

Session 1: The Health Survey for England

Vanessa Higgins
ESDS Government



Today

Session 1: Introduce the data (Vanessa)

Coffee

Session 2: Performing standard single level analyses on the data in Stata and MLwiN, organising the HSE data for multilevel analysis in MLwiN and fitting single level models in MLwiN (Ian)

Lunch

Session 3: Introducing multilevel models (Mark)

Coffee

Session 4: Practical – normal response multilevel (Mark) analysis of the HSE using MLwiN



Housekeeping

- Toilets
- Loud continuous alarm, head back up the stairs and gather on grass outside
- Lunch & coffee breaks

This session

- The Health Survey for England
 - Overview
 - Why use it?
 - Registration and access
 - Resources to help users
- The HSE dataset were using
 - Introduction to the topic of obesity
 - Descriptive analyses of the data

ESDS overview

- National data archiving, dissemination and support service, established 2003 (for key economic and social data). Funded by ESRC
- Distributed service, bringing together centres of expertise
 - UK Data Archive (UKDA)
 - Centre for Census and Survey Research (CCSR)
 - Manchester Information & Associated Services (MIMAS)
 - Institute for Social and Economic Research (ISER)
- Access and User Support (helpdesk, user guides, workshops etc)

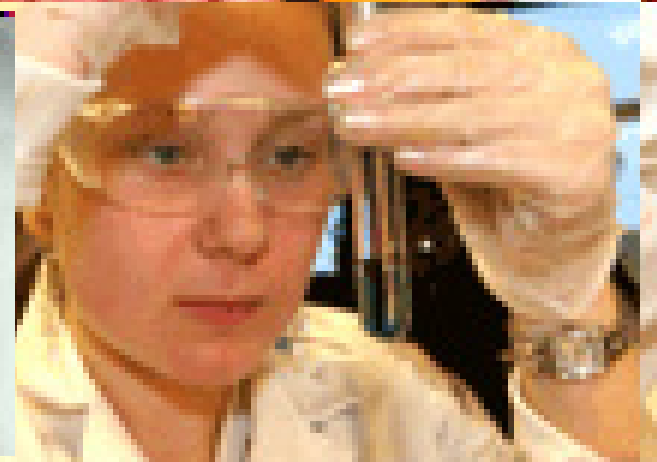
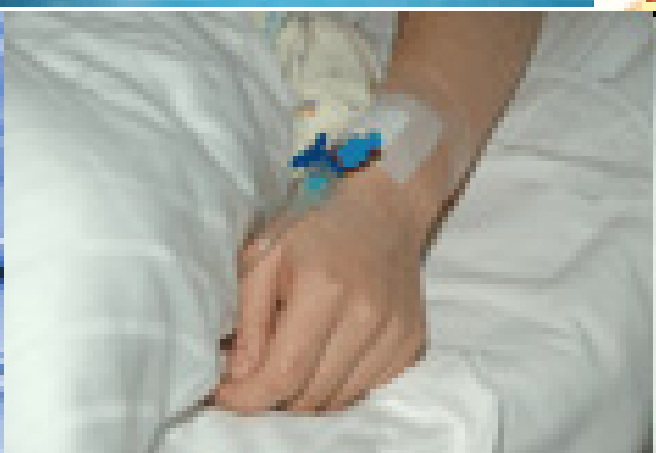


E-Stat

- Exciting new node in the ESRC NCeSS programme comprising:
 1. Statisticians researching statistical methodology and statistical software
 2. Computer science researchers
- Synthesis of statistical and computing expertise
- Enable cutting edge methodological developments to help quantitative researchers do new and better research



The Health Survey for England



The Health Survey for England

- Commissioned by The NHS Information Centre for health and social care and conducted by NatCen and UCL
- Key indicators for health (used by govt)
- Annual since 1991 (children since 1995)
- Cross-sectional: snapshot data
- Sample data
- Computer-assisted personal interviewing (CAPI) & CASI face-to-face interview followed by a nurse visit for a clinical examination



Economic and Social Data Service

The Health Survey for England

Topics:

Core questions each year plus topic modules

Sample size, design and questionnaire vary to reflect topic e.g.....

2003 cardiovascular disease

2004 ethnic minority groups

2005 older people

2006 cardiovascular disease

2007 knowledge and attitudes

2008 physical activity and fitness

Changes to Health and Social Care Survey in 2011



Tuesday, 13 June, 2000, 15:06 GMT 16:06 UK

Ethnic health inequalities revealed



Smoking is high among some ethnic groups

A major government survey has found high rates of disease among people from ethnic minority groups.

Ministers have responded by pledging to reduce high rates of smoking among ethnic groups - a factor they see as key to the problem.

The Health Survey for England found smoking rates among Bangladeshi

“
It is completely unacceptable that in

The English get fatter



The USA obesity problems are spreading to the UK

A fifth of English women and 17% of men are obese, according to a report for the Department of Health.

Official statistics, published on Wednesday, show the proportion of the population now classified as obese has risen by 4% since 1993.

In 1993, 13% of men and 16% of women were classified as obese.

The government says it is taking action to reverse the trend, but it does not expect an overnight solution.

Accessing the reports

- <http://www.ic.nhs.uk/statistics-and-data-collections/health-and-lifestyles-related-surveys/health-survey-for-england>
- **Health Survey for England - 2008: Physical activity and fitness**
- **Health Survey for England 2007: Healthy lifestyles: knowledge, attitudes and behaviour**
- **Health Survey for England 2006: CVD and risk factors adults, obesity and risk factors children**
- **Health Survey for England 2005: Health of Older People**
- **Health Survey for England 2004: Health of Ethnic Minorities - Full Report**



Why should you want to use the data for secondary analyses?

Because the data are.....

- Very cost effective: data free of charge to not for profit researchers
- Saves time: no need to conduct survey
- Access to high quality, well documented data
- Can provide nationally representative data - allows generalisation to population
- Allows historical and geographical comparisons to be made
- Look at sub-populations
- Using the flexibility of the data to look at alternative definitions
- Use the structure of the data to do multilevel modelling!!



Data access via ESDS/UK Data Archive



Economic and Social Data Service

Accessing data

- All users can access study descriptions, online documentation, including questionnaires, free of charge without registering with ESDS
- In order to access the datasets you need to register with ESDS
 - Register online using your UKFederation username and password
 - Simple online form, takes about 10 minutes
 - You need to register a usage of the data/project as part of this process
 - Non-commercial users: free of charge
 - Commercial users: £500 Charge for per study and will need to apply for UKFed username and password
 - You need to agree to the End User Licence when you register



The End User Licence

This is not public data!

We need to know who you are, how to contact you and what you are using the data for

Agree not to attempt to identify individuals

Only use the data for your stated purpose (you can re-register if you want to use the data for another use)

Do not pass the data to unregistered parties (that includes deleting the data before passing on PCs!)

Tell us if you publish using the data

Some more sensitive or detailed datasets require more stringent licensing procedures

- Special conditions
- Special licences



Special Licence

- A Special Licence (SL) is used for special licence data which poses a higher risk of disclosure
- A SL requires the signature(s) of the researcher(s) and their institution. It also needs the explicit permission of the data owner
- Surveys such as the Annual Population Survey and the Labour Force Survey include special licence versions
- We hope to deposit the HSE data we are using today under special licence



Obtaining data once registered

- Users then
 - Download the data to their local machine selecting their preferred format (SPSS, STATA < TAB etc)

Or

- Place an order for the data (Special Licence; commercial projects; special conditions if no online agreement) and complete all relevant forms

NB: Depositor permission: in most cases, ESDS has to request permission from the data owner to supply data for commercial purposes.



Finding data

- [Overview](#)
- [Catalogue search](#)
 - [Help on searching](#)
 - [About catalogue](#)
 - [Glossary of terms](#)
- [Variables search](#)
- [Browse by subject](#)
- Major studies**
 - [Major depositors](#)
 - [New releases](#)
 - [HASSET thesaurus](#)
 - [Other archives](#)

[About the data](#)

[Login](#)

NEW USERS



How do I find data?
How do I register?
What data are available?

More frequently asked questions...

Health Survey for England list of datasets

Users should obtain the data and documentation using the table below.

19 April 2010: Body Mass Index (BMI) variables

The children's BMI variables included in Health Survey for England datasets from 1995 onwards are now available. Users should note that the original variables bmicut, bmicut2 and bmicut3 are unreliable and should not be used. Further information is available from the Information Centre for Health and Social Care [Health Survey for England](#) web page.

Users are advised to visit the [Health Survey for England](#) web pages for support in using these data, additional resources, and news and events.

See [Exploring differences in health SPSS workbook for students](#)

SN	Study Description	Explore Online	Doc	Download / Order
6397	Health Survey for England, 2008			<input type="checkbox"/>
6112	Health Survey for England, 2007			<input type="checkbox"/>
5809	Health Survey for England, 2006			<input type="checkbox"/>
5675	Health Survey for England, 2005			<input type="checkbox"/>
5439	Health Survey for England, 2004	-		<input type="checkbox"/>
5098	Health Survey for England, 2003	-		<input type="checkbox"/>
4912	Health Survey for England, 2002			<input type="checkbox"/>
4628	Health Survey for England, 2001	-		<input type="checkbox"/>
4487	Health Survey for England, 2000			<input type="checkbox"/>
4365	Health Survey for England, 1999			<input type="checkbox"/>
4150	Health Survey for England, 1998			<input type="checkbox"/>
3979	Health Survey for England, 1997			<input type="checkbox"/>
3886	Health Survey for England, 1996	-		<input type="checkbox"/>
3796	Health Survey for England, 1995	-		<input type="checkbox"/>
3640	Health Survey for England, 1994	-		<input type="checkbox"/>
3316	Health Survey for England, 1993	-		<input type="checkbox"/>

Documentation

- Questionnaires
- Codebooks
- Information about derived variables
- Technical information on sampling, weighting etc.



4. Using the data

The 2002 data consists of one individual level file and one household level file:



HSE02ai.sav	18,398 records	contains data for all individuals in Household who gave a full interview. It contains information from the household questionnaire, main individual schedule, self-completions and the nurse visit (where one occurred).
-------------	----------------	--

HSE02ah.sav	29,184 records	contains data on household, and sex, age and marital status for all individuals in co-operating households.
-------------	----------------	---

4.1 Variables on the files

Each of the data files contain questionnaire variables (excluding variables used for administrative purposes) and derived variables. The variables included in the individual file are detailed in the **“List of Variables”** document in the data section of the documentation. This document is the best place to look at in order to plan your analysis. It includes:

- Major categories of variables (eg Accidents, Anthropometric measurements)
- Sub categories of variables (eg Attitudes to cycling, Major accidents within the Accidents category)
- Source of each variable (eg Individual questionnaire, Nurse visit, Derived variable etc.)

Once you have decided which variables to include in your analysis, you can look up details of the question wording using the interview section documentation (all variables on the data file are given by name in the copy of the interview schedules provided), or use the **“Derived Variables Specification”** document in the data section of the documentation for derived variables.

4.2 Weighting variables

In HSE 2002, the sample was boosted in order to obtain greater numbers of children, young adults (aged 16-24) and mothers of infants under 1. While children aged 0-15 and young adults aged 16-24 were sampled from all selected addresses, adults aged 25 and over were selected only at Sample II addresses (i.e. they were selected at only 9 out of the 38 addresses included within each postcode

Smoking (Aged 18+)

IF Age of Respondent = 18 to 24 THEN

BookChk

INTERVIEWER CHECK: *(name of respondent)* IS AGED *(age of respondent)*. RESPONDENT TO BE...:

- 1 Asked Smoking/Drinking questions
- 2 Given SELF-COMPLETION BOOKLET FOR YOUNG ADULTS

ENDIF

IF Age of respondent = 16 to 17 AND (is in joint session with Adult aged 25+ OR with adult aged 18-24 AND BookChk=1) THEN

YAIntro

INTERVIEWER: Prepare lilac or grey self-completion booklet for young adults by entering serial numbers. Check that you have the correct person number.

Press <1> and <Enter> to continue.

YAInt2

At this point, I would now like you to answer some questions by completing all of this booklet on your own. The questions cover smoking, drinking and attitudes to health. I will need to ask you a few more questions from the laptop in a little while, and I will ask you to close the booklet for a few minutes while I do this.

INTERVIEWER: Explain how to complete booklet and show example in booklet.

Press <1> and <Enter> to continue.

IF (Age of Respondent is 18 years or over) OR (BookChk = Asked) THEN

SmokEver

May I just check, have you ever smoked a cigarette, a cigar or a pipe?

- 1 Yes
- 2 No

IF SmokEver = Yes THEN

SmokeNow

Do you smoke cigarettes at all nowadays?

- 1 Yes
- 2 No

ENDIF

IF SmokeNow = Yes THEN

DlySmoke

About how many cigarettes a day do you usually smoke on weekdays?

INTERVIEWER: IF LESS THAN ONE A DAY, ENTER 0. IF RANGE GIVEN AND CAN'T ESTIMATE, ENTER MID POINT. IF RESPONDENT SMOKES ROLL UPS AND CANNOT GIVE NUMBER OF CIGARETTES, CODE 97.

Range: 0..97

Smoking

Adult Current Smokers

CIGDYAL: (D) Number of cigarettes smoke a day - inc. non-smokers

SPSS Syntax

```
IF cigwday>=0 & cigwend>=0 cigdial=((5*cigwday)+(2*cigwend))/7.
IF ANY(-9,cigwday,cigwend) cigdial=-9.
IF ANY(-8,cigwday,cigwend) cigdial=-8.
IF age<16 cigdial=-1.
RECODE cignow(-9,-8=COPY) (2=0) INTO cigdial.
RECODE smkevr(-9,-8=COPY) (2=0) INTO cigdial.
RECODE cigevr(-9,-8=COPY) (2=0) INTO cigdial.
VARIABLE LABELS cigdial "(D) Number of cigarettes smoke a day - inc. non-smokers".
```

Adults Cigarette Smoking General

CIGST1: (D) Cigarette Smoking Status - Never/Ex-reg/Ex-occ/Current

- 1 Never smoked cigarettes at all
- 2 Used to smoke cigarettes occasionally
- 3 Used to smoke cigarettes regularly
- 4 Current cigarette smoker

SPSS Syntax

```
IF any(2,cigevr,smkevr) cigst1=1.
RECODE cigreg (3=1) (2=2) (1=3) INTO cigst1.
IF cignow=1 cigst1=4.
IF ANY(-9,smkevr,cignow,cigevr,cigreg) cigst1=-9.
IF ANY(-8,smkevr,cignow,cigevr,cigreg) cigst1=-8.
IF smkevr=-1 cigst1=-1.
IF age<16 cigst1=-1.
VARIABLE LABELS cigst1 "(D) Cigarette Smoking Status - Never/Ex-reg/Ex-occ/Current".
VALUE LABELS cigst1
  1 "Never smoked cigarettes at all"
  2 "Used to smoke cigarettes occasionally"
  3 "Used to smoke cigarettes regularly"
  4 "Current cigarette smoker".
```

CIGSTA3: (D) Cigarette Smoking Status: Current/Ex-Reg/Never-Reg

- 1 Current cigarette smoker
- 2 Ex-regular cigarette smoker
- 3 Never regular cigarette smoker

ESDS Government resources

- [HSE home](#)
- [Datasets](#)
- [Resources](#)
- [Questionnaires](#)
- [Starting Analysis](#)
- [Citing this data](#)
- [Variables Search](#)
- [Links](#)
- [Registered Uses](#)
- [FAQ](#)
- [Other surveys](#)

Health Survey for England

The Health Survey for England (HSE) is a series of annual surveys about the health of people living in England. Since 1994 the survey has been carried out by the Joint Health Surveys Unit of the National Centre for Social Research and the Department of Epidemiology and Public Health, Royal Free and University College Medical School, London. The survey is sponsored by the Department of Health to provide better and more reliable information about various aspects of people's health and to monitor selected health targets.

The HSE began in 1991 and has been carried out annually since then. A number of core questions are included every year but each year's survey also has a particular focus on a disease or condition or population group. Topics are brought back at appropriate intervals in order to monitor change. The survey combines questionnaire-based answers with physical measurements and the analysis of blood samples. Blood pressure, height and weight, smoking, drinking and general health are covered every year. An interview with each eligible person in the household is followed by a nurse visit.

The 'core' includes: questions on general health and psycho-social indicators, smoking, alcohol, demographic and socio-economic indicators, questions about use of health services and prescribed medicines and measurements of height, weight and blood pressure. The modules may be about a single topic, several topics or about population groups. The modules to date have been:

- 1993 cardiovascular disease
- 1994 cardiovascular disease
- 1995 asthma, accidents and disability
- 1996 asthma, accidents and special measures of general health (Euroqol, SF36)
- 1997 children and young people
- 1998 cardiovascular disease
- 1999 ethnic groups
- 2000 older people and social exclusion
- 2001 respiratory disease and atopic conditions, disability and non-fatal accidents
- 2002 children and young people (aged 0-24)
- 2003 cardiovascular disease
- 2004 ethnic minority groups
- 2005 older people
- 2006 cardiovascular disease
- 2007 knowledge and attitudes
- 2008 physical activity and fitness

The early surveys, from 1991 to 1994, were confined to adults ages 16 and over but since 1995 HSE has also included children aged 2-15. In 2001, for the first time, this lower age limit was removed and the survey covers all ages. However, lower age limits are imposed for questions on certain topics. Information is obtained directly from persons aged 13 years and over. However, information about children aged under 13 years is obtained from a parent, with the child present.

Download data via UKDA

Variable search

Help to get started: FAQs and Starting analysis

ESDS Government resources

- Thematic guides (Health, Employment etc)
- Analysis guides (SPSS, Stata, Weighting, complex surveys, Change over time, Small area estimation)
- Keeping in touch:
Workshops/events/newsletter
- Helpdesk
- Teaching datasets
- Variable and publication searches



Variable search

Datasets for Health Survey for England

For each survey entry all variables are listed with further information giving the distribution of each variable and any additional information available. The Information and Download link take you to the ESDS Access pages for that survey.

Search variables
hearing hse
only

When searching you may wish to use common abbreviations for words (govt for government) if you don't get results you're expecting. This is because the survey documentation searched has used these abbreviations.

- Health Survey for England, 2008
- Health Survey for England, 2007
- Health Survey for England, 2006
- Health Survey for England, 2005
- Health Survey for England, 2004

- [Variables Information and Download](#)
- [Variables Information and Download](#)
- [Variables Information and Download](#)
- [Variables Information and Download](#)
- [Variables Information and Download](#)

ESDS Government Variable Database: Search Results

You searched for **hearing** in the ESDS Government variables database for the [Health Survey for England](#), not in the [variable catalogue](#). There are 123 results:

- [HSE comhear](#) (Are hisher[PNo] communication problems to do with hisher[PNo] hearing):
[2001](#), [2000](#), [1995](#).
- [HSE disaba3](#) (Cannot follow a TV programme at a volume others find acceptable (with hearing aid if normally worn):
[2005](#), [2001](#), [1995](#).
- [HSE disabae3](#) (Cannot follow a TV programme at a volume others find acceptable (WITH HEARING AID IF NORMALLY WORN):
[1995](#).
- [HSE dqtres04](#) (Reason quit: Hearing about new treatment):
[2008](#), [2007](#).
- [HSE hear](#) ((D) hearing disability):
[2001](#), [2000](#), [1995](#).

Health Survey for England, 2001 Variables

hear – (D) hearing disability

Value Label	Count	Percentage	Filename: hse01ai
-9.00 No answer/refused	2	n/a%	
-8.00 Don't know	5	n/a%	
-7.00 Refused/not obtained	0	n/a%	
-6.00 Schedule not obtained	0	n/a%	
-2.00 Schedule not applicable	0	n/a%	
-1.00 Item not applicable	2466	n/a%	
0.00 None	16367	95%	
1.00 Moderate	713	4%	
2.00 Severe	87	0%	

19640 cases , missing values -1, -99 .

Copyright: Crown Copyright material is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland

[Back to the variable list.](#)

Documentation references

- [a4628ucb.pdf](#) (page 78, 1 mention)

ESDS Contacts

Helpdesk:

govsurveys@esds.ac.uk

(0161) 275 1980

Website:

www.esds.ac.uk/government

www.esds.ac.uk

Jisc list:

esds-govsurveys@jiscmail.ac.uk



Economic and Social Data Service



BESITY

defusing the health time bomb

***Data from national surveys
can be used to monitor
prevalence and develop
appropriate public health
interventions***

BMI

- BMI=weight/height²
- BMI categorical measure

Category	BMI
Underweight	18.5 or less
Desirable	Over 18.5 to 25
Overweight	Over 25 to 30
Obese	Over 30

NB: child bmi is calculated differently

Obesity is multi-factorial

- Factors contributing to obesity:
 - Energy intake (diet)
 - Physical activity levels (exercise)
 - The environment & locality
 - Social class (poverty)
 - Disease and disability
 - Genetics
 - Ethnic origin
 - Psychology and behaviour
 - Birth weight and rate of infant weight gain
 - Parental obesity

Patterns of obesity

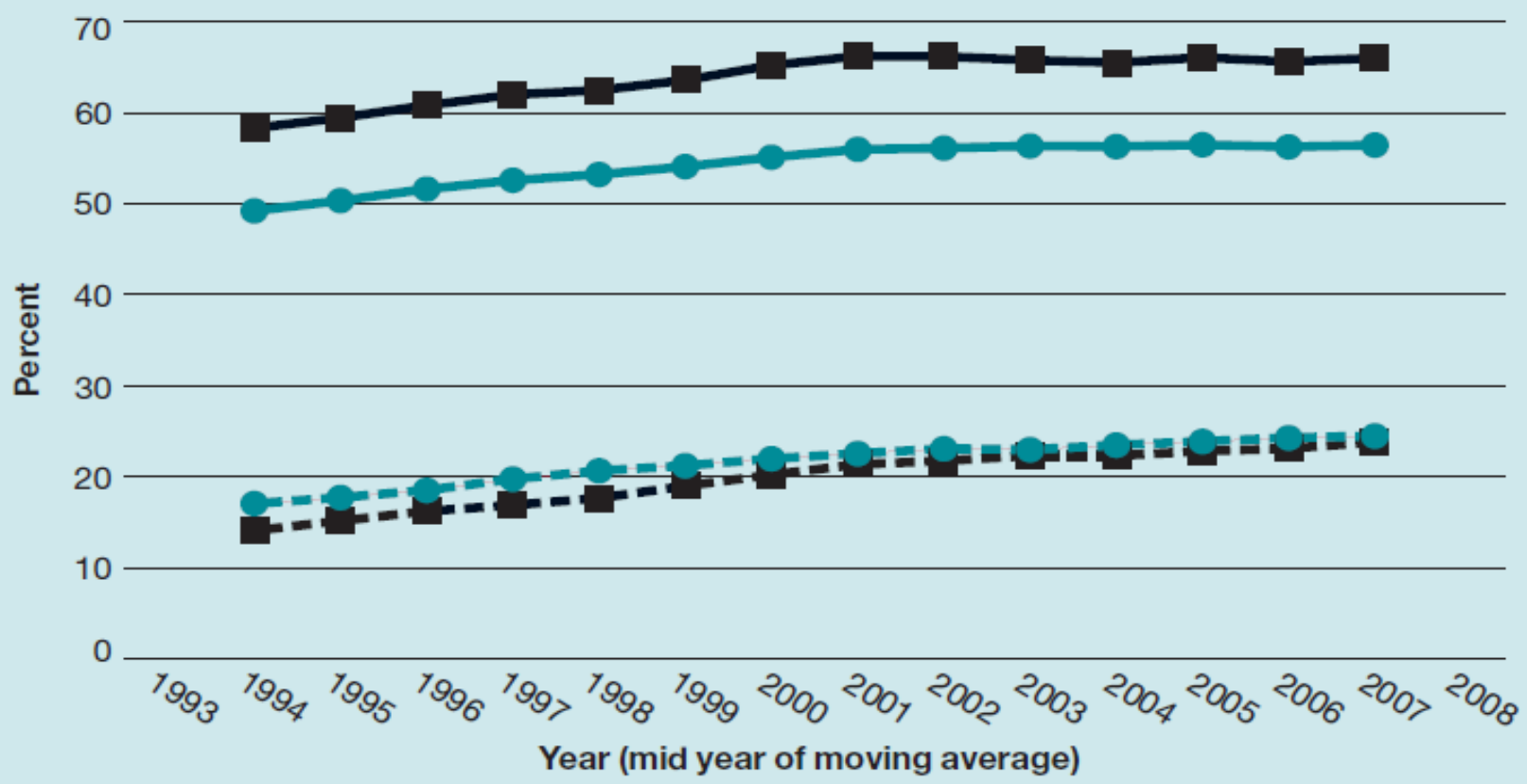
- Obesity is more prevalent for those with lower qualifications
- Obesity is more prevalent amongst those who are ill
- Obesity is more common in the North compared to the South
- Levels of obesity appear to be on the increase

Figure 7D

Overweight and obesity prevalence, 1993-2008, by sex (three year moving averages)

- Men obese
- Men overwt incl obese
- Women obese
- Women overwt incl obese

Base: Aged 16 and over with valid BMI measurement



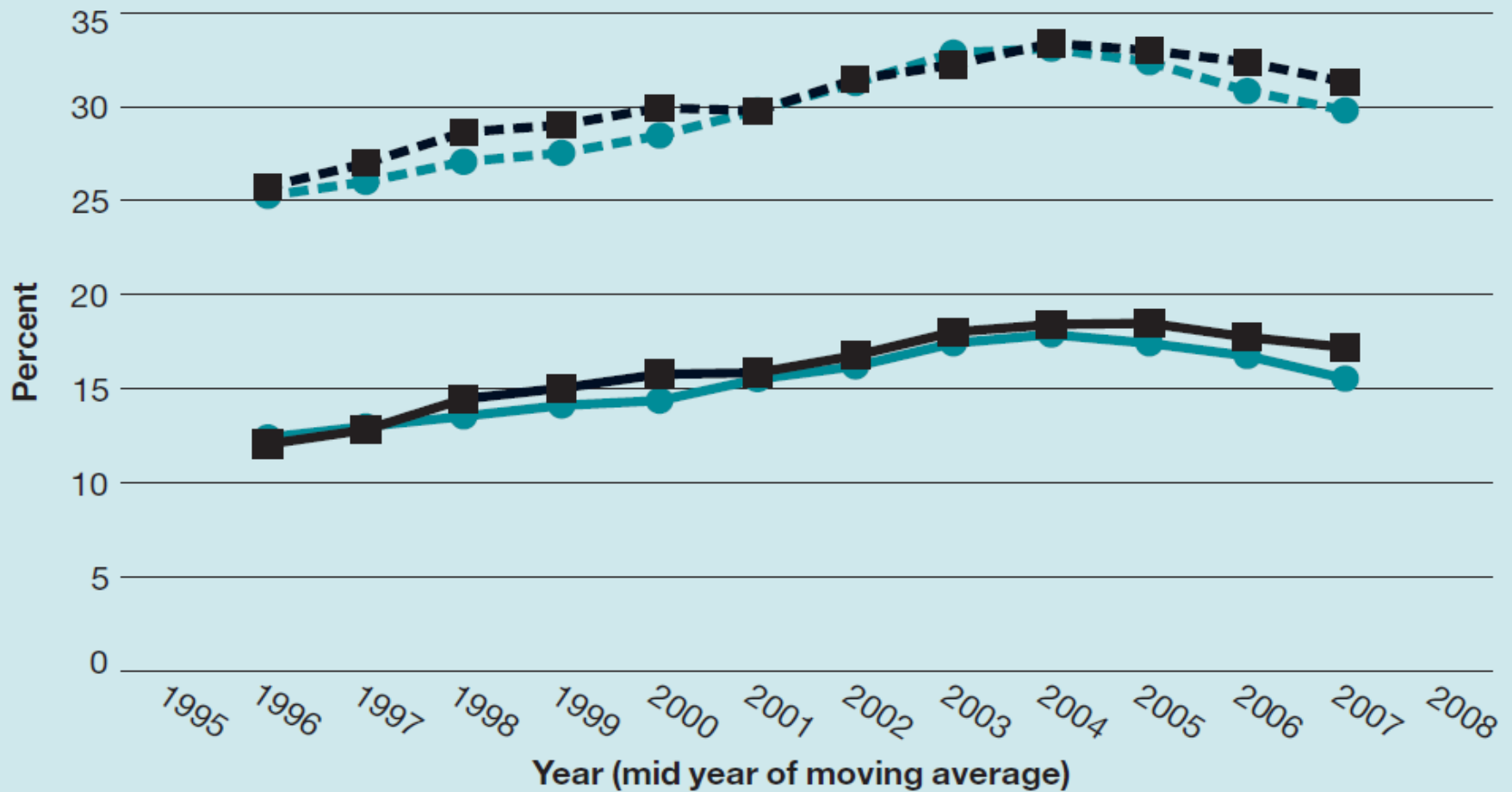
Note: Data from 1993 to 2002 are unweighted. Data from 2003 onwards are weighted for non-response.

Figure 13D

Overweight and obesity prevalence of children aged 2-15, 1995-2008, by age and sex (three year moving averages)

Base: aged 2-15 with valid BMI measurement

- Boys obese
- Boys overweight incl obese
- Girls obese
- Girls overweight incl obese



Geographies of obesity

Figure 4a

London : Men's BMI status

Base: Aged 16 and over with valid BMI

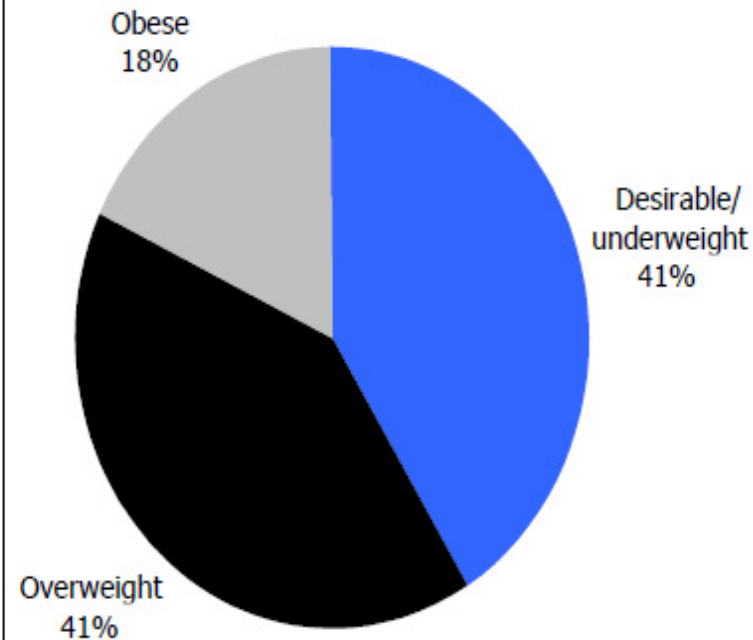
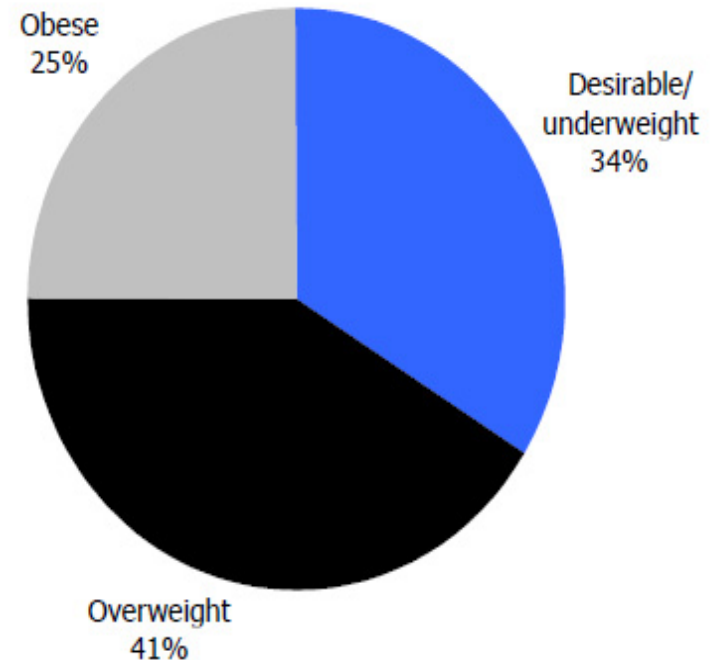


Figure 4b

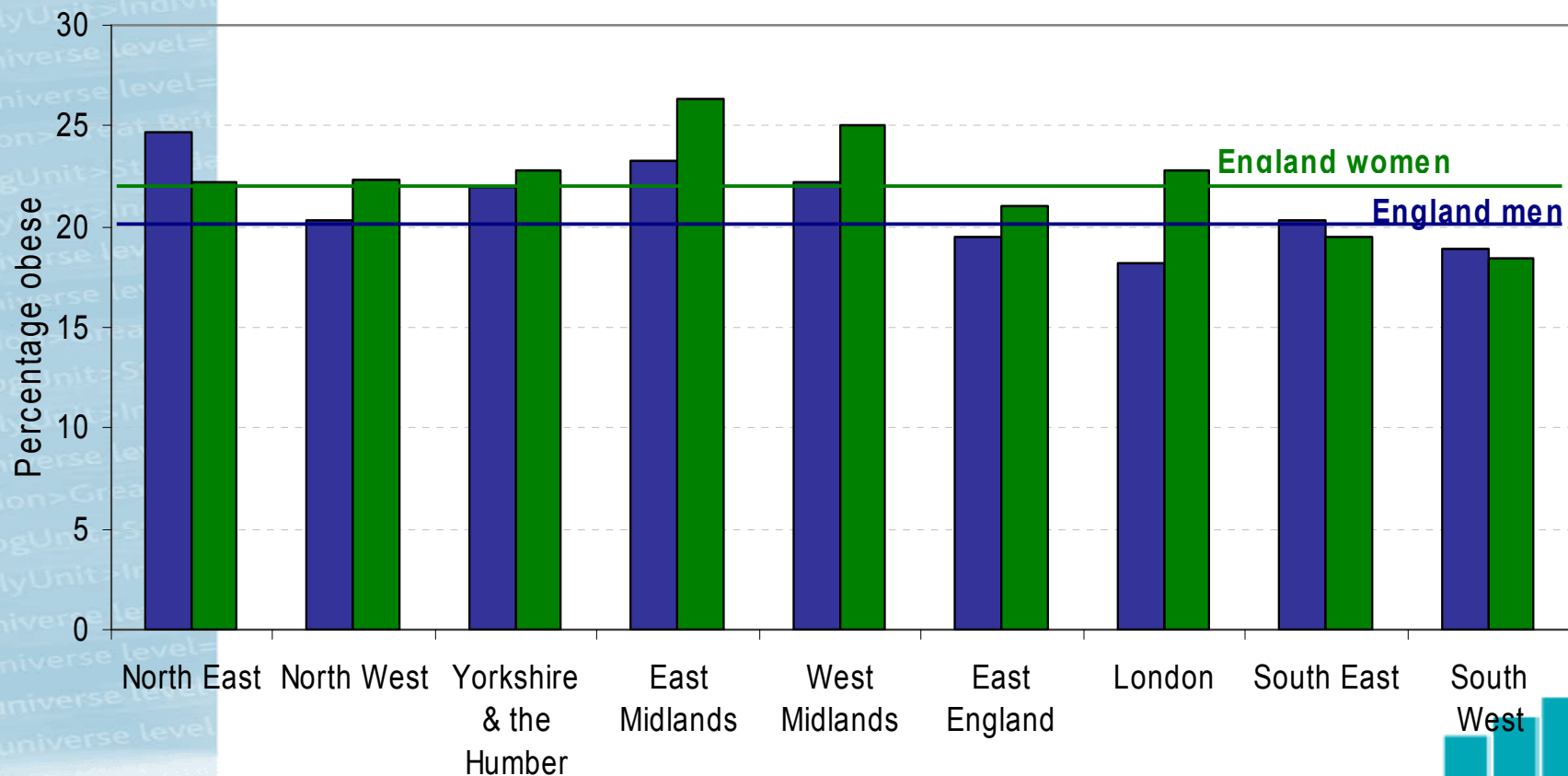
Yorkshire and the Humber: Men's BMI status:

Base: Aged 16 and over with valid BMI



Source: HSE 2003 taken from *Forecasting Obesity to 2010* (Zaninotto et al, 2006)

HSE: adult obesity by region



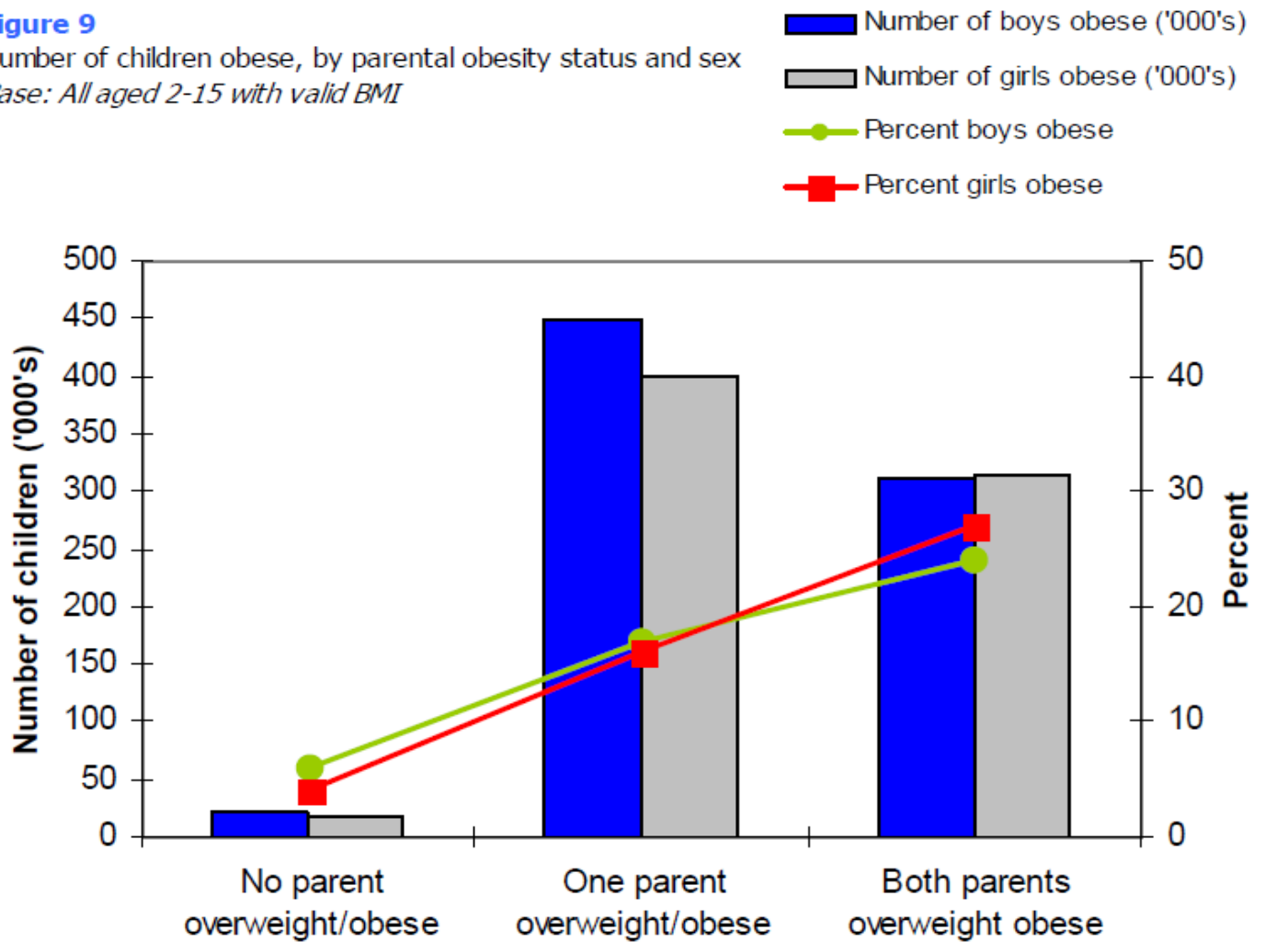
■ men ■ women
Source: HSE 2001-2

Does obesity run in families?

Figure 9

Number of children obese, by parental obesity status and sex

Base: All aged 2-15 with valid BMI



Source: HSE 2003 taken from *Forecasting Obesity to 2010* (Zaninotto et al, 2006)

Introduction to the dataset we're using today

Year	2003	2004	2005
Sample size	16,062	6,803	7,795

Note: boost samples in 2004 and 2005 are excluded



Sample design

Sampling frame – Postcode address file

PSU – Postcodes (half of a ward in 2004)

PSU's stratified by LA and within LA by the % of hh with a non-manual head of household (and by ethnicity in 2004)

Systematic sample of households drawn from each selected postcode

All adults in selected households interviewed. Up to 2 children aged 0-15 included in the survey

Types of variables in the workshop data

Variable	Example
BMI	'bmival' – valid measures of bmi
Socioeconomic	'nssec' – social class
Household	'tenureb' - tenure
Multilevel	'area' - postcode
Sampling	'wt_int' – sampling and non response weight

BMI variables

- Bmiok – Whether bmi measure is valid
- Bmi – body mass index including non-valid heights and weights
- Bmival – body mass index excluding non-valid heights and weights
- Height – persons height (cm)
- Weight – persons weight (kg)

Differences in obesity between men and women

	Men	Women
Desirable/under weight	35%	44%
Overweight	43%	33%
Obese	22%	23%

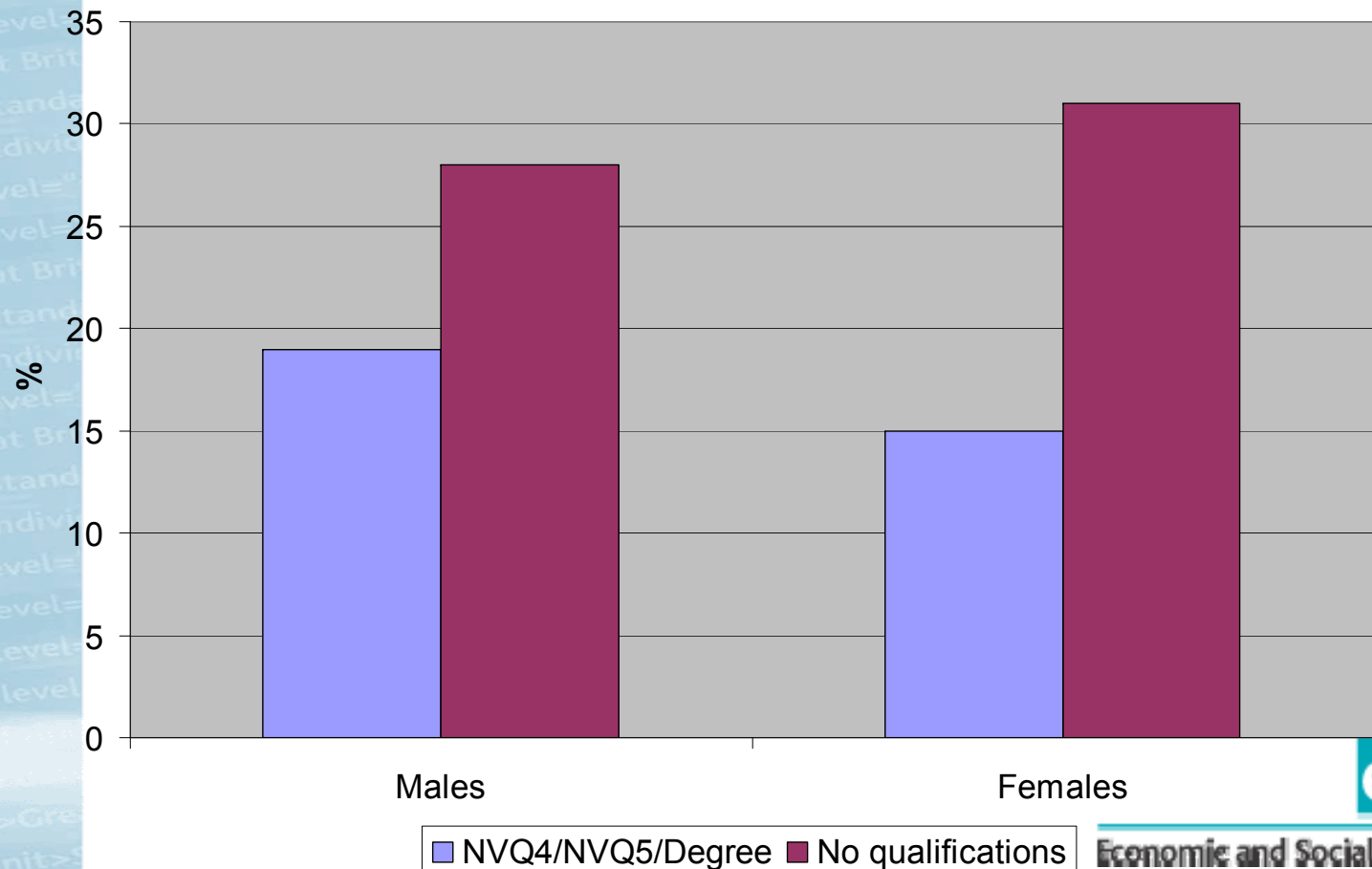
Source HSE 2003/05 - Adults aged 16+ with a valid BMI

Socioeconomic variables

- The data includes a number of variables that are known to be associated with bmi:
 1. NSSEC – social class
 2. Age
 3. Sex
 4. Limiting long term illness
 5. Ethnicity
 6. Qualifications
 7. Car ownership
 8. Marital status

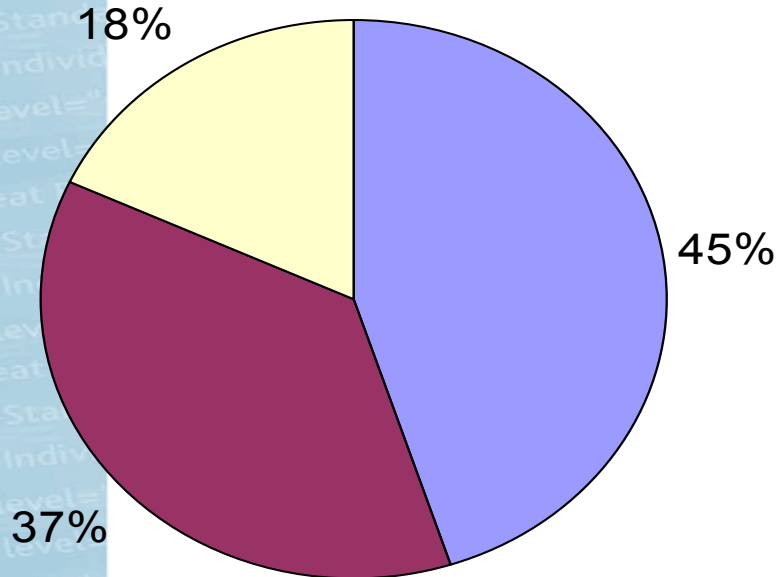
Differences between BMI according to qualifications

Obesity prevalence according to level of qualification (HSE 2003/05)

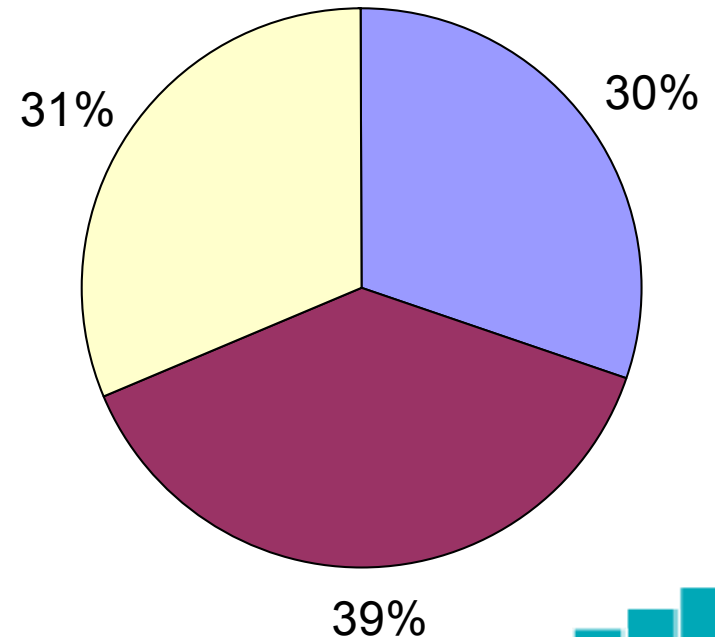


Differences in BMI status for the ill and well populations

BMI status – no illness



BMI status – limiting illness



■ Desirable or underweight ■ Overweight ■ Obese

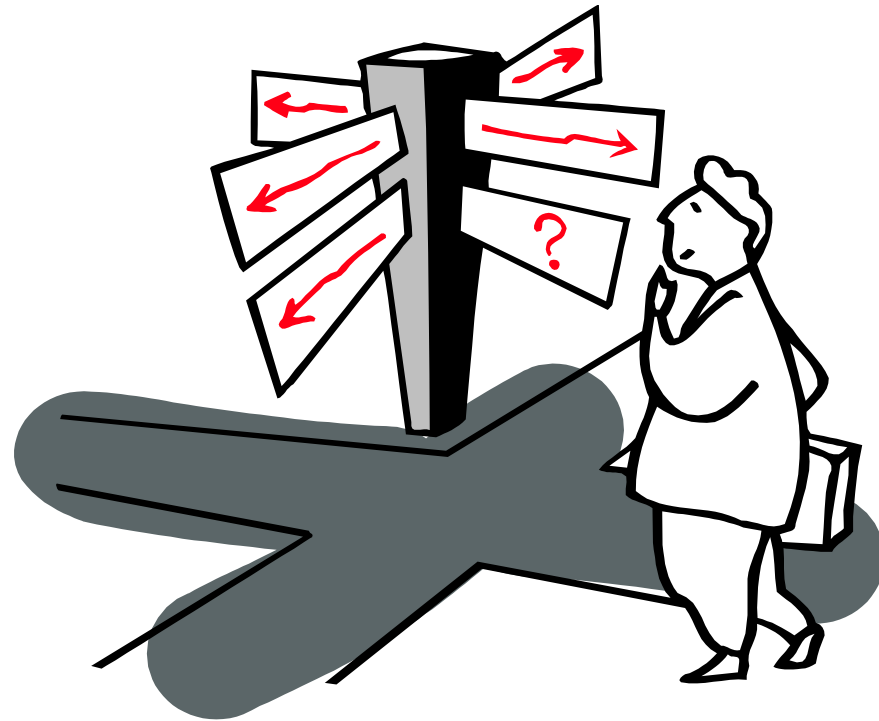
Household variables

- The models we will fit later include household as a level in multilevel models
 - A number of household variables are included in the data:
 - NSSEC of household reference person
 - Household size
 - Household type (e.g. small family)
 - Number of households at the address
-and a number of variables to define the geographical area in which the household falls

Sampling variables

- There are three variables that give information on sampling design:
 1. Wt_int (weighting variable)
 2. Area (PSU variable)
 3. Cluster (Stratification variable)
- Non response/disproportionate sampling
- Standard errors – complex survey design

Discussion / Questions



Thanks to Louisa Ellis from Durham University & colleagues from Teeside University for re-use of some of these slides. Go to www.noo.org.uk for latest obesity data.