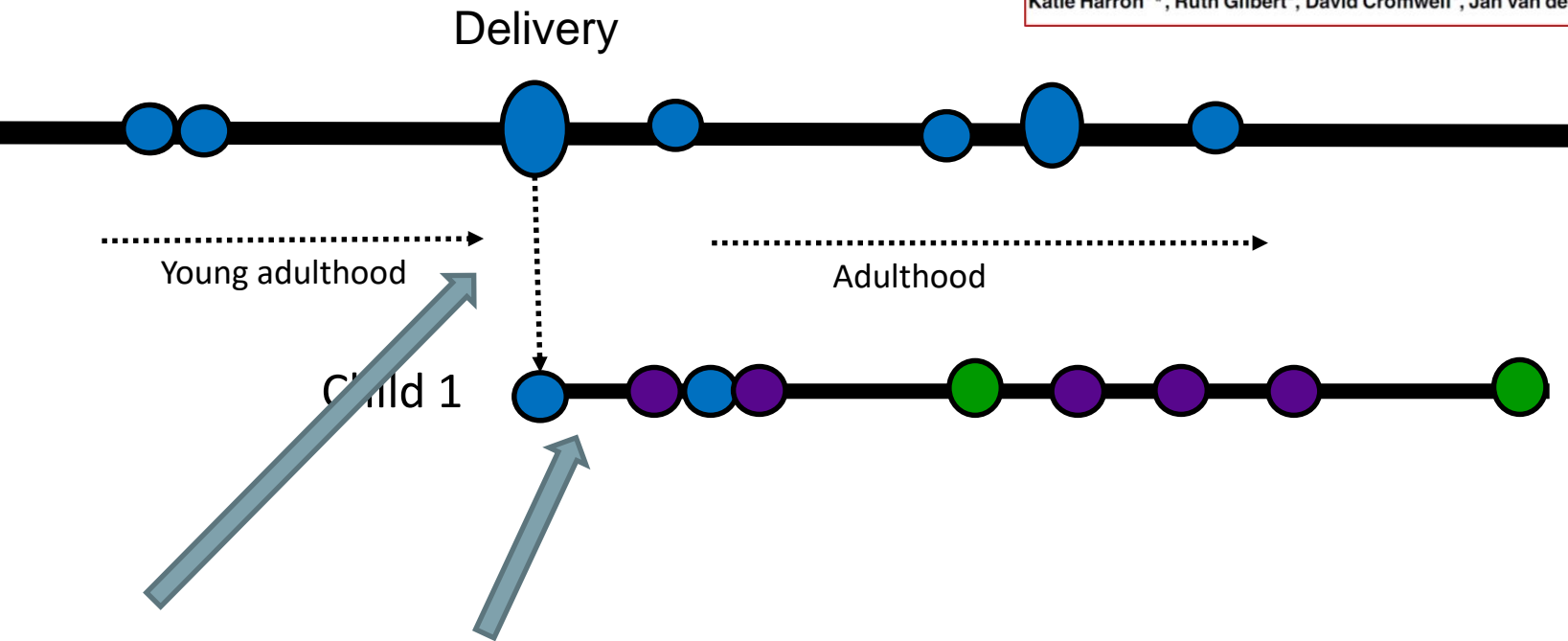
RESEARCH ARTICLE

Linking Data for Mothers and Babies in De-Identified Electronic Health Data

Katie Harron^{1*}, Ruth Gilbert², David Cromwell¹, Jan van der Meulen¹

2016

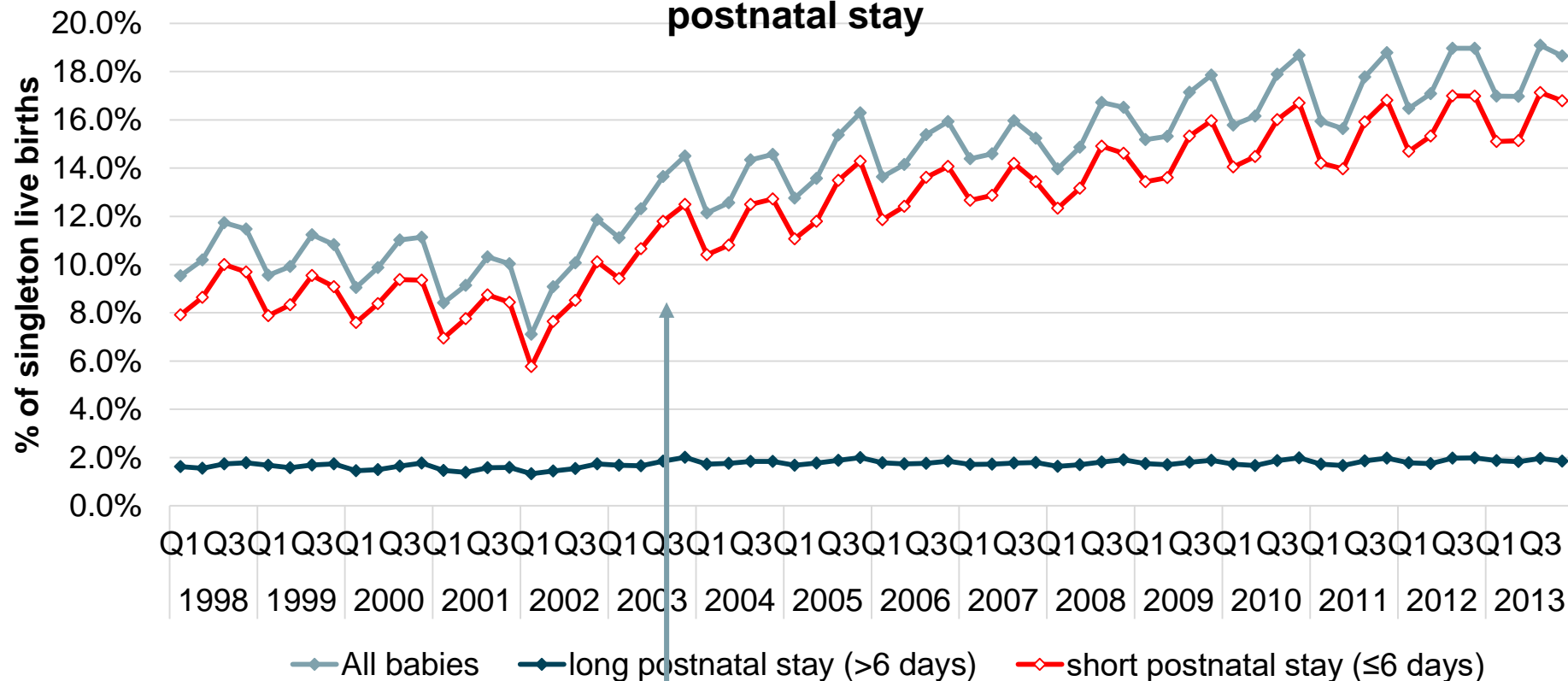
Mother-baby linkage in England – 7 million pairs



Linkage accuracy?

Validating data : Hospital episode statistics for England

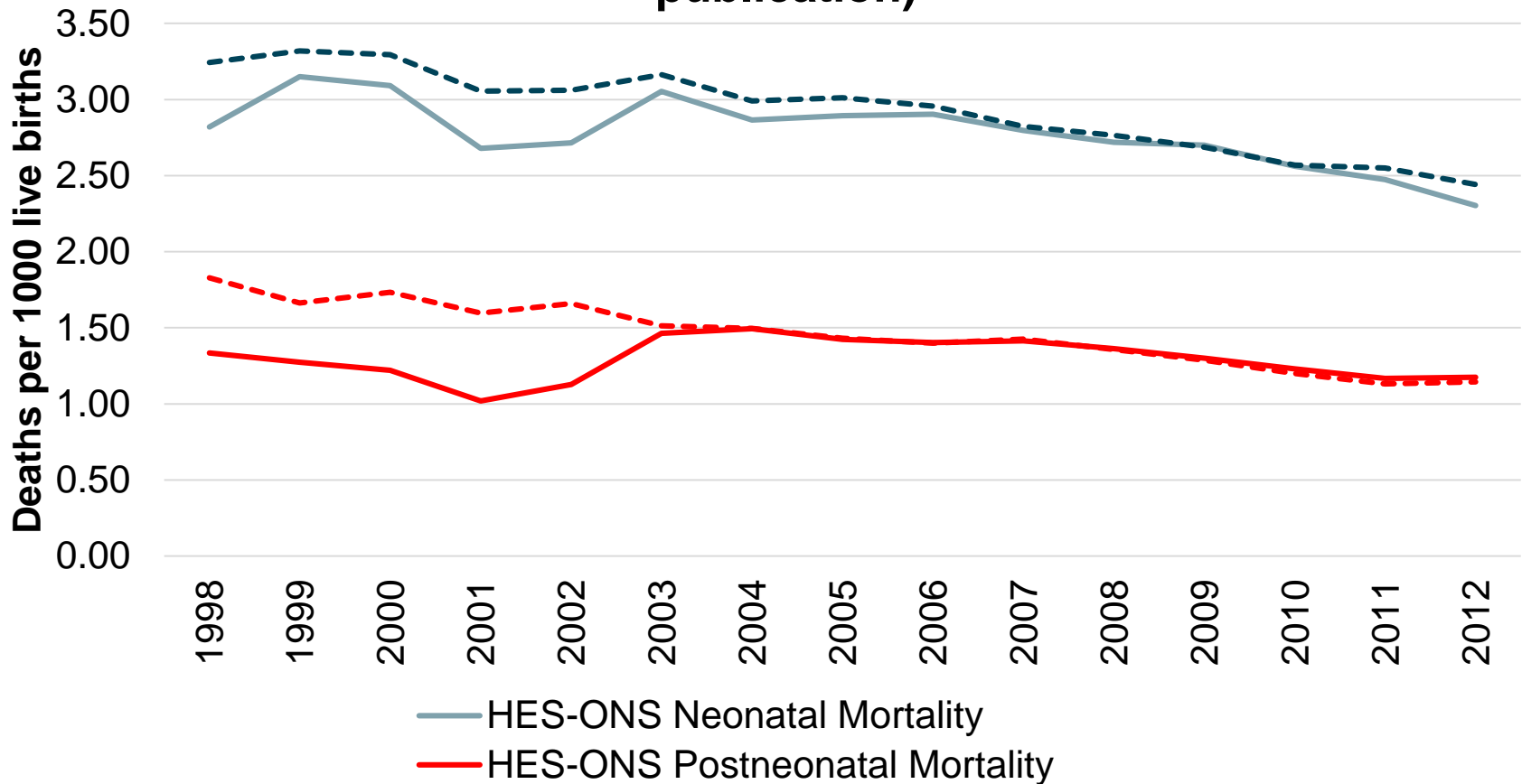
% of children with at least one hospital readmission by length of postnatal stay



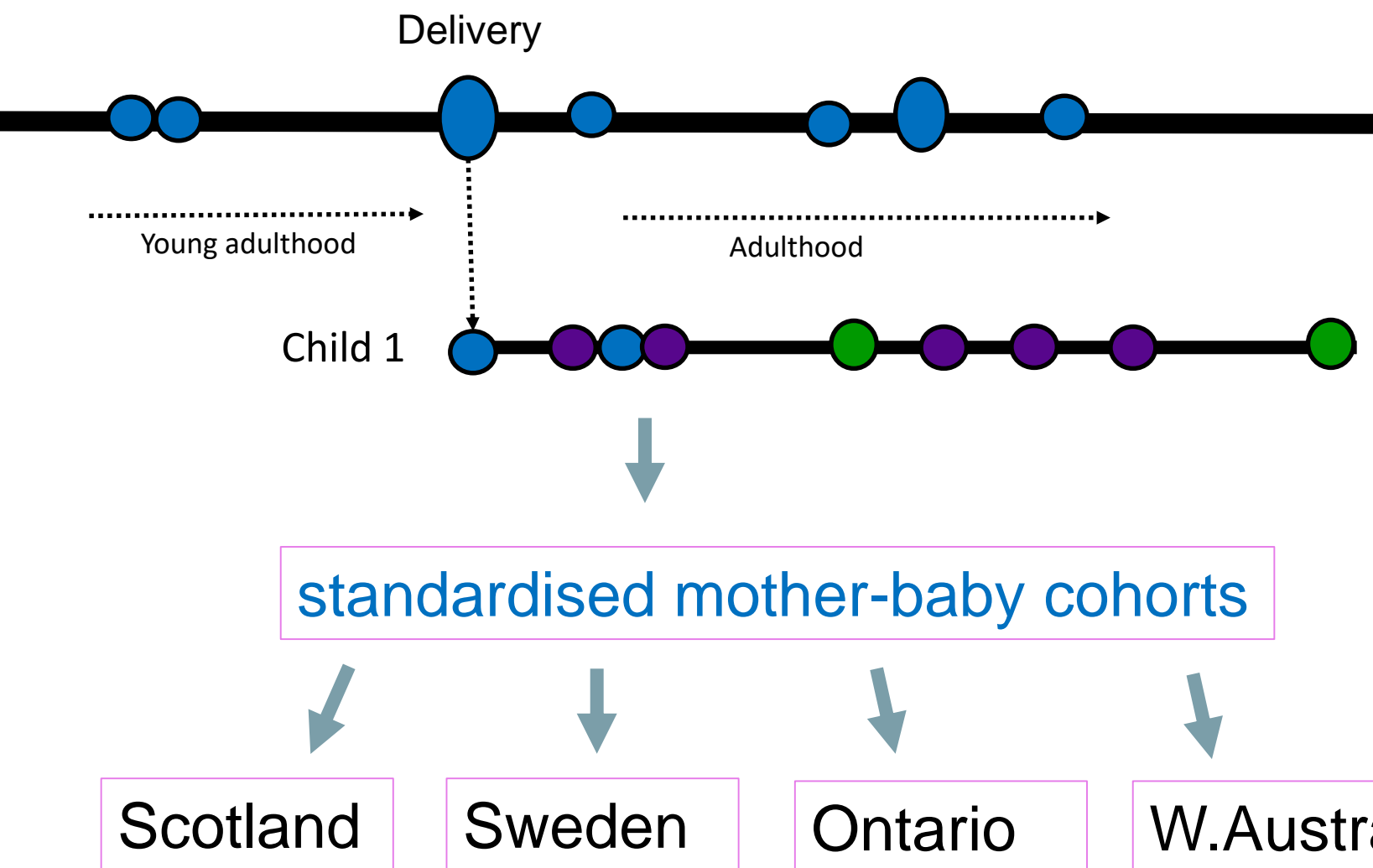
Apparent step change in rates of emergency readmissions due to improved linkage accuracy between birth episodes and subsequent admissions in HES due to availability of NHS number

Validating data : Hospital episode statistics for England

Neonatal (1-27 days) and post-neonatal (28-364 days) mortality in HES-ONS compared to gold standard (ONS publication)



Mother-baby linkage in England – 7 million pairs



International comparison of emergency hospital use for infants: data linkage cohort study in Canada and England

Harron K, et al. *BMJ Qual Saf* 2017;**0**:1–9. doi:10.1136/bmjqs-2016-006253

Katie Harron,¹ Ruth Gilbert,² David Cromwell,¹ Sam Oddie,³ Astrid Guttman,⁴ Jan van der Meulen^{5,6}

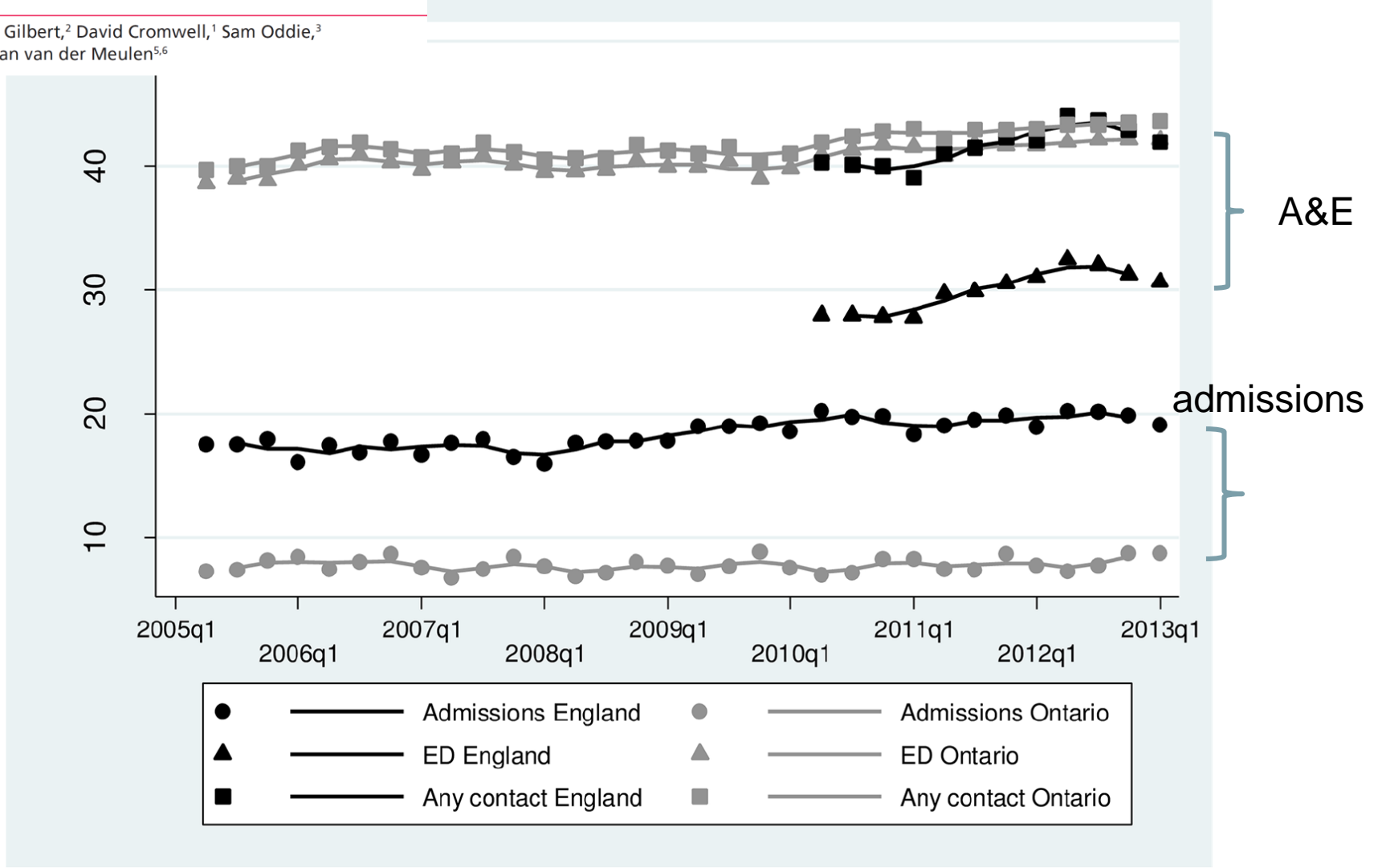
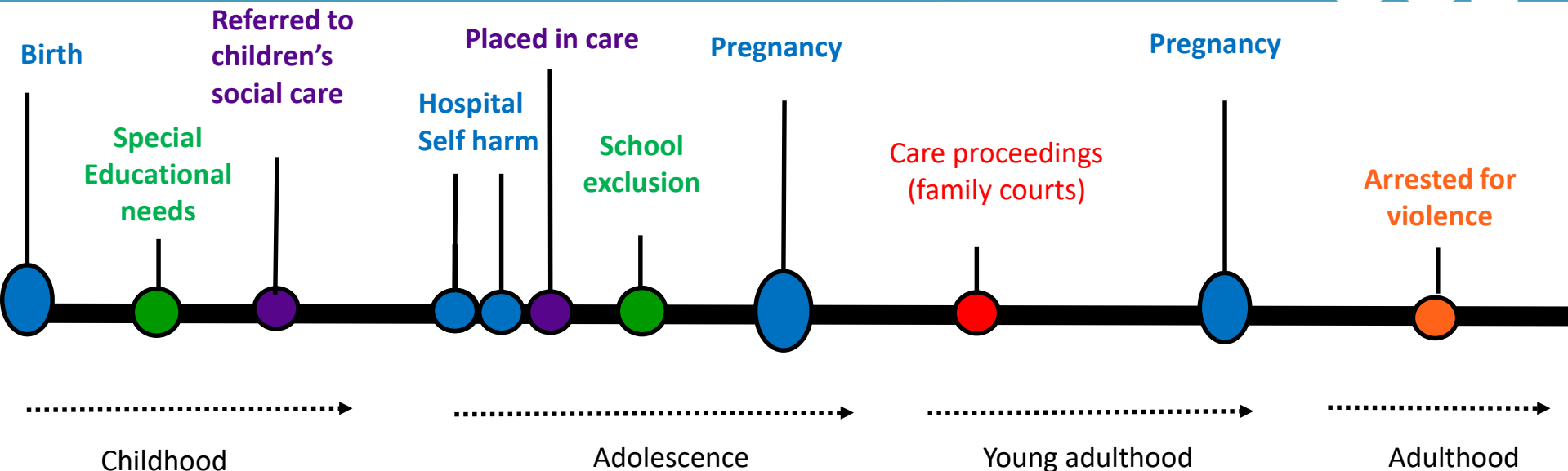


Figure 1 Trends in the percentage of infants with ≥ 1 unplanned admission, ED visit or any contact (admission or ED) within 12 months of postnatal discharge. Symbols=observed rates; lines=three-quarter moving average. ED, emergency department.

What data can we draw on?



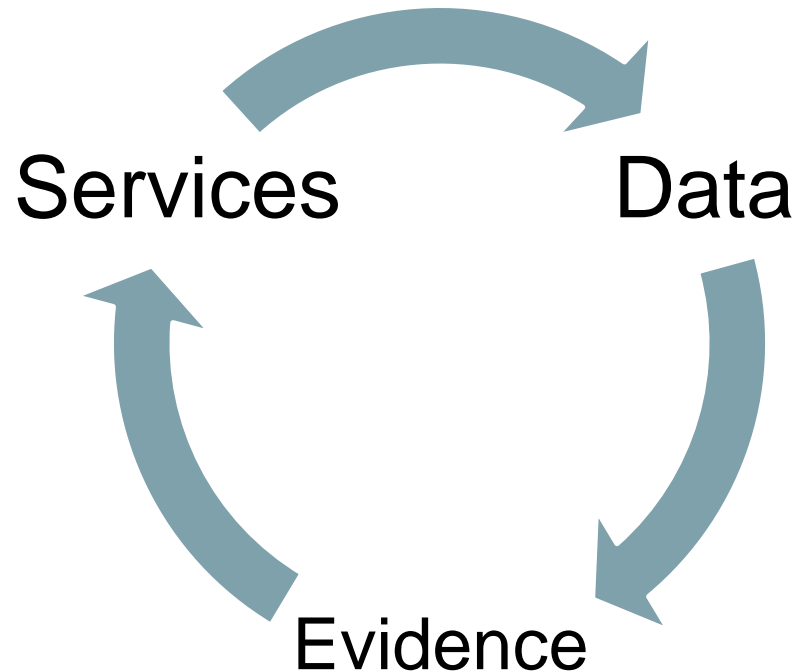
Health

Education

Social care

Family justice

Crime, employment.....



ADRC England exemplar studies in development

- Health-education-pollution indicators
- hospitalisation-birth registration-census mother-baby cohorts
- renal transplantation ~ welfare benefits/earnings
- dormant trials ~ education and health
- household survey (labour force survey) ~ health for family members

England is lagging behind

Wales

GP registration spine, hospital, GP and community health, social care, education, environment, housing, crime

SAIL – independent linkage and research environment

Scotland

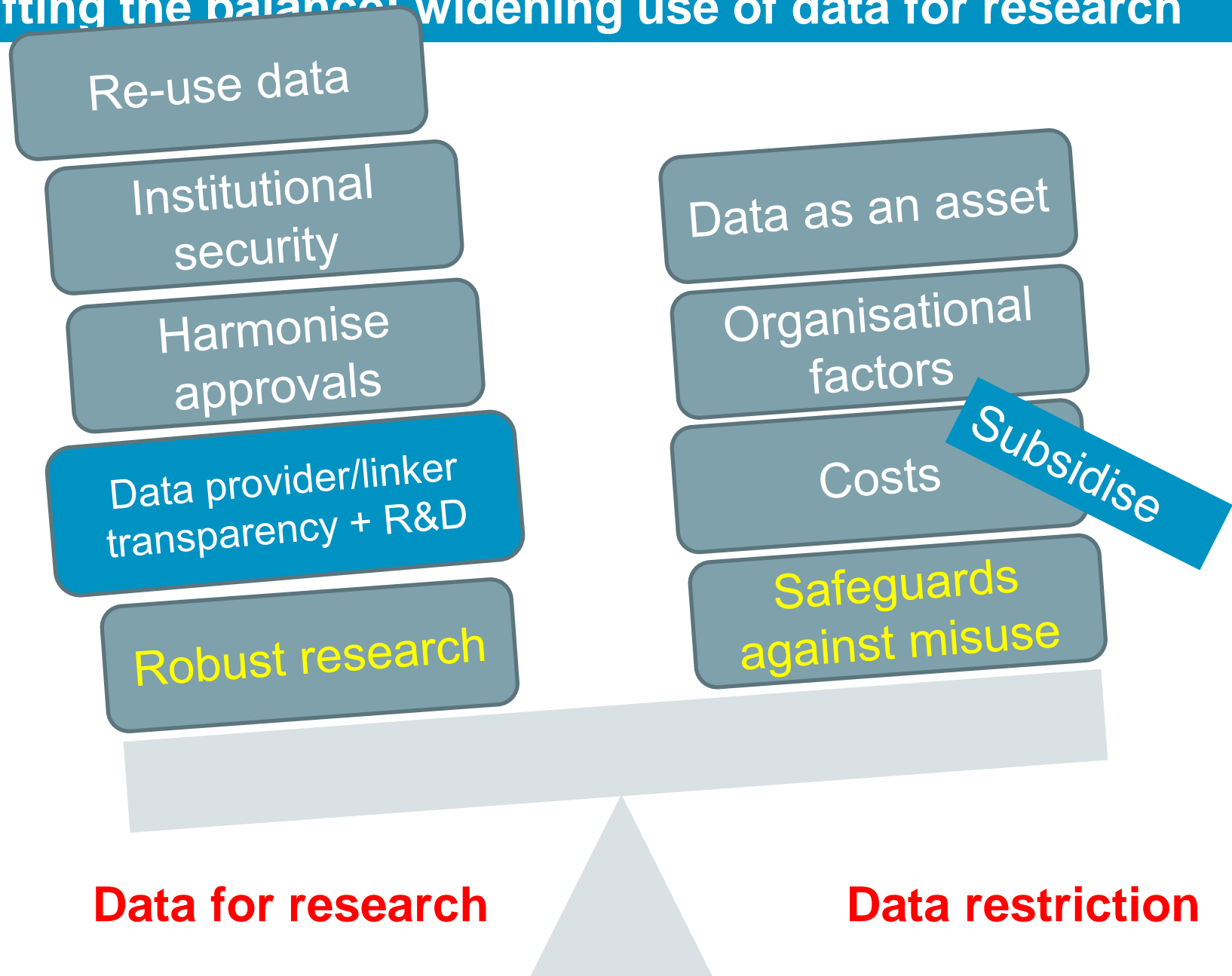
Registration spine, hospital, community prescribing, education, census, crime

eDRIS – independent linkage and research environment

England

- many large data providers
- established access arrangements
- no central relationship trusted third party

Shifting the balance: widening use of data for research



Thanks to:

- Lorraine Dearden, Bilal Nasim, Efrosini Setakis, Sam Sims, Louise Mc Grath Lone, James Doidge, Ruth Blackburn, Linda Wijlaars, Pia Hardelid, Ania Zylberstein, Max Verfuerden, Harvey Goldstein, Katie Harron, Jenny Woodman, Hannah Morris (UCL)
- Paul Wilkinson, Ai Milojevic, Jan van der Meulen (LSHTM)



Administrative Data
Research Centre
England