



This is an author produced version of *Cohort Profile: The Yorkshire Health Study*.

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/84163/>

---

**Article:**

Green, M.A., Li, J., Relton, C., Strong, M., Kearns, B., Wu, M., Bissell, P., Blackburn, J., Cooper, C., Goyder, E., Loban, A. and Smith, C. (2014) Cohort Profile: The Yorkshire Health Study. *International Journal of Epidemiology*. ISSN 1464-3685

<http://dx.doi.org/10.1093/ije/dyu121>

---



*promoting access to  
White Rose research papers*

[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<http://eprints.whiterose.ac.uk/>

# Cohort Profile: The Yorkshire Health Study

Mark A Green\*<sup>1</sup>, Jessica Li<sup>1</sup>, Clare Relton<sup>1</sup>, Mark Strong<sup>1</sup>, Benjamin Kearns<sup>1</sup>, Mengjun Wu<sup>2</sup>, Paul Bissell<sup>1</sup>, Joanna Blackburn<sup>1,3</sup>, Cindy Cooper<sup>1</sup>, Elizabeth Goyder<sup>1</sup>, Amanda Loban<sup>1</sup>, Christine Smith<sup>3</sup>.

<sup>1</sup> School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK.

<sup>2</sup> Institute of Mental Health, University of Nottingham, Nottingham, UK.

<sup>3</sup> Barnsley Hospital National Health Service (NHS) Foundation Trust, Barnsley, UK.

\* Corresponding author: Dr Mark A Green PhD, Public Health Section, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, United Kingdom. E-mail: [mark.green@sheffield.ac.uk](mailto:mark.green@sheffield.ac.uk). Tel: (+44) 0114 222 0838. Fax: (+44) 0114 272 4095.

Word count: 2374

## Summary

The Yorkshire Health Study is a longitudinal observational regional health study collecting health information on the residents from the Yorkshire and Humberside region in England. The second wave of data collection is currently underway. The study aims to inform National Health Service (NHS) and local authority health related decision making in Yorkshire, with wider implications from findings as well. The first wave contains records for 27,806 individuals (2010-2012), aged between 16 and 85 from one part of Yorkshire (South Yorkshire), with the second wave expanding data collection to the whole of the Yorkshire and

Humberside region. Data were collected on current and long-standing health, health care usage and health-related behaviours, with a particular focus on weight and weight management. The majority of individuals have also given consent for record linkage with routine clinical data, allowing the linking to disease diagnosis, medication use and health care usage. The study encourages researchers to utilise the sample through the embedding of randomised controlled trials, other controlled trials and qualitative studies. To access the anonymised data or use the sample to recruit participants to studies, researchers should contact Dr Clare Relton ([c.relton@sheffield.ac.uk](mailto:c.relton@sheffield.ac.uk)).

MeSH keywords: Cohort Studies, Longitudinal Studies, Patient Selection, England, Obesity.

## **Key Messages**

- The cohort research platform facilitates multiple randomised controlled trials to be run in the same population. This allows greater comparability between studies.
- The cohort offers greater focus for research surrounding obesity and weight management, particularly their association to long-term health conditions. No other cohorts offer the detail required to inform policy decisions in this area.
- Analysis of this data set has shown that the health-related quality of life measure EQ-5D to be a useful tool for capturing differences in health and well-being after accounting for various long-standing health conditions.

## **Why was the cohort set up?**

The study began in 2009 as the South Yorkshire Cohort with the aims of collecting information on the health of the residents of South Yorkshire and providing a research platform for the region. In 2014, the study was expanded and rebranded as the Yorkshire Health Study to collect information for the whole of the Yorkshire and Humberside region. A wide range of variables were collected to provide information on demographic characteristics, current and long-standing health, well-being, health care usage and health-related behaviours. The original focus of the cohort was weight and weight management, particularly the relationships between obesity, health care and long-standing health conditions. This focus was based on the recommendations of the 'Healthy Weight, Healthy Lives' government report which had proposed five key aims for future research to tackle obesity levels in England (1):

- A longitudinal approach to research (incorporating the short, medium and long term).
- Exploring the association of obesity to other key health issues.
- Analysing the impact of policies, with a focus on reducing social inequalities.
- Evaluating the effectiveness of long term policies and interventions.
- The targeting of population groups and incorporation of 'natural experiments'.

The Yorkshire Health Study allows the investigation of each of these aims through a focused set of variables and a large sample size (2). The depth of information provided allows a detailed investigation of variations in obesity to other dimensions of health, particularly chronic health conditions. Other population cohort studies are less relevant to fulfilling these identified needs of the National Health Service (NHS) for obesity-related research (for example (3–5)).

The cohort is uniquely set up to facilitate the ‘cohort multiple randomised controlled trial’ approach to experimental research (6). This approach utilises the large cohort study as a research platform to facilitate recruitment to multiple pragmatic randomised controlled trials (with treatment as usual as the control), as well as other health studies. This approach provides a number of benefits to studies utilising the Yorkshire Health Study research platform: detailed information on potential research participants and their efficient (quick and inexpensive) identification and ongoing information as to the natural history of the condition and treatment as usual (2). The particular benefits for randomised controlled trials include long term outcomes as standard, increased comparability between each trial conducted within the cohort and increased efficiency particularly for expensive or high risk interventions (6).

The study completed its first wave of data collection (2010 to 2012) and the second wave (2013 to 2015) to extend the data longitudinally is ongoing. The study is expected to run for a minimum of ten years. New variables will be added in future waves in response to the findings of research and the recommendations of health and social care commissioning bodies.

Ethical approval for the cohort was obtained from the Leeds East NHS Research Ethics Committee on the 22<sup>th</sup> April 2010 (ref: 09/H1306/97).

## **Who is in it?**

### Sample region

In the first wave, the sample region was South Yorkshire, an English metropolitan county in the Yorkshire and Humber region in England. It consists of four large urban areas (the

metropolitan boroughs of Barnsley, Doncaster and Rotherham, along with the city of Sheffield) and surrounding rural areas, with a total population of 1,343,601 recorded at the 2011 Census. Data were also collected for a small number of residents of Derbyshire. For subsequent waves, the sample region was expanded to the Yorkshire and Humberside 'Government Office Region' with a total population of 5,283,733 (2011 Census). Previous research has shown that in much of South Yorkshire and Yorkshire as a whole, rates of obesity are higher than compared to the national average (7). As such, the sample regions provide a useful setting for a cohort study that has a weight and weight management focus.

### Sampling strategy

A two stage approach was used for the initial data collection (2). Firstly, General Practitioner (GP) surgeries were contacted with a view to participating in the study (43 agreed; 50% acceptance). Consenting GP surgeries then mailed letters of invitation for all patients aged between 16 and 85. Included with the letter of invitation was an eight page questionnaire for data collection (see 2). An online version of the questionnaire was also available. Of the 156,866 questionnaires sent out, 27,806 were returned (a response rate of 15.9%). Of the 27,806 participants who returned questionnaires, 22,179 (81.7%) gave consent to being contacted again and 22,150 (79.7%) gave consent for researchers to access their health records. As from 2014, a 'citizen recruitment' strategy is being used to increase coverage and sample size. Information about the study is promulgated through a regional media campaign, and all residents in the Yorkshire and Humber region are invited to participate via a questionnaire and online website (<http://www.yorkshirehealthstudy.org/>). Data will also be collected from those who agreed to be contacted again in the first wave to extend the original dataset longitudinally.

### Cohort characteristics

Table 1 presents the baseline characteristics of participants in the first wave of the cohort by age, sex and deprivation quintile (using the Indices of Deprivation 2010 measure (8)) based on the residential postcode of the participant. For comparison, the category proportions for the South Yorkshire population (based on the 2011 Census) are also included.

(Table 1 here)

Participants in the cohort are older than in the total South Yorkshire population and the proportion of females in the sample is also somewhat higher. Participants from the lowest two quintiles of deprivation are over represented, and participants from the highest three areas are under-represented. As such, there is slight bias in the data across these three variables in the cohort. The majority of participants also reported being of White ethnicity (94.1%), which was over-representative of the ethnic group (2011 Census; 90.5%).

### **What has been measured?**

Table 2 summarises the variables collected. The inclusion of 12 long-standing conditions (along with a free-text field for additional conditions) reflects the growing interest in the impact of multi-morbidity (9,10). The EuroQoL EQ-5D health-related quality of life questionnaire (11) is included as part of the cohort questionnaire. This measure is used within the NHS and is the National Institute for Health and Care Excellence's preferred measure of quality of life in economic evaluations (2,12). In order to understand more about NHS and non NHS healthcare resource use, medication usage is also included.

(Table 2 here)

The postcode of the individual allows each record to be geo-located. From this, additional spatial information can be gathered (for example, see Table 1; the deprivation measure was collected from area data), allowing the analysis of phenomenon not included in the questionnaire (e.g. exposure to pollution). GP surgery information can also be linked to participants, providing information relating to the characteristics of the primary healthcare provider used by each individual.

Included in the questionnaire is an option for individuals to consent to allow researchers to access their health records. Data linkage of individual records in the cohort to GP clinical records (e.g. morbidity, health care and medication usage) and other sources (e.g. Hospital Episode Statistics) is possible and discussions are currently ongoing.

The second wave of data collection will collect the variables in the first wave (Table 2) excluding occupation, education or weight management. It will additionally include data on levels of income, provision (and receipt) of informal and formal care of individuals, dietary intake, severity of long-term health conditions, and more detail on physical activity.

## **What has it found?**

With the first wave of cohort data only recently collected, research is at an early stage. The focus so far has been on understanding differences in weight management strategies and exploring the associated factors to health-related quality of life. Some early results from these studies are presented here.

The cohort has been used to help inform the local councils' policy development, planning and resource allocation. For example, a report to inform the 'Joint Strategic Needs Assessment' of health and well-being for Sheffield City Council showed evidence of wide social inequalities in levels of obesity (13). There were twice as many people classified as obese in the most deprived areas than compared to the least deprived (with the relative difference in the morbidly obese being even greater). The cohort has also been used to provide estimates on the prevalence of long term health conditions and health-related behaviours.

A second study explored the uptake of slimming clubs and weight loss medication (manuscript submitted). There is little evidence surrounding how these weight management strategies varied by demographic factors, particularly deprivation. The study found varying social gradients in the uptake of each strategy independent of age, gender and health (Figures 1 and 2), with slimming clubs less likely to be used in deprived areas but weight loss medication more likely to be used (although the effect size for the latter was lower). The results demonstrate the importance of targeting population groups with different strategies to effectively tackle obesity.

(Figures 1 and 2 here)

A number of studies have focused on the relationship between reported EQ-5D and long-standing health conditions (11). Firstly, a study examined whether the EQ-5D measure is too narrowly focused to reflect the impact of long-standing health conditions on subjective well-being (manuscript submitted). It was observed that long-standing health conditions did not have an independent and consistent impact on life satisfaction. Therefore, it was concluded that the EQ-5D measure is a useful instrument to reflect the impact of health conditions on life satisfaction and no additional dimensions need to be added.

A second study focused on the impact of long-standing health conditions on reported EQ-5D (14). Previous studies have shown an inconsistent relationship between chronic ill-health and quality of life measures, with the suggestion that this may be due to differences between the studies in the method used to measure quality of life. The analysis showed that the long term conditions pain, depression, anxiety/nerves, and osteoarthritis had the largest negative effect on health-related quality of life (although all long term health conditions measured reported negative effects, independent of socio-demographic factors). These results were consistent across each statistical methodology employed in the analysis.

Researchers have also used the data to assess the relationship between BMI and EQ-5D, and whether the relationship is attenuated by long term health conditions (15). A deviation from a healthy weight (i.e. the 'normal' BMI category) was found to have a negative effect on health-related quality of life (EQ-5D) independent of social, demographic and health factors.

Removing each long term health condition from the analysis (compared to a single variable for long term condition) allowed an assessment to be made regarding the effect it had on the relationship-association between BMI and EQ-5D for individuals with a BMI  $\geq 25$  kg/m<sup>2</sup> (Figure 3). Diabetes, heart disease, osteoarthritis and high blood pressure were important mediating factors in the relationship and the results were consistent across each dimension of EQ-5D.

(Figure 3 here)

A number of research groups have used the platform to identify and invite potential participants to their studies. Researchers investigated how people make decisions concerning weight management (16). Decisions about managing an individual's weight were found to be impacted by stigma surrounding obesity, as negative feelings affected clear thinking in the choices individuals made. Drawing on their findings, a decision aid booklet was created to

help influence individuals to make good decisions about their weight control. The results also led to the development of a questionnaire to assess the impact of obesity stigma (17).

The first randomised controlled trial utilising the cohort multiple randomised controlled trial design is now being implemented (6). This trial is testing the effectiveness of treatment by homeopaths for people with self-reported depression (DEPSY) (<http://www.controlled-trials.com/ISRCTN02484593>).

A list of current projects can be viewed at <http://www.yorkshirehealthstudy.org/#!your-research-studies/c17r5>.

## **What are the main strengths and weaknesses?**

Strengths:

- A large sample size covering a wide range of population sub-groups.
- Longitudinal design for data collection.
- Data collected on a broad range of health related variables.
- A ‘research platform’ designed to facilitate the running of ‘cohort multiple randomised controlled trials’ and facilitate participant identification and recruitment to health related research (6).
- Consent to link the majority of individuals to their NHS health records.
- Questions designed to be policy relevant, to encourage the collaboration between academia, local authorities and the NHS.

Weaknesses:

- Some demographic bias with a higher proportion of elderly, female, White ethnicities and individuals from less deprived areas. This was partly due to the low response rate.
- The first wave was collected over three years, rather than at a single time point.

## **Can I get hold of the data? Where can I find more?**

The data are managed by the University of Sheffield's Clinical Trials Research Unit.

Anonymised data and details regarding using the resource for recruiting participants to studies can be gathered by contacting Dr Clare Relton ([c.relton@sheffield.ac.uk](mailto:c.relton@sheffield.ac.uk)). Multi-disciplinary collaboration is strongly encouraged.

## **Funding**

This work was supported by the 'National Institute for Health Research' (NIHR)

'Collaborations for Leadership in Applied Health Research and Care' (CLAHRC) for South Yorkshire (Obesity Theme) and the University of Sheffield.

## **Acknowledgements**

We are grateful to all the individuals who have enrolled in the cohort. We also acknowledge the GP practice staff for their contribution in the recruiting process. This publication is the work of the authors and does not necessarily reflect the views of the Yorkshire Health Study Management Team or Steering Committee.

This paper presents independent research by the NIHR CLAHRC for South Yorkshire. The views and opinions expressed are those of the authors, and not necessarily those of the NHS, the NIHR or the Department of Health. CLAHRC South Yorkshire would also like to acknowledge the participation and resources of our partner organisations. Further details can be found at [www.clahrc-sy.nihr.ac.uk](http://www.clahrc-sy.nihr.ac.uk).

Thanks to Professor Jon Nichols for advice on an early manuscript. We would also like to thank the Clinical Trials Research Unit at the University of Sheffield, particularly Sheila Tarleton and Tim Chater, for data entry and management. Also thanks to Barnsley Hospital research governance team (Mike Bramall and Jamie Matthews) for ethics and governance support.

**Conflict of interest:** None declared.

## References

1. Department of Health. Healthy weight, healthy lives: A cross-government research and surveillance plan. London: DoH; 2008.
2. Relton C, Bissell P, Smith C, Blackburn J, Cooper CL, Nicholl J, et al. South Yorkshire Cohort: a “cohort trials facility” study of health and weight - protocol for the recruitment phase. *BMC Public Health*. 2011;**11**:640.
3. Craig R, Mindell J, (eds). Health Survey for England 2011. London: NatCen; 2012. <https://catalogue.ic.nhs.uk/publications/public-health/surveys/heal-surv-eng-2011/HSE2011-All-Chapters.pdf> (accessed 27 Jan 2014).
4. Medical Research Council. National Survey for Health and Development. 2013. <http://www.nshd.mrc.ac.uk/default.aspx> (accessed 27 Jan 2014).
5. Raynor P. Born in Bradford, a cohort study of babies born in Bradford, and their parents: protocol for the recruitment phase. *BMC Public Health*. 2008;**8**:327.

6. Relton C, Torgerson D, O’Cathain A, Nicholl J. Research Methods & Reporting: Rethinking pragmatic randomised controlled trials: introducing the “cohort multiple randomised controlled trial” design. *BMJ*. 2010;**340**:c1066.
7. National Obesity Observatory. Local obesity prevalence. 2013. [http://www.noo.org.uk/LA/obesity\\_prev](http://www.noo.org.uk/LA/obesity_prev) (accessed 27 Jan 2014).
8. Department for Communities and Local Government. The English Indices of Deprivation 2010. London: DCLG; 2011. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/6871/1871208.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6871/1871208.pdf) (accessed 27 Jan 2014).
9. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet*. 2012;**380**:37–43.
10. Mangin D, Heath I, Jamouille M. Beyond diagnosis: rising to the multimorbidity challenge. *BMJ*. 2012;**344**:e3526.
11. Rabin R, Charro F. EQ-SD: a measure of health status from the EuroQol Group. *Ann Med*. 2001;**33**:337–43.
12. NICE. Guide to the methods of technology appraisal 2013. 2013. <http://www.nice.org.uk/media/D45/1E/GuideToMethodsTechnologyAppraisal2013.pdf> (accessed 27 Jan 2014).
13. Bogaert P, Relton C, Li J, Green MA, Strong M. Informing the Sheffield JSNA: Lifestyle behaviours and health inequalities in Sheffield. Sheffield: University of Sheffield; 2013.
14. Wu M, Brazier JE, Kearns B, Relton C, Smith C, Cooper CL. Examining the impact of 11 long-standing health conditions on health-related quality of life using the EQ-5D in a general population sample. *Eur J Heal Econ*. 2014;Advance Access. DOI:10.1007/s10198-013-0559-z.
15. Kearns B, Ara R, Young T, Relton C. Association between body mass index and health-related quality of life, and the impact of self-reported long-term conditions -- cross-sectional study from the south Yorkshire cohort dataset. *BMC Public Health*. 2013;**13**:1009.
16. Brown I, Deighton M. Obesity & Weight Loss Therapy A Decision Aid Intervention to Improve Decisions about Weight Management Referral in Primary Care: Development and Feasibility Study. *Obes Weight Loss Ther*. 2013;**3**:6.
17. Brown I, Flint S, Kilner K, Stride C. Understanding the impact of obesity stigma on weight control actions: development of the Thoughts about Body Size (TaBS) measure. Presented at The Obesity Society (TOS) International Conference. Atlanta. 2013.

## Tables

**Table 1.** Distribution of demographic characteristics of the sample and the total population.

Variable	Cohort (%)	Actual (%)
<b>Age</b>		
16-25	6.9	14.8
26-35	9.7	12.4
36-45	13	22.3
46-55	16.9	13.4
56-65	22.2	11.4
66-75	20.1	8.5
76+	9.9	7.0
Missing	1.2	
<b>Sex</b>		
Male	43.8	49.3
Female	56.2	51.7
<b>Indices of Deprivation 2010 Quintile (8)</b>		
Least deprived	14.5	9.3
Low deprivation	25.2	15.5
Average	16.6	18.1
High deprivation	18.3	22.3
Most deprived	25.1	34.8
Missing	0.3	

**Table 2:** Data collected in the first wave of the cohort.

Dimension	List of data available
Personal	Age, sex, ethnicity, number of children, height, weight, waist, body mass index, employment status, occupation, education, postcode.
Health	Long-standing conditions (tiredness/fatigue, pain, insomnia, anxiety/nerves, depression, diabetes, breathing problems, high blood pressure, heart disease, osteoarthritis, stroke, cancer), health-related quality of life (Euro-QoL-5D;(11)), medication (name, strength, type, prescription, what for).
Health care	Use of health services (hospital, General Practitioner, social/welfare, mental, alternative therapists).
Health behaviour	Smoking status, alcohol consumption, exercise (physical activity levels and type), concern over personal weight management, weight management strategy use (exercise, healthy eating, slimming clubs, weight loss medication, meal replacements).
Well-being	Life satisfaction.

## Figures

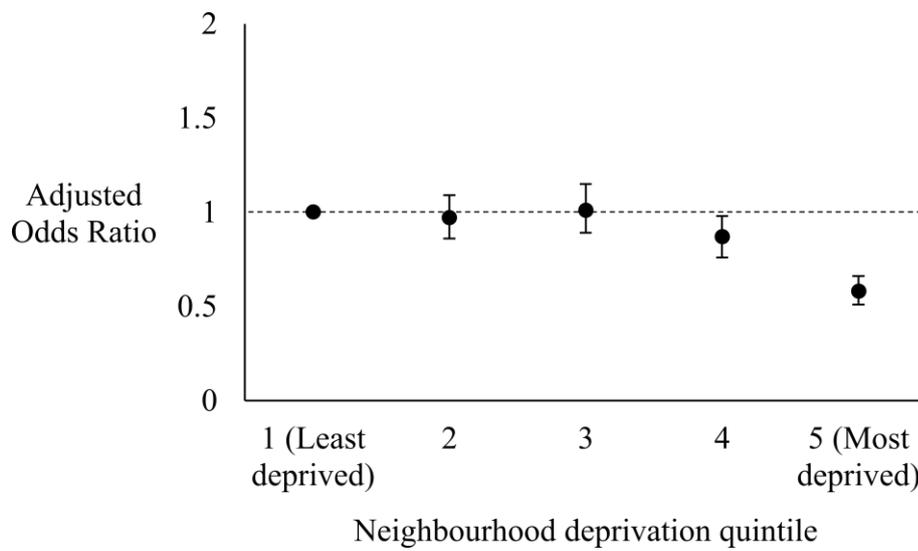


Figure 1: Adjusted odds ratios for neighbourhood deprivation quintile in analysing usage of slimming clubs (adjusted for age, sex and long-term health condition) (after manuscript submitted).

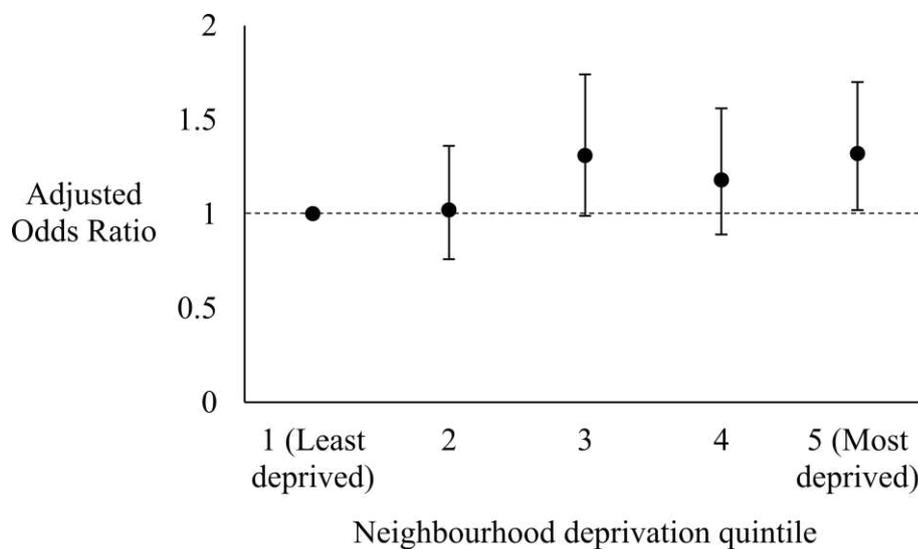


Figure 2: Adjusted odds ratios for neighbourhood deprivation quintile in analysing usage of weight loss medication (adjusted for age, sex and long-term health condition) (after manuscript submitted).

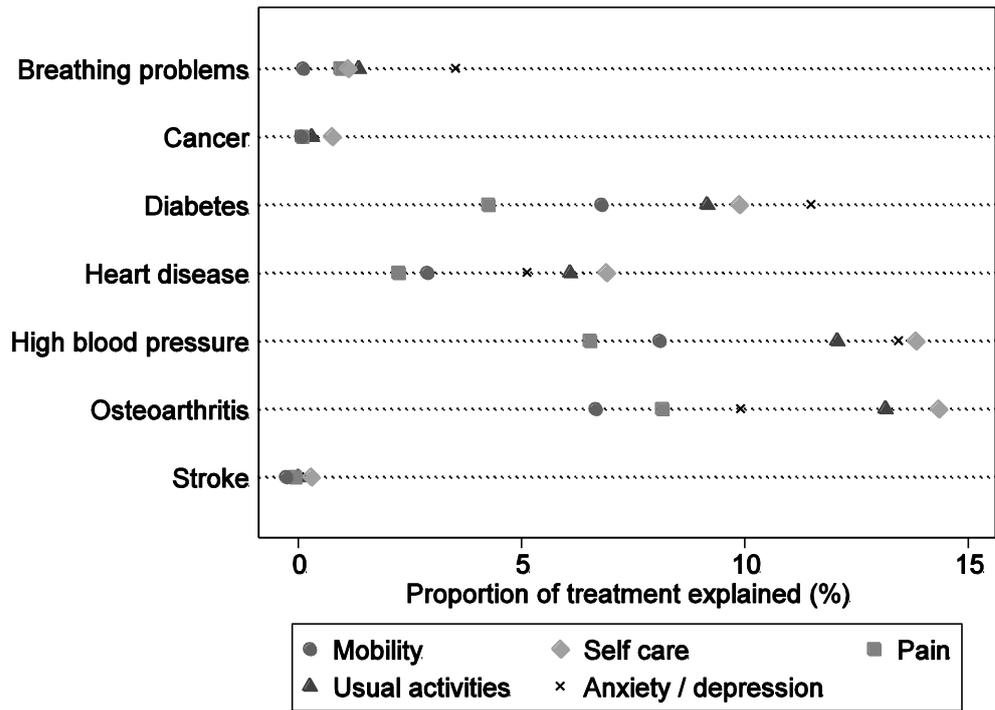


Figure 3: Percentage of the association between body mass index and EQ-5D ‘explained’ by removing each long term health condition in comparison to a single variable for long term condition for individuals with BMI  $\geq 25$  kg/m<sup>2</sup> (after (15)).