

# The ONS Longitudinal Study

Dr Oliver Duke-Williams

twitter: @oliver\_dw email: o.duke-williams@ucl.ac.uk

Making the most of Census microdata:  
An introductory workshop

21 November 2018, University of Manchester

## Census data in the UK / Great Britain

- Censuses held since 1801
- Current arrangement is three separate but coordinated censuses (E&W, S, NI)
- Users are generally most familiar with aggregate data

## Longitudinal data

- There are three longitudinal studies in the UK
- They have different sample sizes and cover different time periods
- They also differ in the range and amount of linked data
- All have secure access arrangements

# UK Census Longitudinal Studies

	Sample size	1971	1981	1991	2001	2011	Approx. sample members <sup>1</sup> , 2011
ONS Longitudinal Study	4/365.25	✓	✓	✓	✓	✓	614k
Scottish Longitudinal Study	20/365.25			✓	✓	✓	290k
Northern Ireland Longitudinal Study	104/365.25		✓	✓	✓	✓	512k

<sup>1</sup> Based on sample size and published populations



## What is in the LSes?

- Similarly to census microdata, all variables apart from individual identifiers
  - More detail than the safeguarded and open microdata files
  - Some (detailed) variables have additionally restricted access
- Imputed fields are included (and identified)
- Imputed records are not included

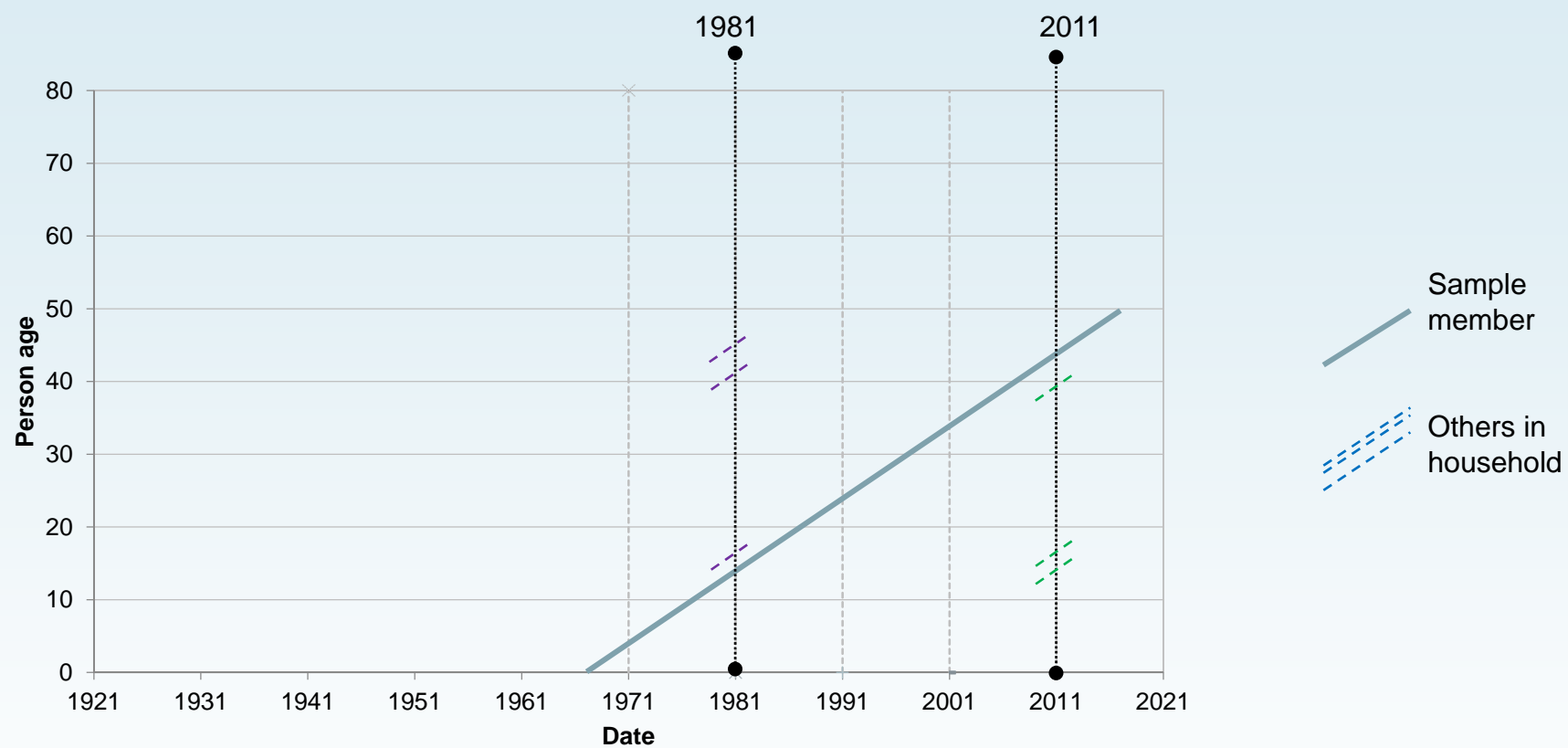
## What is in the LSes?

- As well as ‘standard’ variables, there are a variety of restricted variables
  - It can be possible to use these for analysis where appropriate
    - Example: birth dates will never be shown to the user, but might be used to create a derived indicator
    - Example: low level geography such as Output Areas can be used to attach area-level data, but analysis cannot allow a small area to be identified in output

## What is in the LSes?

- As well as LS sample members, the studies also include equivalent records for other persons in the household, referred to as ‘non-members’
- Non-members are not linked over time
  - It is sometimes possible to make reasonable assumptions about whether or not a non-member observed at two times is in fact the same person
  - For example, consistent date of birth and relationship to others in the household





# Census Data

## From each census

- Age, sex, marital status, country of birth
- Family and household types, communal establishments
- Housing: tenure, rooms and amenities
- Qualifications, economic activity, occupation, industry and social class
- Travel to work and one-year migration
- Geographical information

## More recent censuses

- Ethnicity (1991-2011)
- National identity (2011)
- Year of arrival (2011)
- Limiting long-term illness (1991-2011) & self-rated health (2001, 2011)
- Care-giving (2001, 2011)
- Religion (2001, 2011)
- Short-term migration (2011)
- Main language (2011)

## England/Wales (LS)

Civil registration system  
 Births of sample members  
 Births to sample mothers  
  
 Stillbirths/infant deaths  
 Deaths of sample members  
 Widow(er)hoods  
  
 NHS Patient Register  
 Immigration into England/Wales  
 Emigration from England/Wales  
  
 Cancer registrations  
 Cancer data

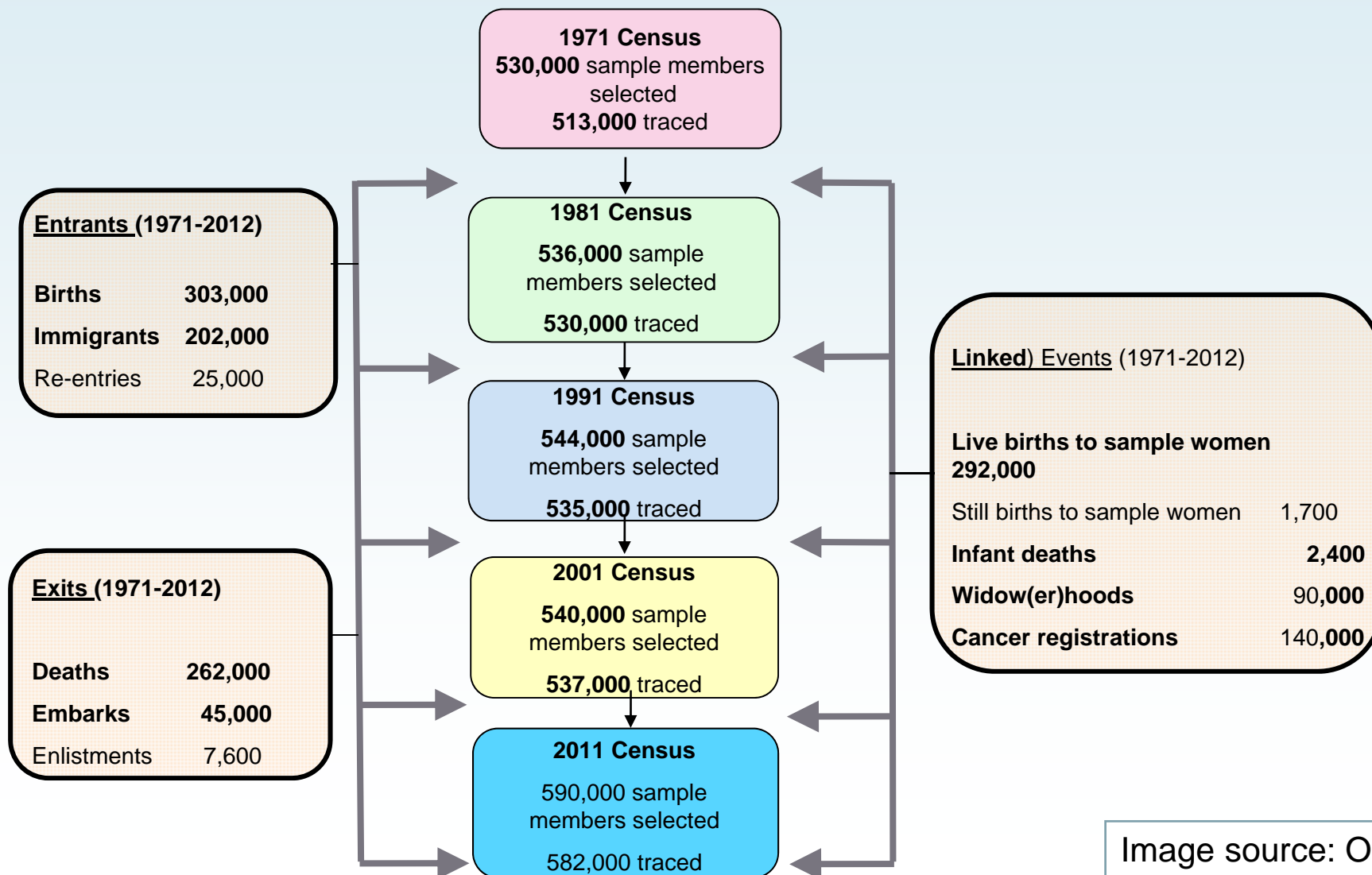
## Scotland (SLS)

Civil registration system  
 Births of sample members  
 Births to sample mothers  
 Births to sample fathers  
 Stillbirths/infant deaths  
 Deaths of sample members  
 Widow(er)hoods  
 Marriages  
 NHS Patient Register  
 Immigration into Scotland  
 Emigration from Scotland  
  
 Education data from ScotXed  
 Individual-level data from Schools Census,  
 attendance, absences/exclusions,  
 SQA attainment, qualifications  
  
 Special linkage, subject to approval  
 Hospital attendances  
 Maternity data  
 Cancer data  
 Prescribing data  
 Also: Weather and pollution data

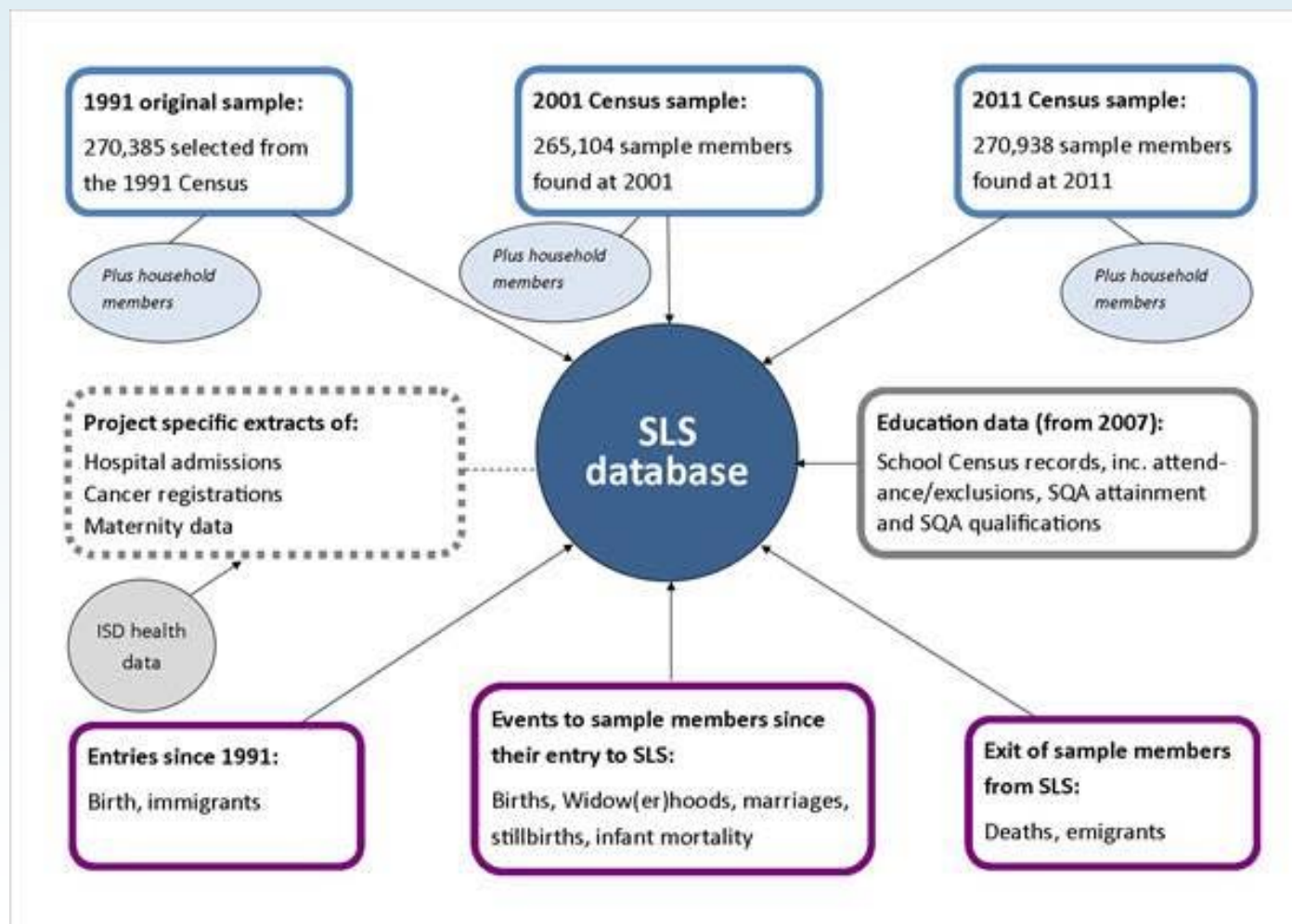
## Northern Ireland (NILS)

Civil registration system  
 Births of sample members  
 Births to sample mothers  
 Births to sample fathers  
 Infant deaths  
 Deaths of sample members  
  
 Health card registration system  
 Immigration into N. Ireland  
 Emigration from N. Ireland  
 Internal migration  
 Land and Property Services  
 Type of accommodation, value in 2005,  
 urban/rural etc.  
  
 Special linkage, subject to approval  
 Health data, including breast screening,  
 dental treatments, prescription of  
 antibiotics

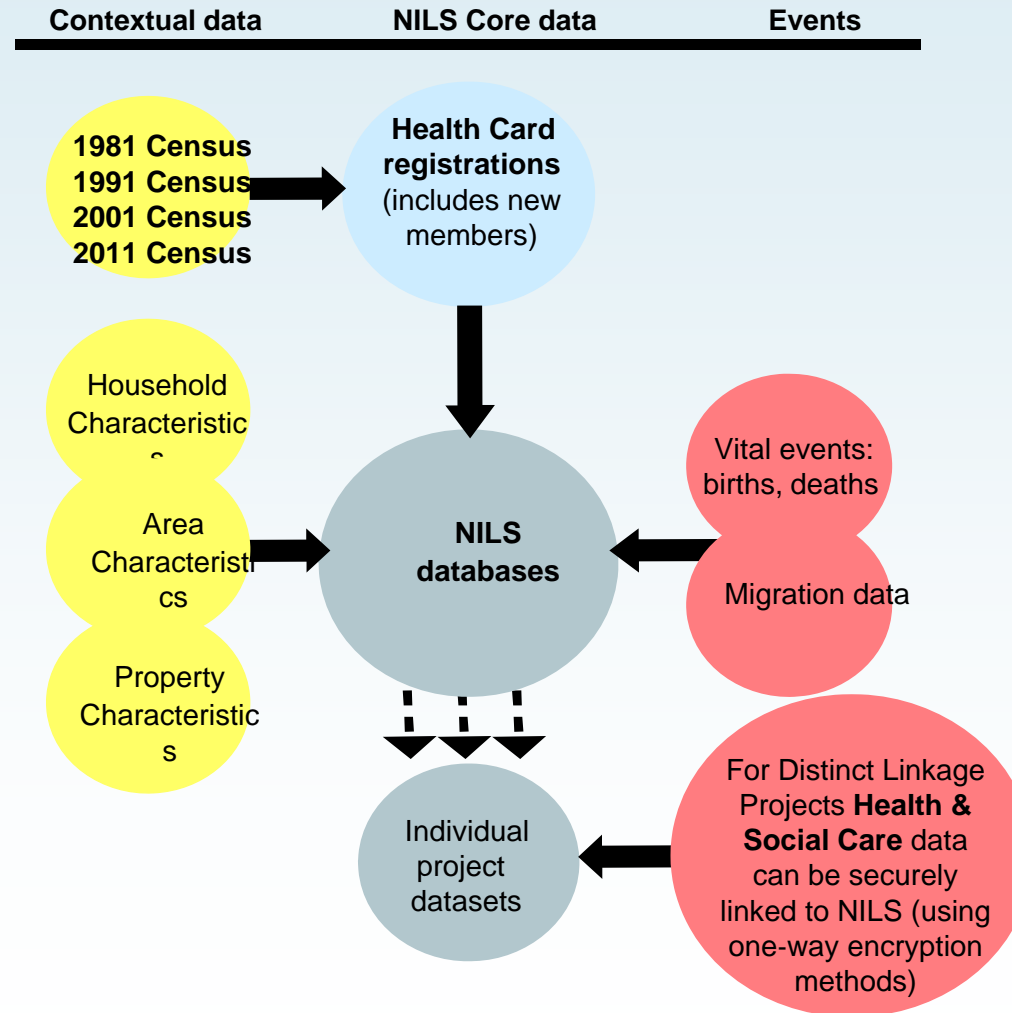
# LS structure: England & Wales



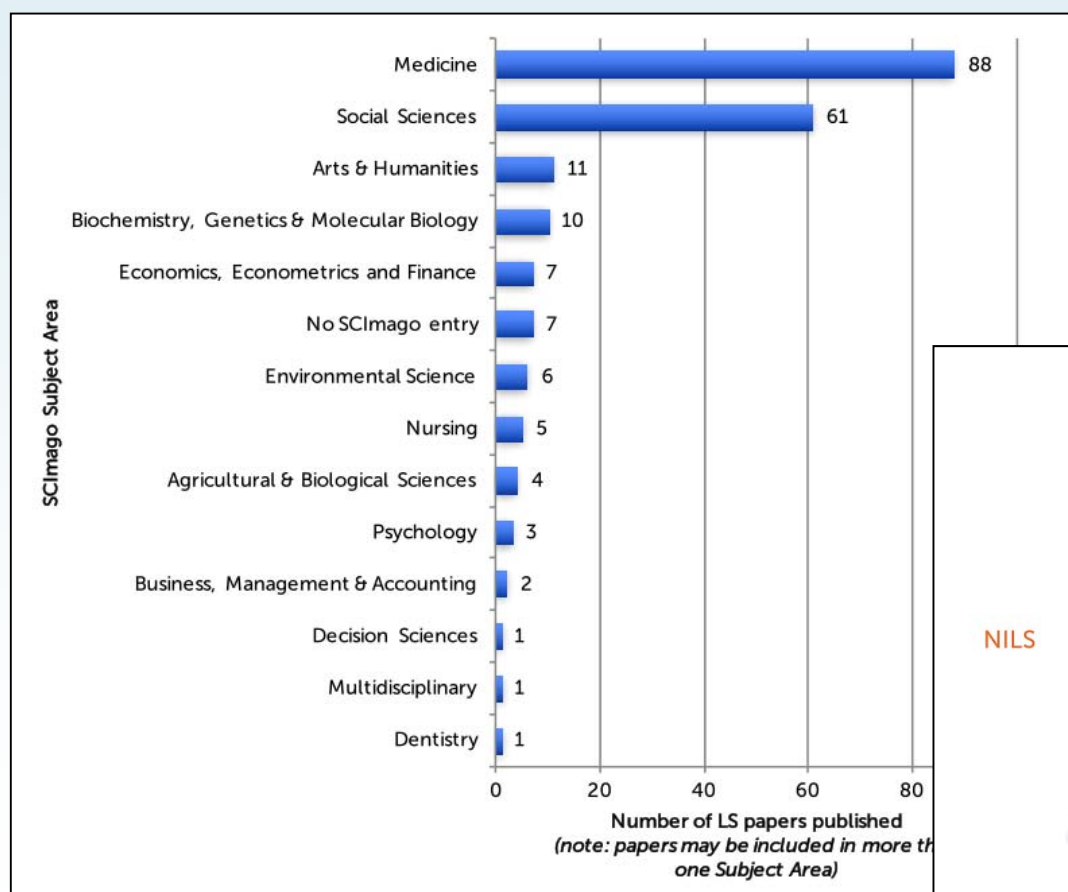
# LS structure: Scotland



# LS structure: Northern Ireland



# Who uses the LSes?



Source: Cox, F (2017) CALLS Hub  
Citation Analysis  
<https://calls.ac.uk/research-blog/>

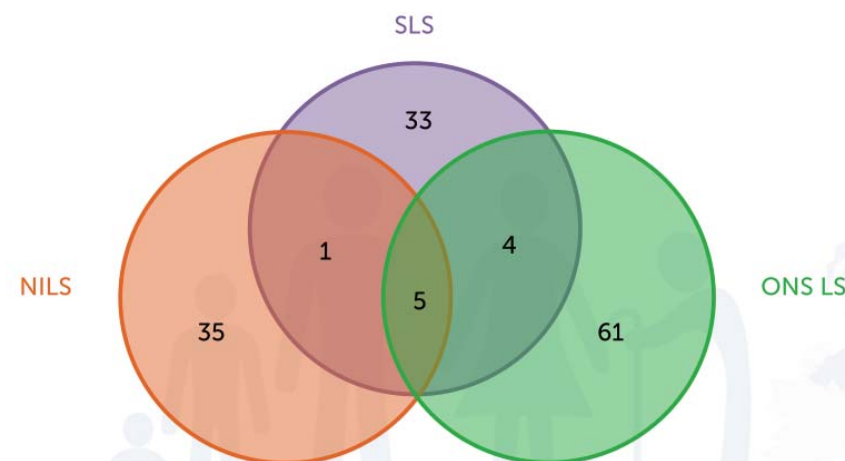


figure 1. Number of published papers per LS, Jan 2010 - Oct 2017. n = 139

## Example: how do people change transport mode used over time?

- We can compare cross-sections easily enough
  - But: we don't know whether those that used (mode x) in 2001 were the same people that used (mode x) in 2011, unless we use longitudinal data...

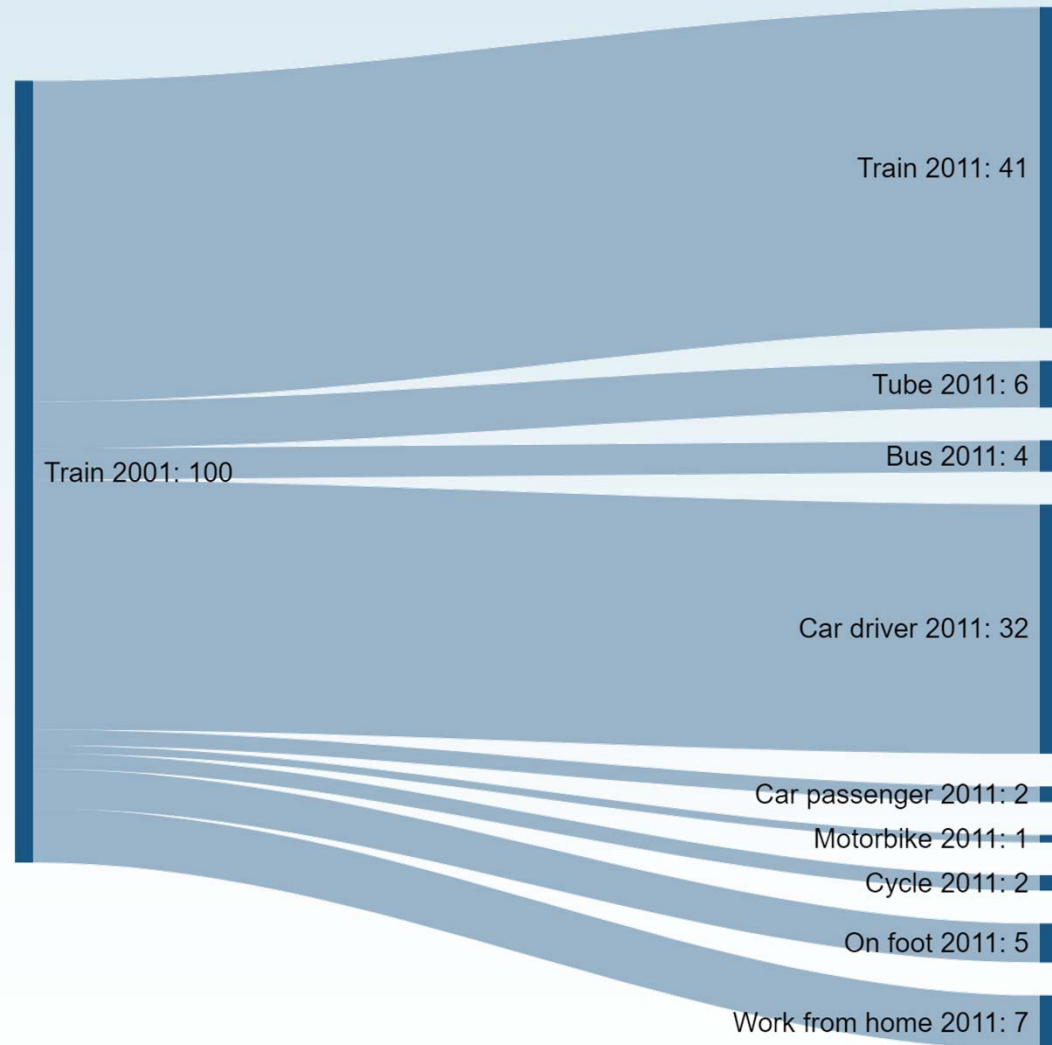


# Travel to work 2001 v 2011: longitudinal transitions

		Mode in 2011									Total
		Train	Tube	Bus	Driving a car or van	Passenger in a car or van	Motorcycle	Bicycle	On foot	Work mainly at or from home	
Mode in 2001	Train										
	Tube/Metro/Light rail/tram (E&W)										
	Bus/Minibus/Coach										
	Drive car/van										
	Passenger car/van										
	Motor cycle/Scooter/Moped										
	Bicycle										
	On foot										
	Works at/from home										
Total											

## Train

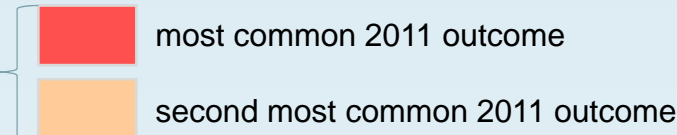
- Diagram shows percentage splits for train commuters in 2001 by travel to work mode in 2011
  - Persons are present at both times, and employed / self-employed at both times



Source: ONS Longitudinal Study

# Travel to work 2001 v 2011: longitudinal transitions

For each 2001 mode



		Mode in 2011									Total
		Train	Tube	Bus	Driving a car or van	Passenger in a car or van	Motorcycle	Bicycle	On foot	Work mainly at or from home	
Mode in 2001	Train	41%	6%	4%	32%	2%	1%	2%	5%	7%	6324
	Tube/Metro/Light rail/tram (E&W)	17%	33%	8%	24%	2%	1%	3%	5%	7%	3849
	Bus/Minibus/Coach	5%	3%	27%	39%	8%	0%	2%	12%	3%	10638
	Drive car/van	2%	1%	2%	82%	2%	1%	1%	4%	5%	93087
	Passenger car/van	3%	1%	8%	53%	20%	1%	2%	10%	3%	10522
	Motor cycle/Scooter/Moped	4%	2%	3%	56%	4%	19%	6%	4%	3%	1840
	Bicycle	3%	1%	4%	43%	4%	2%	30%	10%	3%	4415
	On foot	3%	2%	7%	40%	6%	1%	3%	34%	4%	14621
	Works at/from home	3%	1%	2%	54%	3%	0%	1%	7%	28%	14035
Total		7381	3428	7323	105507	6879	1248	4053	13075	10437	159331

- 30% retention rate for cyclists
- 34% retention rate for walkers

## Comparison with birth cohorts

- People unfamiliar with the LSes are often more familiar with the idea of birth cohort studies
  - Birth cohorts draw a sample of persons born in a particular year
  - Census longitudinal studies draw a sample across all persons regardless of age

## Comparison with birth cohorts

- Sample sizes
  - LS total samples are much bigger
  - LS individual year-of-age samples are smaller in England and Wales, but less subject to attrition
    - Starting cohort sizes 17-19K
    - ONS LS, 2011 per single year of age < 65 c. 5-8K
- Content
  - Cohort studies have much broader content
  - Sample size allows LS to have more detailed geography etc

## Comparison with birth cohorts

- Cohort studies are affected by sample attrition over time
  - We can produce subsets of LS sample members who have the same birth year as a cohort study
  - LS has much lower attrition
  - We can compare characteristics of the two groups in order to get a better idea of how representative the cohort sample remains
  - Comparison 1958 NCDS cohort with LS: Archer et al (forthcoming)

## Using the ONS LS

- Two access routes
  - In person at a secure setting
  - Submission of Stata etc scripts to be run remotely
- No data will be transferred out of the secure setting until it has had disclosure clearance

## Secure access

- Researcher Accreditation
  - Required for both Secure Lab and Secure Research Service
  - Experience + training + undertakings
- SL: Secure Access agreements
- SRS: Approved Research Projects



## Access to the the Lses: RSUs

- CeLSIUS
  - <https://ucl.ac.uk/celsius>
  - [celsius@ucl.ac.uk](mailto:celsius@ucl.ac.uk)
- NILS-RSU
  - <https://www.qub.ac.uk/research-centres/NILSResearchSupportUnit/>
  - [rsu@nisra.gov.uk](mailto:rsu@nisra.gov.uk)
- SLS-RSU
  - <https://sls.lscs.ac.uk/>
  - [sls@lscs.ac.uk](mailto:sls@lscs.ac.uk)

## Support in planning research

- Data dictionaries
- Advice from support officers

# Data dictionaries

**CeLSIUS**

UCL Home · CeLSIUS · CeLSIUS data dictionary · Details

**CeLSIUS**

- Home
- About CeLSIUS
- About the ONS Longitudinal Study
- Using the ONS Longitudinal Study
- Projects using the ONS LS
- Resources
- Contact us

**Variable details**

[ Back to Data Dictionary home | previous page ]

**LS variable metadata**

TEND11	
Table	ME11 2011 Census
Short description	Dwelling tenure, 2011.
Source	2011 Census. Questions H12/H13.
Code notes	0-9, -0, -9

**Variable coding**

Value	Label
0	Owens outright
1	Owens with a mortgage or loan
2	Shared ownership
3	Social rented, Rented from council
4	Social rented, Rented from Registered Social Landlord or housing association
5	Private rented, Private landlord or letting agency
6	Private rented, Employer of a household member
7	Private rented, Relative or friend of household member

**Tweets by @celsiusnews**

Looking forward to our conference tomorrow - we have a full programme of student research using #Census data research from the cross sectional and longitudinal anonymised samples [eventbrite.co.uk/e/student-](http://eventbrite.co.uk/e/student-)

CeLSIUS data dictionary

[ucl.ac.uk/celsius](http://ucl.ac.uk/celsius)

Variable	LS	Variable Description	Table	Table Description	Similar	More	Save
NS_DEP_TENH0	NILS	Census 2001 - National Statistics Deprivation Indicator - Tenure   A household is defined as deprived in the housing dimension if it is social housing	C012	2001 Census - Household Information	Yes	→	+
TENUREH0	NILS	Census 2001 - Tenure of Household	C012	2001 Census - Household Information	Yes	↓	+

**Variable Name** TENUREH0

**Table name** C012

**Full Description** Census 2001 - Tenure of Household

**Similar variables in other LSs (Guidance)**

- 6 ONS LS - TENH0 (table: ME01) - Household tenure, 2001.  
**Comments:** ONS category 9 - 'Lives rent free' not present in NILS. ONS 'Other' category appears more general than NILS' 'Private Rented: Other'
- 6 SLS - TENH0 (table: C22) - Household tenure, 2001.  
**Comments:** SLS category 9 - 'Lives rent free' not present in NILS.

**Similar variables across time within NILS**

- TENUREH8 (table: C812) - Census 1981 - Tenure
- TENUREH9 (table: C912) - Census 1991 - Tenure
- TENH1 (table: C112) - Census 2011 - Tenure of Household (NI)
- TENUREH8 (table: C812) - Census 1981 - Tenure
- TENUREH9 (table: C912) - Census 1991 - Tenure

CALLS-Hub data dictionary

[calls.ac.uk](http://calls.ac.uk)

# Questions?

## Acknowledgements

- The permission of the Office for National Statistics (ONS) to use the Longitudinal Study is gratefully acknowledged, as is the help provided by staff of the Centre for Longitudinal Study Information & User Support (CeLSIUS). CeLSIUS is supported by the ESRC Census of Population Programme (Award Ref: ES/K000365/1). The authors alone are responsible for the interpretation of the data.