Health Data in Practice lecture series

Electronic health records: geographical data

Carol Dezateux

27th October 2020

Institute of Population Health Sciences

wellcome

Learning Objectives

At the end of this lecture you will be able to

- Define output areas and understand their use in area-based measures

- Describe the main area-based measures of multiple deprivation hat can be linked to electronic health records

- Understand the limitations of area-based measures

- Define unique property reference numbers, their relation to postcodes and addresses, and their potential uses for individual measures

Output areas

- Output Areas (OA) OAs have been used for data collection and output of Census data since the 2001 Census

- OAs were created for Census data, specifically for the output of census estimates.

- The OA is the lowest geographical level at which census estimates are provided.

- They were introduced in Scotland at the 1981 Census and in all the countries of the UK at the 2001 Census

- See A Beginner's Guide to UK Geography Edition: 2020 v1.0 ONS Geography, Office for National Statistics in reading list for fuller account

Super Output areas (SOAs)

- SOAs were designed to improve the reporting of small area statistics and are built up from groups of output areas (OA).

- Lower Layer Super Output Areas (LSOAs): LSOAs have an average population of 1500 people or 650 households

- Middle Layer Super Output Areas (MSOAs): MSOAs have an average population of 7500 residents or 4000 households

- SOAs are defined by 9 characters: The first character is either E for England or W for Wales. The next two characters are either 01 for Lower Layer, or 02 for Middle Layer and the remaining six characters make up the unique 6-digit tag for each zone.

- Statistics for lower layer super output areas (LSOA) and middle layer super output areas (MSOA) were originally released in 2004 for England and Wales. Scotland also released statistics for data zones (DZ), that were equivalent to LSOAs, in 2004 and intermediate geographies (IG), that were equivalent to MSOAs, in 2005. Northern Ireland introduced LSOAs in 2005 but do not have an MSOA geography.

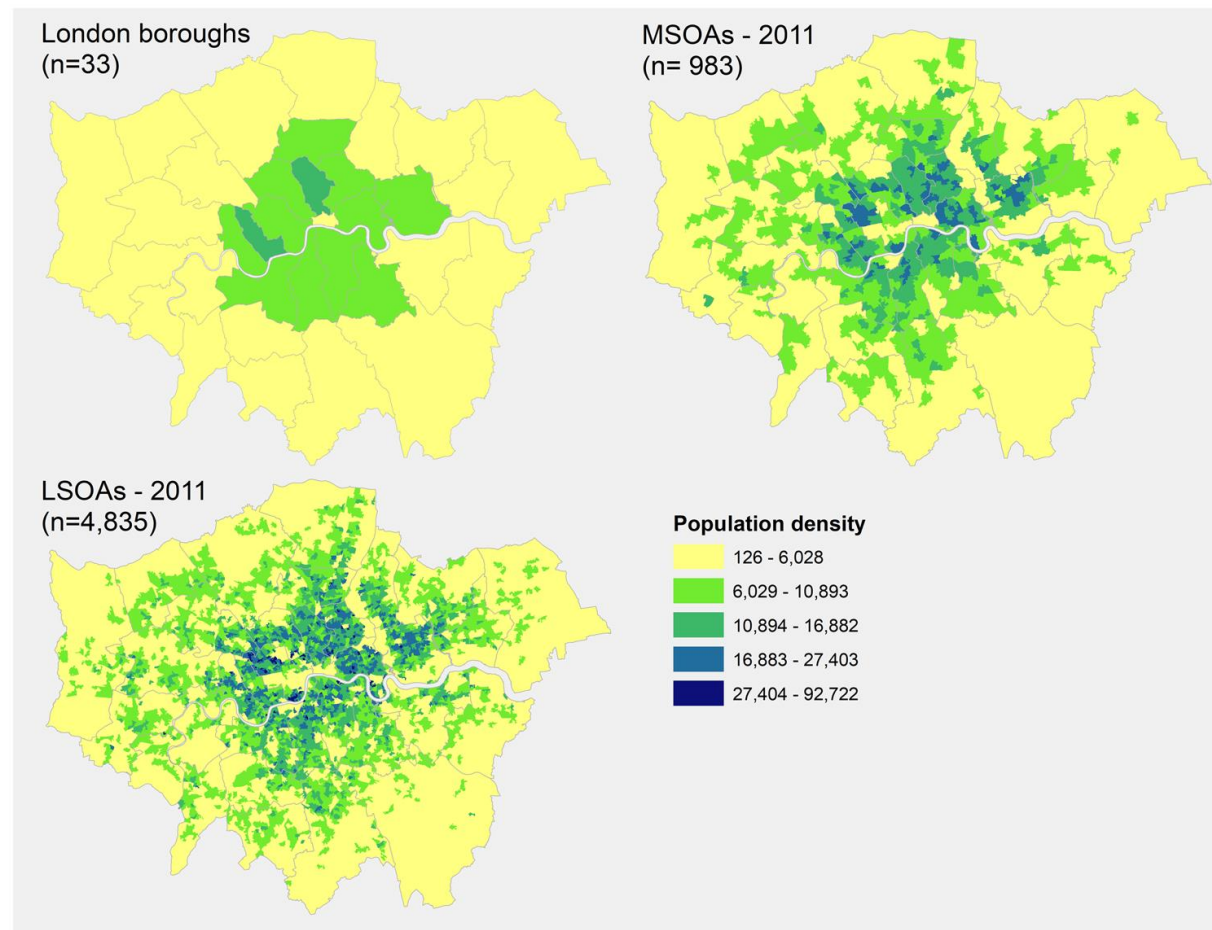| Geography	England	Wales	Scotland	Northern Ireland
LSOA	32,482	1,896	6,505*	890
MSOA	6,781	413	1,235**	n/a
Notes:				
1. * DZ				
2. ** IG				

Figure 5 Population density (inhabitants per square kilometre) in the Greater London area. The three maps depict the ...
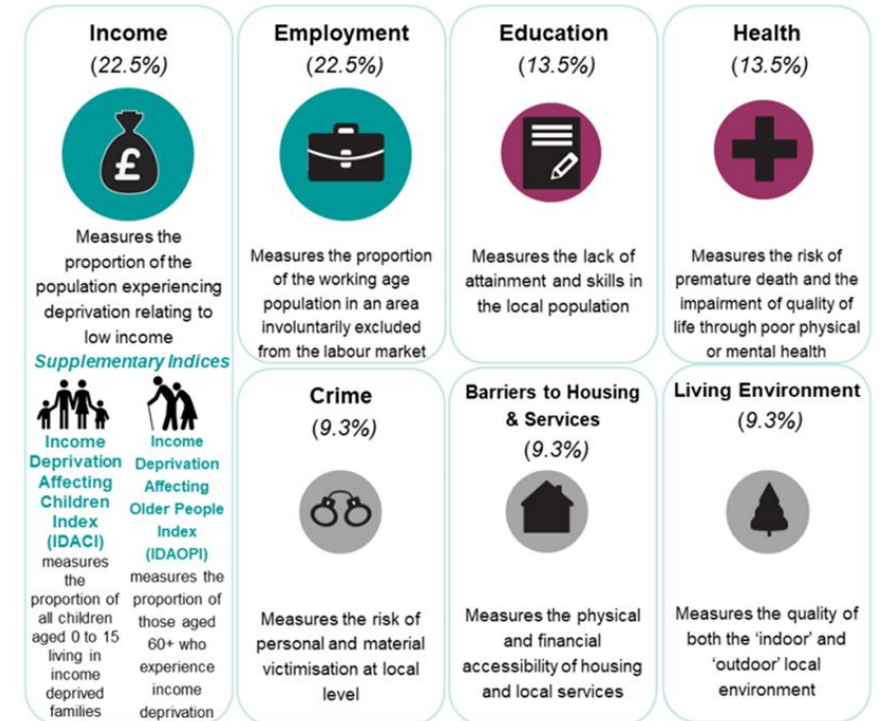
OXFORD
UNIVERSITY PRESS

Area based measures of deprivation

1. **Index of Multiple Deprivation**

- Includes 7 domains of deprivation plus overall score

- Issued by Department of Communities & Local Government

- Issued at LSOA level- can be aggregated by population-weighted averages.

- Cannot make direct comparisons between versions by year as scores
- are calculated differently by year.

- Can only comment on relative rank differences between years.

There are 7 domains of deprivation, which combine to create the Index of Multiple Deprivation (IMD2019):

Income (22.5%) Measures the proportion of the population experiencing deprivation relating to low income

Employment (22.5%) Measures the proportion of the working age population in an area involuntarily excluded from the labour market

Education (13.5%) Measures the lack of attainment and skills in the local population

Health (13.5%) Measures the risk of premature death and the impairment of quality of life through poor physical or mental health

Supplementary Indices

Income Deprivation Affecting Children Index (IDACI) measures the proportion of all children aged 0 to 15 living in income deprived families

Income Deprivation Affecting Older People Index (IDAOPI) measures the proportion of those aged 60+ who experience income deprivation

Crime (9.3%) Measures the risk of personal and material victimisation at local level

Barriers to Housing & Services (9.3%) Measures the physical and financial accessibility of housing and local services

Living Environment (9.3%) Measures the quality of both the 'indoor' and 'outdoor' local environment

Area based measures of deprivation

2. Jarman Score

- Initially developed as a measure of GP workload.

- Used to determine deprivation payments to GPs by DH.

- Includes:
    - % people in households aged >65 living alone
    - % people in households aged <5
    - 1 person >16 plus 1 or more children <1
    - those >16 seeking work
    - persons in households with >1 person per room
    - moved within last year
    - households headed by new commonwealth /Pakistan

- Data is linked to last census so can be out of date.
- Does not indicate the proportion of people in an area who are deprived.
- Biased towards urban population.

Area based measures of deprivation

3. Townsend Material Deprivation Score

• Uses census data

• 4 domains are:
  - Unemployment: % economically active who are employed
  - car ownership
  - owner occupation
  - overcrowding (>1 person per room)

• Can be used to look at small areas.
• Highly correlated with measures of ill-health.
• Easy to calculate because sum of standardised scores.

• Data can be out of date as linked to last census- particularly in terms of housing tenure.
• Does not indicate the proportion of people in an area who are deprived.
• Better indicator of deprivation in urban areas that rural areas.

Area based measures of deprivation

4. Carstairs

- Uses census data
- 4 domains are:
    - Unemployment: % economically inactive males aged >16 who are employed
    - % households without car ownership
    - overcrowding (>1 person per room)
    - % residents in households with an economically active head of household in Social Class IV or V approximated from NS-SEC

- Based on census so allow objective results for whole population.
- Can be used to look at small areas.
- For an outcome such as low birthweight, Carstairs scores have been shown to perform better than individual social class in describing the extent of inequalities in the population

- Data can be out of date as linked to last census- particularly in terms of housing tenure.
- Better indicator of deprivation in urban areas that rural areas.

Area based measures : use in COVID
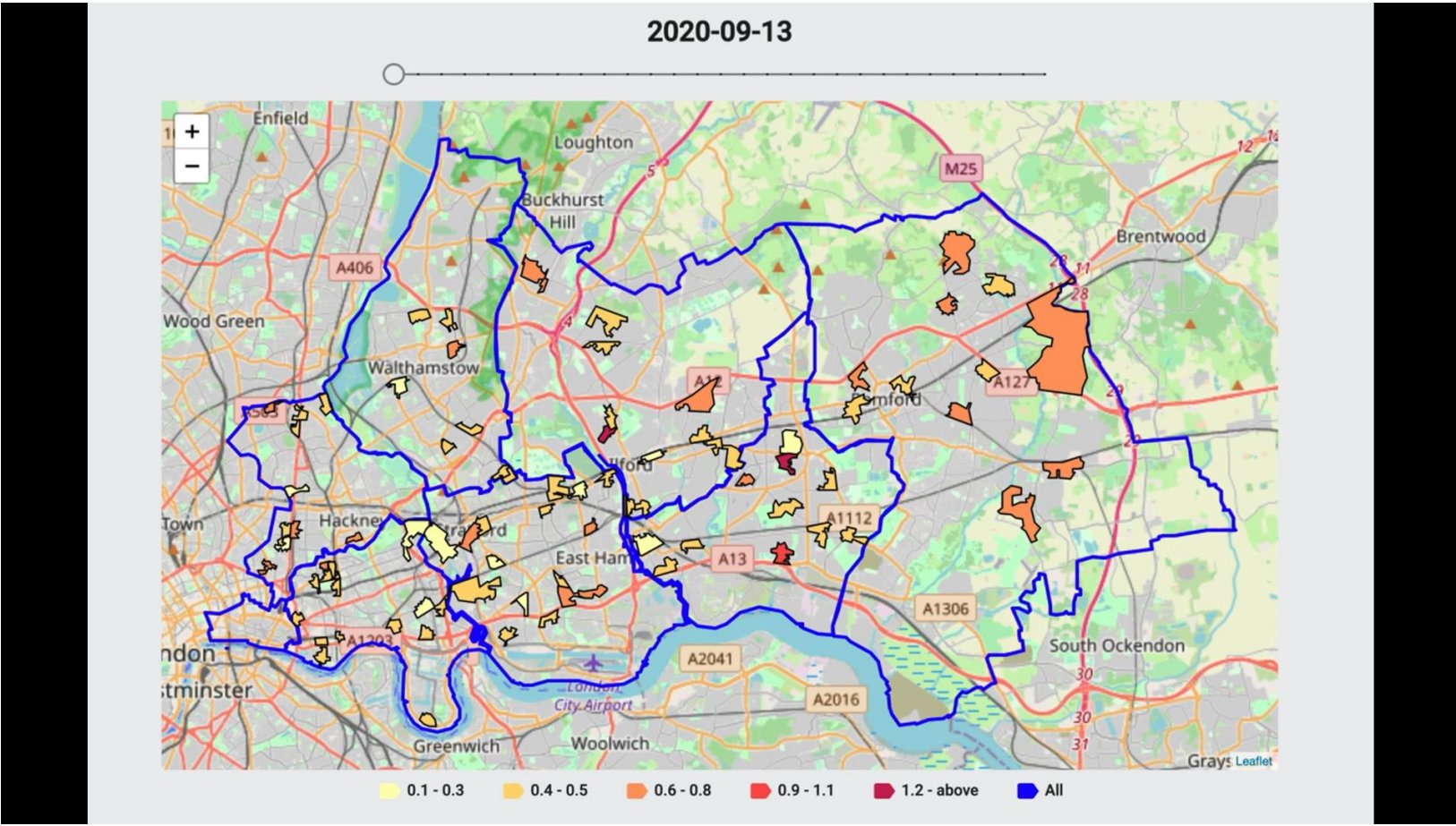
Area based measures of deprivation: limitations

- Comparisons between different UK countries limited

- Measures which rely on census data become out of date

- Comparisons over time limited

- Most measures limited for rural areas

- Ecological measure of deprivation: may misclassify individuals

- Car ownership and housing tenure measures may not have same meaning over time and in different settings

Questions???

Postcodes

- Postcodes are alphanumeric references comprising an outward code of 2–4 characters and an inward code of three characters.

- It is not a unique identifier of your property

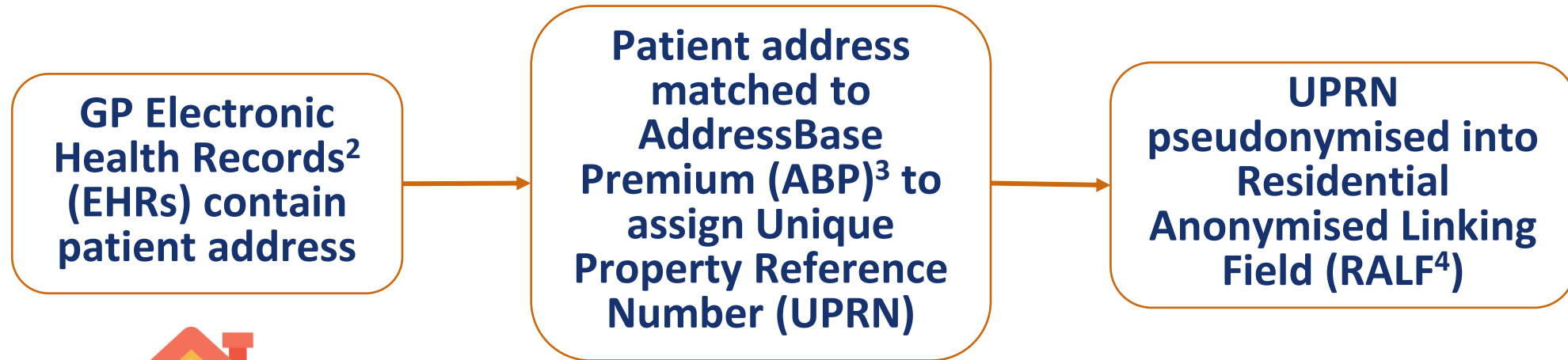The postcode is structured hierarchically, supporting four levels of geographic unit:

Example	Geographic Unit	Number in UK*
PO	Postcode Area	124
PO15	Postcode District	3,118
PO15 5	Postcode Sector	12,433
PO15 5RR	Unit Postcode	Approximately 1.77 million (live)**

Address level geography: Unique Property Reference Numbers

- Each property in Great Britain is assigned a Unique Property Reference Number (UPRN)

- The UPRN is the unique identifier for every spatial address in Great Britain.

- It provides a comprehensive, complete, consistent identifier throughout a property's life cycle – from planning permission through to demolition.

- Most people know their address, so this is used as a geographic reference when collecting data.

- This reference can then be related to any geographic unit used for statistical production, such as a local authority district or electoral ward.

- The UPRN is numeric and can be up to 12 digits. As of September 2020, there are 39.3 million GB UPRNs.

Households

| GP Electronic Health Records[2] (EHRs) contain patient address	→	Patient address matched to AddressBase Premium (ABP)[3] to assign Unique Property Reference Number (UPRN)	→	UPRN pseudonymised into Residential Anonymised Linking Field (RALF[4])

Household = **Patients sharing the same RALF**

[2] East London Discovery Data Service

[3] AddressBase Premium (Ordnance Survey)

[4] Rodgers et al. 2009

Age standardised prevalence: Household composition
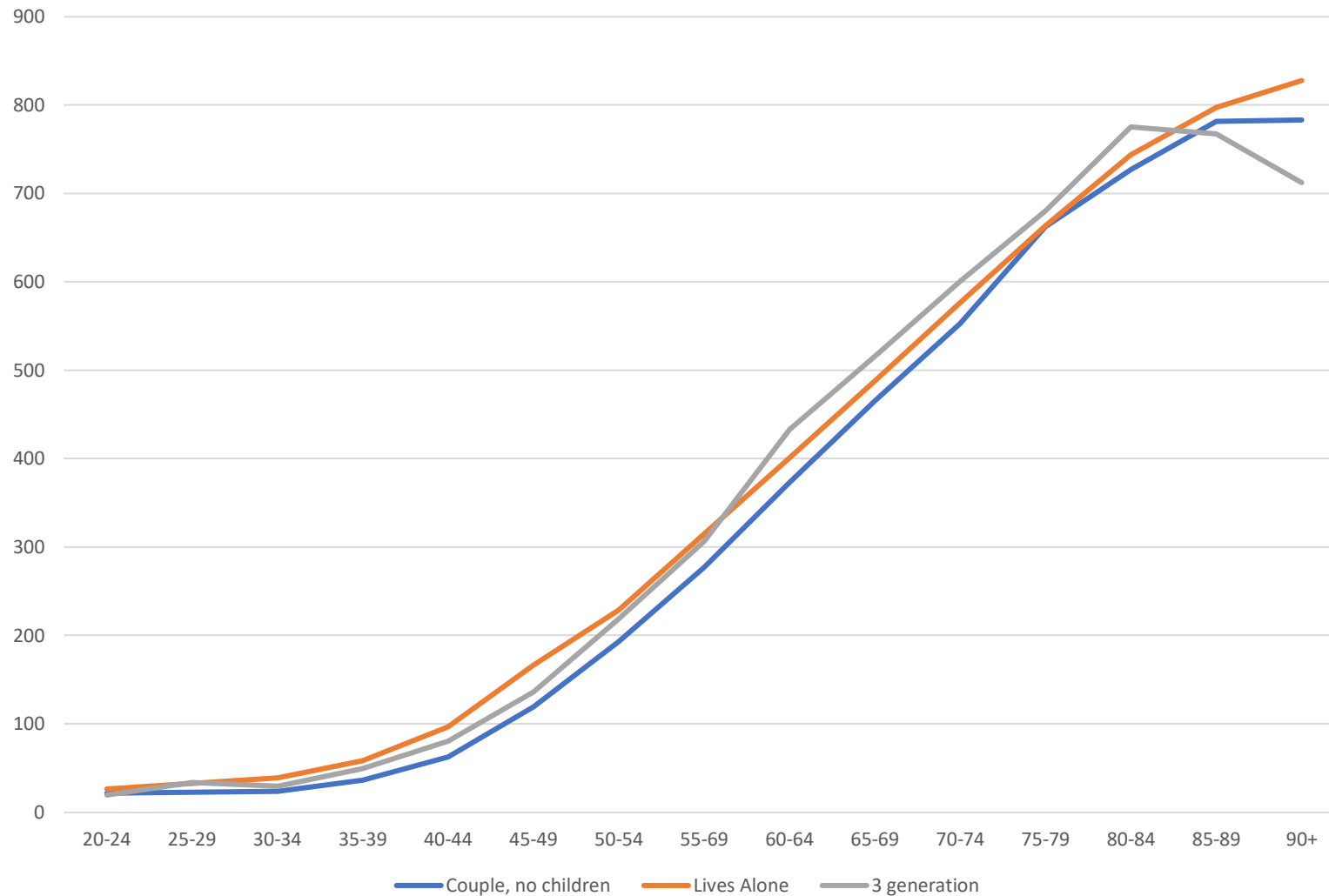
Age standardised prevalence per 1000 (95% CI)

Alone   268 (265,271)
3-Gen   266 (263,269)
Couple   243 (240,246)

Couple, no children — Lives Alone — 3 generation

Questions???

Tutorial

You will be assigned an exercise to bring for discussion at next week's tutorial based on today's lecture

This slide set will be made available to you on QMPlus together with a reading list