

Health & Wealth across England's population at local authority level: Methodology

IPPR

September 2021

This document sets out methodology used in the analysis presented to IPPR in July 2021. Providing a high level overview of how the data was compiled and how the analysis was performed.

Summary of approach

Health, education, employment and economic administrative data over 2015-2020 for England was collated. The raw data used and methodology was consistent with that of the health index for those variables that are within the ONS's [Health Index](#), where possible. Any variation in methodology to the health index for those variables is set out below in Table 1, which was due to data availability.

Variables not within the health index were primarily sourced from the Annual Population Survey, obtained from NOMIS.

Data was analysed at an Upper Tier Local Authority (UTLA) level, where the raw data was at a lower level (such as Lower Tier Local Authority (LTLA)) the data was aggregated up to UTLA level weighted by population, data at a higher level (such as regional) was applied equally to all UTLAs within the region due to difficulties disaggregating data, as discussed within the health index methodology.

Change over time analysis was carried out for all years available. Due to data availability the correlation and clustering analysis used 2019 data, as very few variables had 2020 data, where data was not available at 2019 the most recently available data was used.

Pearson correlation coefficients were calculated between all variables to produce a heatmap. Linear regression was used to explore the relationship between key pairs of health and wealth indicators. Hierarchical clustering was used to cluster geographical areas by wealth variables and analyse how both the wealth and health variables vary by cluster.

Data variables

Health, education, employment and economic data was collated from a variety of sources as set out in Table 1. Where possible the data and methodology used was consistent with that of the ONS [Health Index](#). The methodology used for the health index can be found [here](#). Any methodological variations due to data availability are set out in Table 1. The methods used for those variables not used within the Health Index, where possible, are consistent with those of the Health Index.

Where reasonable, raw data was used, however, for physical health conditions, wellbeing and behaviour risk factors the composite variable calculated for the Health Index was used to allow a wide range of data to be easily comparable. The raw data used within the health index for these variables is set out in Appendix 1. This data is only available to 2018 currently.

Table 1: List of health, education, employment and economic variables used in this analysis

Variable	Source	Comment
Physical Health Condition	Health index	Composite data directly from the health index. See Appendix 1 for raw data sources.
Personal wellbeing		
Behavioural risk factors		
Disability that impacts daily activities	ONS	As per the health index
Life satisfaction	ONS	
Suicide	ONS	
Public green space	ONS	
Early years development	PHE Fingertips	
Rough sleeping	Ministry of Housing, Communities and Local Government	
Child poverty	Department of Work and Pensions	
Workplace safety	Health and safety executive	
Depression	NHS Digital QOF indicators	
Infant mortality	PHE Fingertips	
Child obesity	PHE Fingertips	
Unemployment	PHE Fingertips	Similar to health index but uses PHE Fingertips data
Travel to GP	Department of transport	Similar to health index but uses travel time (walking) to service by local authority
Travel to sports/leisure centre	Department of transport	Similar to health index but uses travel time (walking) to service by local authority
Healthy life expectancy	PHE Fingertips	Similar to health index but uses PHE fingertips data combined assuming a 50:50 male female split across UTLAs.
Life expectancy	PHE fingertips	Additional to health index - as per healthy life expectancy but using life expectancy data
Gross Domestic Product (GDP)	ONS	Additional to health index - real terms GDP by LTLA aggregated to UTLA and calculated per head

IMD	ONS	Additional to health index – only available at 2015 and 2019
PIP/DLA	Department of work and pensions	Additional to health index - benefit data mapped from LTLA to UTLA using weighted average population to calculate to per 1,000 population. Adjusted for inflation using the average CPI inflation over each year
NHS Spend per head	NHS England	Additional to health index – only available at 2019, using ONS mapping files the data was mapped from CCG to GP and then aggregated to LSOA level and UTLA level weighted by population
Household income	ONS	Additional to health index - income data mapped from LTLA to UTLA using weighted average population to calculate to per 1,000 population. Adjusted for inflation using the average CPI inflation over each year
NVQ4+	NOMIS	Percentage of working age population with an NVQ4+ qualification
Total wealth per head	ONS	Only available at regional level, equally applied to each UTLA. Adjusted for inflation using the average CPI inflation over each year.

Geographical aggregation

Where possible raw data was used at UTLA level. Where any aggregation was required this was in line with the health index methods, specifically using weighted population adjustments and official mapping files to UTLA. Over the time period analysed a number of boundary changes have occurred, the same approach as the Health Index methodology was taken in our analysis to allow for boundary changes. Population estimates have been obtained from PHE Fingertips to allow for aggregation or boundary changes. Population data required for the UTLA Buckinghamshire was unavailable and so estimates from Eurostat were obtained.

Where data was at a higher level (e.g. regional) the value for each UTLA within a region was as per the raw data. This is due to difficulties of disaggregating data as set out within the health index methodology.

Managing date differences

Where data differs from calendar year (i.e. data by tax year) this is assigned to the year in which most of the source period falls (e.g. 2015/16 becomes 2015), this is consistent with the approach used in the health index. Where data is across three-year aggregates the data has been assigned to the final year covered (e.g. 2016-2018 is assigned to 2018), in line with the Health Index.

Analysis

Where possible, the correlation, regression and cluster analysis used 2019 data, if unavailable the most recent data was used. Isles of Scilly and City of London data was excluded from the analysis due to small populations.

Correlation analysis

A correlation heatmap was produced to understand the relationship between the selective health and wealth variables. Pairwise Pearson correlations were calculated between each of the variables with negative values

coloured in purple and positive ones coloured in red. Low correlations, which with absolute values less than 0.5, were masked in white.

Regression analysis

To investigate how key health indicators were associated with key wealth indicators, linear regression was used to explore the linear relationship between pairs of health and wealth indicators and their effect sizes, at UTLA level. Table 2 sets out the pairings used, as part of the regression analysis. For each pair, the health indicator (the outcome) was regressed against the wealth indicator (the exposure), both unadjusted and adjusted for mean age in each UTLA. The associations between these pairs of indicators (R^2) and the effect sizes were compared across pairs, for the unadjusted and adjusted analyses.

Table 2: Health indicator (outcome) and wealth indicator (exposure) pairs used as part of the regression analysis

Health indicator (outcome)	Wealth indicator (exposure)
Life expectancy	Child poverty rate
Child obesity	
Healthy life expectancy	
Infant mortality	
Work-limiting disability	Education level (NVQ4+)
Depression prevalence	Disability benefit payment per capita
Healthy life expectancy	Unemployment rate
Life satisfaction index	Travel time to GP
Personal well-being index	
Physical health conditions	Public green space
Physical health conditions	Total wealth per capita
Personal well-being index	Travel time to town centre

Cluster analysis

To understand the geographical distribution of wealth indicators we conducted clustering analysis. A hierarchical clustering algorithm was used to derive the clusters, with a distance matrix defined as the Manhattan distance between standardised values for these wealth indicators among UTLAs. This was conducted using complete linkage (or farthest neighbour clustering).

Change over time analysis

How the variables changed over time both nationally and regionally was charted to observe trends, where data was available (e.g. where the raw data provided a regional breakdown).

At a UTLA level the data was ordered for each variable to identify which UTLA's were in the top or bottom 10 in the earliest and latest year available (usually 2015 and 2019). The absolute change between variables in the earliest and latest data were compared and the five UTLAs with the biggest increase and decrease were charted.

Jonathan Pearson-Stuttard
Partner
jonathan.pearson-stuttard@lcp.uk.com

Rebecca Sloan
Consultant
rebecca.sloan@lcp.uk.com

The use of our work

This work has been produced by Lane Clark & Peacock LLP under the terms of our written agreement with the IPPR ("Our Client").

This work is only appropriate for the purposes described and should not be used for anything else. It is subject to any stated limitations (eg regarding accuracy or completeness). Unless otherwise stated, it is confidential and is for your sole use. You may not provide this work, in whole or in part, to anyone else without first obtaining our permission in writing. We accept no liability to anyone who is not Our Client.

If the purpose of this work is to assist you in supplying information to someone else and you acknowledge our assistance in your communication to that person, please make it clear that we accept no liability towards them.

About Lane Clark & Peacock LLP

We are a limited liability partnership registered in England and Wales with registered number OC301436. LCP is a registered trademark in the UK (Regd. TM No 2315442) and in the EU (Regd. TM No 002935583). All partners are members of Lane Clark & Peacock LLP. A list of members' names is available for inspection at 95 Wigmore Street, London, W1U 1DQ, the firm's principal place of business and registered office.

Lane Clark & Peacock LLP is authorised and regulated by the Financial Conduct Authority and is licensed by the Institute and Faculty of Actuaries for a range of investment business activities. Locations in London, Winchester, Ireland, and - operating under licence - the Netherlands.

© Lane Clark & Peacock LLP 2021

<https://www.lcp.uk.com/emails-important-information> contains important information about this communication from LCP, including limitations as to its use.

Appendix 1 – Health Index Composite raw data

The three variables that use the Health Index data directly are:

- Physical Health Conditions
- Personal wellbeing
- Behavioural risk factors

The data is weighted using time series factor analysis and then scaled to a base of 100 for England. Further information on the factor analysis and scaling can be found in section 17 and 20 of the [methodology document](#) for the Health Index. Values higher than 100 indicate better health than England in 2015 and values below 100 indicate worse health.

The raw data used by the Health Index for each of these variables is set out below:

Physical Health conditions

Prevalence data for six broad conditions was collected and calculated from NHS Digital QOF prevalence indicators (the percentage of GP patients on a practice register for the relevant condition). For some broad conditions more than one indicator was used as set out below:

- **Dementia**
 - Consists of the QOF prevalence of dementia (the percentage of GP patients on a practice register for dementia)
- **Musculoskeletal conditions**
 - Consists of two QOF prevalence indicators combined using a simple average. The two indicators combined were:
 - QOF prevalence of rheumatoid arthritis (for patients over 16 years old); and
 - QOF prevalence of osteoporosis
- **Respiratory conditions**
 - Consists of two QOF prevalence indicators combined using a simple average. The two indicators combined were:
 - QOF prevalence of asthma; and
 - QOF prevalence of Chronic Obstructive Pulmonary Disease
- **Cardiovascular conditions**
 - Consists of five QOF prevalence indicators combined using a simple average. The two indicators combined were:
 - QOF prevalence of coronary heart disease;
 - QOF prevalence of stroke and transient ischaemic attack;
 - QOF prevalence of atrial fibrillation
 - QOF prevalence of heart failure; and
 - QOF prevalence of peripheral arterial disease

- **Cancer**
 - Consists of the QOF prevalence of cancer (the percentage of GP patients on a practice register for cancer)
- **Kidney Disease**
 - Consists of the QOF prevalence of chronic kidney disease (the percentage of GP patients over 18 years old on a practice register for chronic kidney disease)

Personal well-being

The mean score (out of 10) of four indicators of respondents (over 16 years old) from the APS Integrated Household Survey was used to create a personal wellbeing composite. The indicators used included:

- **Life satisfaction**
- **Life worthwhileness**
- **Happiness**
- **Anxiety**

Behaviour risk factors

Data for five risk factors was collected and calculated from a variety of sources as set out below:

- **Alcohol misuse**
 - Consists of the hospital episode statistics (HES) hospital admission episodes for alcohol-related conditions. This is a directly age-standardised rate per 100,000. Calculated by PHE from NHS Digital HES and ONS mid-year population estimates.
- **Drug misuse**
 - Consists of two HES admission indicators combined by addition. A previously this is a age-standardised rate per 100,000 produced by NHS Digital. The two indicators combined were:
 - HES hospital admission with a primary diagnosis of drug poisoning by illicit drugs; and
 - HES hospital admissions with a primary diagnosis of drug-related mental health and behavioural disorders
- **Smoking**
 - Consists of the ONS's smoking prevalence in adults aged 18 via the Annual population survey.
- **Physical activity**
 - Consists of the percentage of adults (over 19 years old) who are physically active for 150 minutes or more per week from the Active Lives Survey.
- **Healthy eating**
 - Consists of the percentage of adults eating five or more portions of fruit and vegetables on a "usual day" from the Active Lives Survey.