Harmonisation of body size data

Will Johnson

MRC Human Nutrition Research
Maternal and Child Nutrition

closer

Promoting excellence in longitudinal research

Our areas of work

CLOSER works across four different areas in order to achieve its overarching objective to maximise the use, value and impact of cohort and longitudinal studies.

Click on an area of work below to learn more.

Data harmonisation

CLOSER is working to make the data from longitudinal studies more comparable, so it is easier to find out how things are changing from generation to generation. The first four work strands in this area are:

1. Harmonisation of biomedical measures
2. Harmonisation of socio-economic measures
3. Harmonisation of analysis of biological samples
4. Harmonisation of measures of vision

PLOS | MEDICINE

RESEARCH ARTICLE

How Has the Age-Related Process of Overweight or Obesity Development Changed over Time? Co-ordinated Analyses of Individual Participant Data from Five United Kingdom Birth Cohorts

William Johnson[1]*, Leah Li[2], Diana Kuh[1], Rebecca Hardy[1]

1 Medical Research Council Unit for Lifelong Health and Ageing at University College London, London, United Kingdom, 2 University College London Institute of Child Health, London, United Kingdom

* William.Johnson@mrc-hnr.cam.ac.uk

OPEN ACCESS

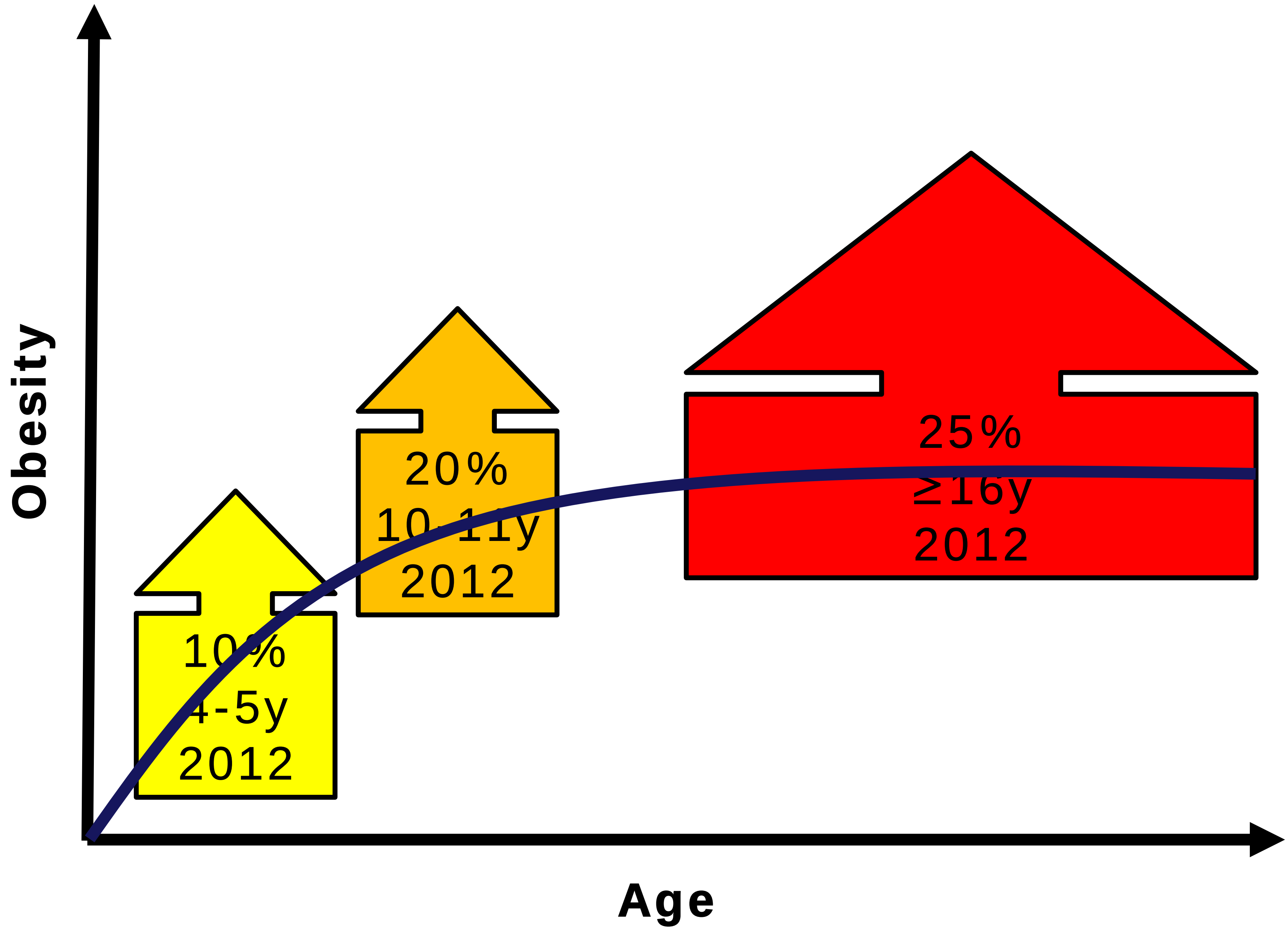
Citation: Johnson W, Li L, Kuh D, Hardy R (2015) How Has the Age-Related Process of Overweight or Obesity Development Changed over Time? Co-ordinated Analyses of Individual Participant Data from Five United Kingdom Birth Cohorts. PLoS Med 12(5): e1001828. doi:10.1371/journal.pmed.1001828

Editor: Richard Lehman, University of Oxford, UNITED KINGDOM

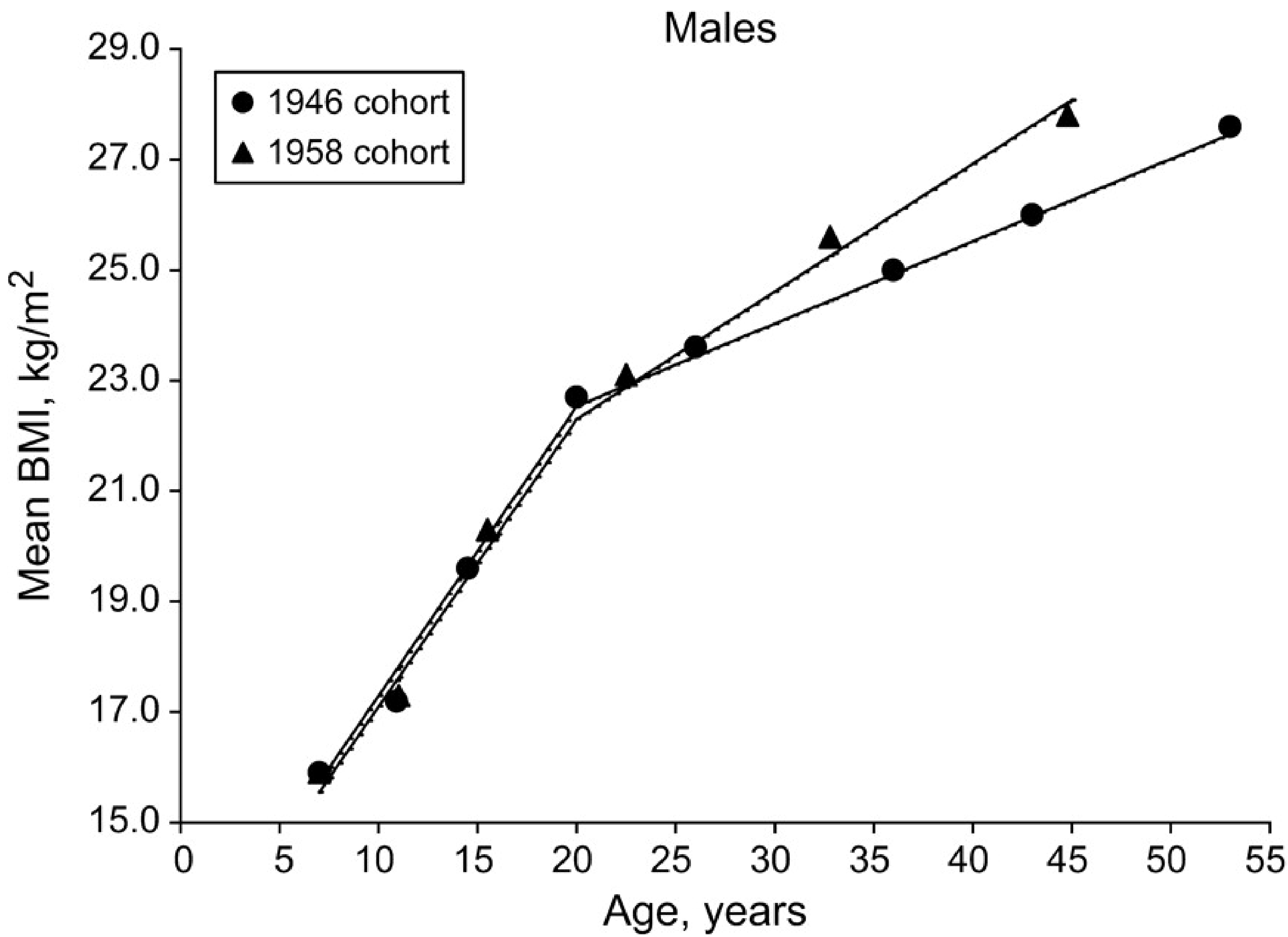
Abstract

Background

There is a paucity of information on secular trends in the age-related process by which people develop overweight or obesity. Utilizing longitudinal data in the United Kingdom birth cohort studies, we investigated shifts over the past nearly 70 years in the distribution of body mass index (BMI) and development of overweight or obesity across childhood and adulthood.

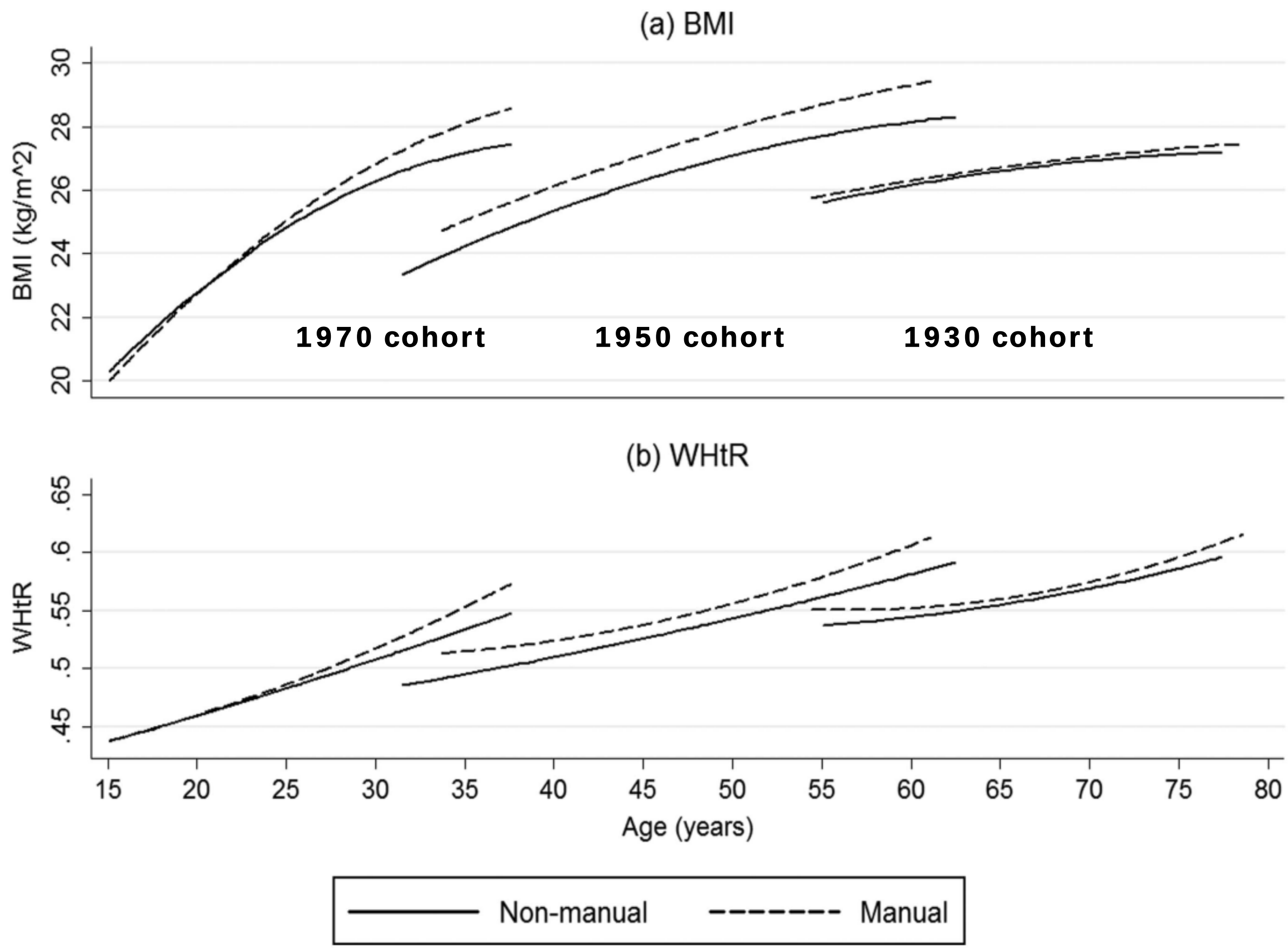

Obesity
10%
4-5y
2012
20%
10-11y
2012
25%
≥16y
2012
Age

One publication had investigated shifts over time in BMI trajectories using data from the UK birth cohort studies



Li L, Hardy R, Kuh D, Lo Conte R, Power C. Child-to-adult body mass index and height trajectories: a comparison of 2 British birth cohorts. Am J Epidemiol 2008;168(9):1008-15

Rest of our knowledge is based on studies in which data have been treated cross-sectionally or studies that are not representative and often span only a small part of the life course

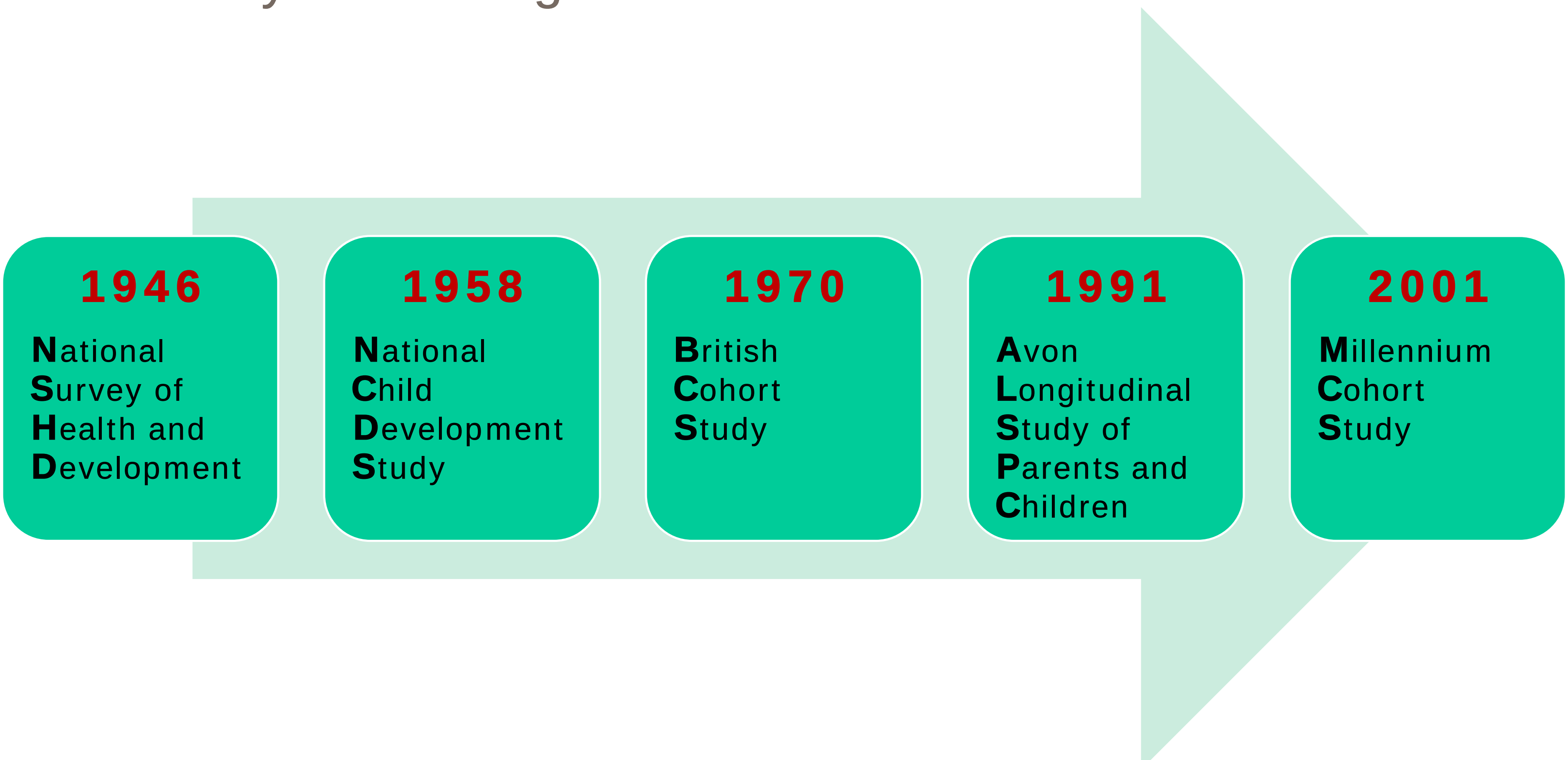


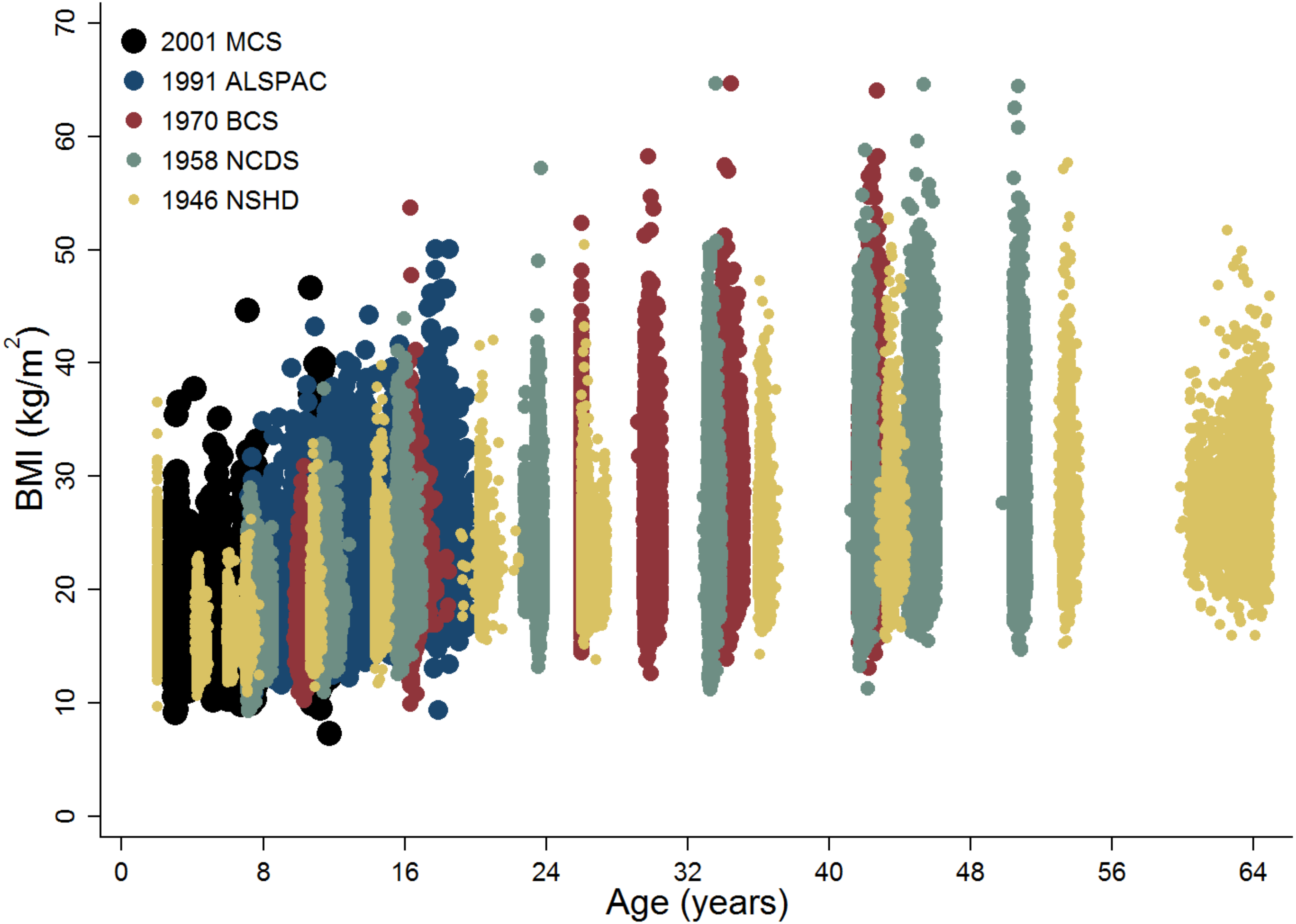
Shaw RJ, Green MJ, Popham F, Benzeval M. Differences in adiposity trajectories by birth cohort and childhood social class: evidence from cohorts born in the 1930s, 1950s and 1970s in the west of Scotland. J Epidemiol Community Health 2014;68(6):550-6

We aimed to utilise the extensive longitudinal BMI data in the UK birth cohort studies to describe

1) Shifts over time in the distribution of BMI across age

2) Shifts over time in the development of overweight or obesity across age

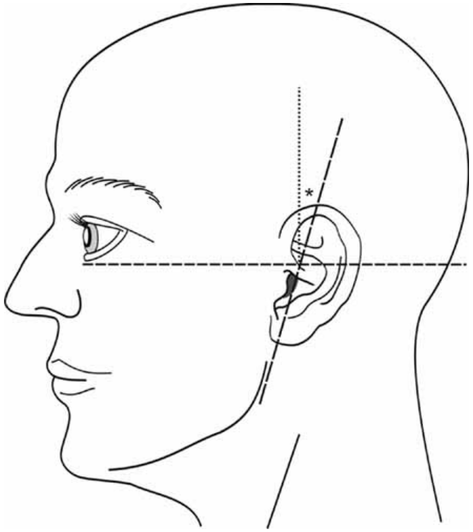




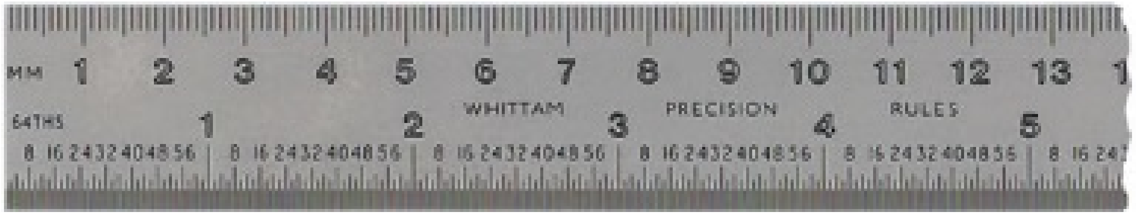
2001 MCS
1991 ALSPAC
1970 BCS
1958 NCDS
1946 NSHD
BMI (kg/m2)
Age (years)
0
10
20
30
40
50
60
70
0
8
16
24
32
40
48
56
64

What are the potential targets for harmonisation?

Measured vs. self-reported data



Measurement protocol differences (including rounding & clothing)

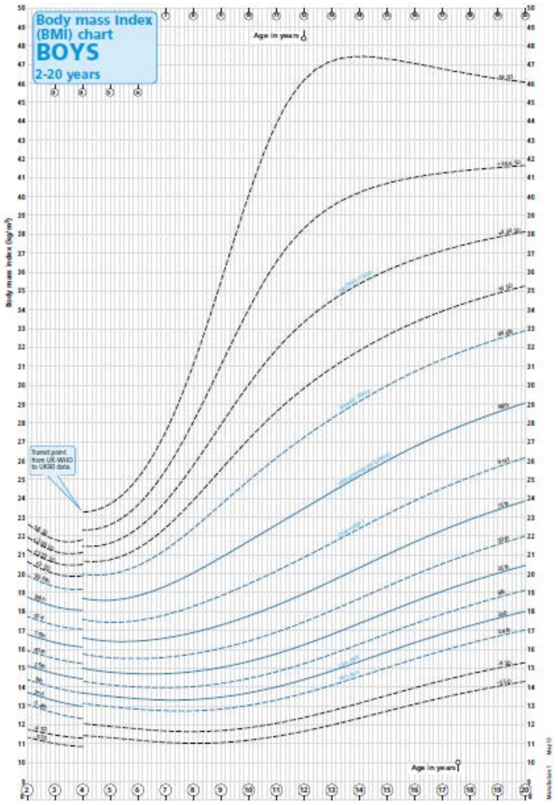


Precision of instrument (including metric vs. imperial)

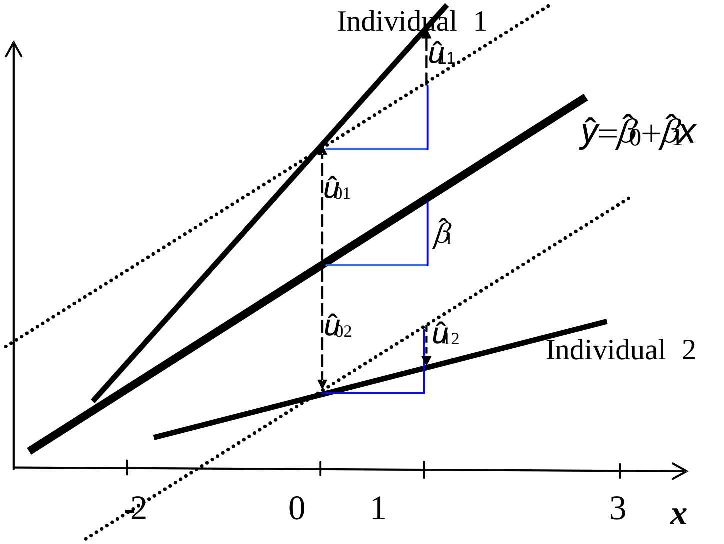
Weights and heights were converted to kg and m.
Measured data were augmented self-reported data at the same age to maximise the amount of available information and to retrieve information from the upper end of the distribution that appeared to have been removed by the employment of a cut-off during data entry or cleaning
Missing observations of adulthood height were filled in with observations of height from previous adulthood sweeps.
Decimal age at assessment variables were computed from existing age variables or as the difference between date of birth and date of assessment (for sweeps that were missing a date or some component of a date variable: day, month, and/ or year was assigned to the whole cohort). Participants who were still missing decimal age were assigned the mean value for that cohort at that sweep.
Measurements taken while a woman was pregnant were excluded, where possible.
A standardised data cleaning protocol was applied. This involved removal of biologically implausible values using sensible yet arbitrary cut-offs (e.g., weight > 250 kg and height > 3 m) and inspection of a connected scatter plot of serial weight or height against age (i.e., a trajectory) for persons with a measurement or change in measurement between two consecutive ages greater than five standard deviations from the sex and study stratified mean.

Sex and study stratified analyses

1. LMS method used to describe the distribution of BMI across age

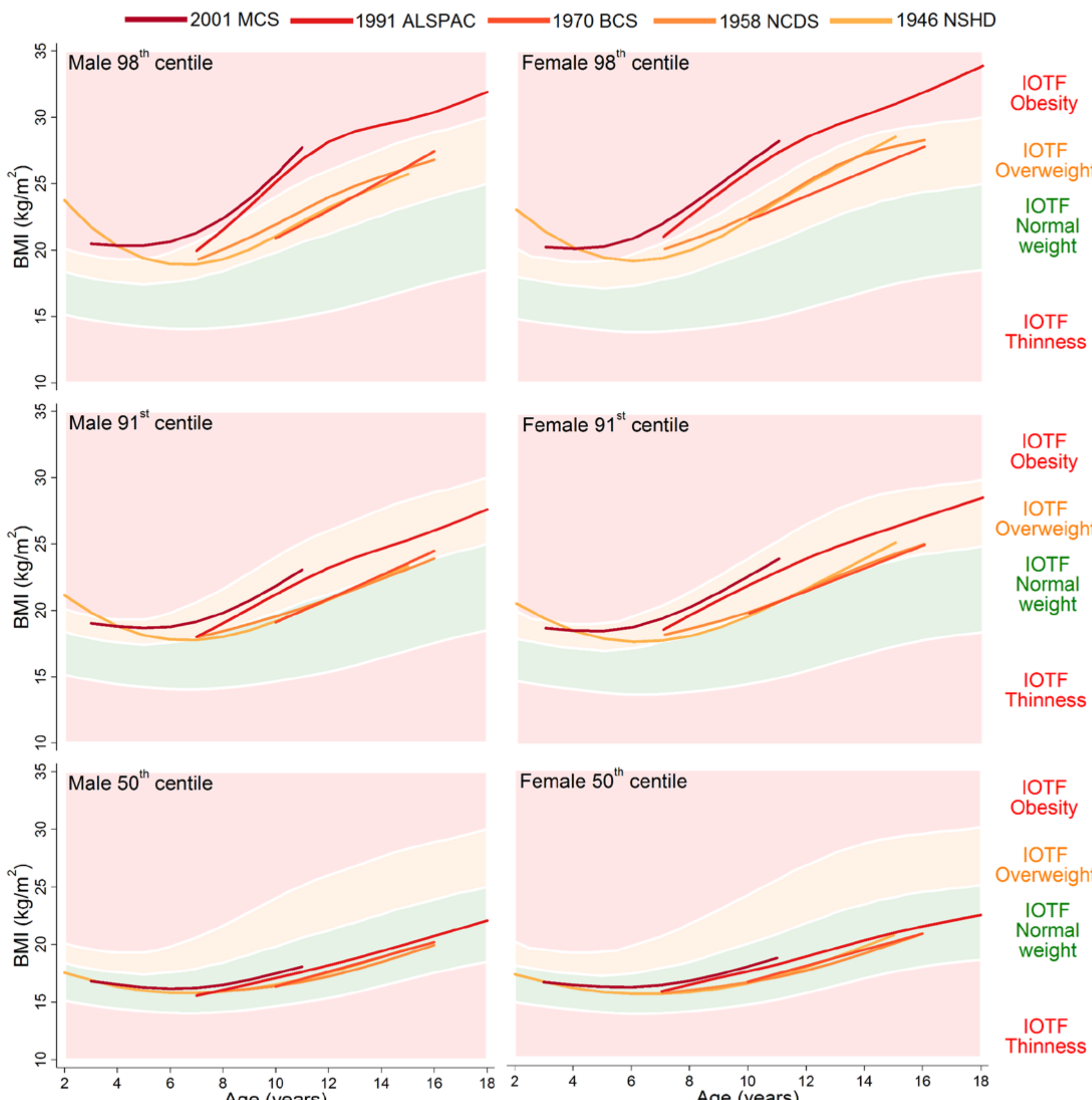


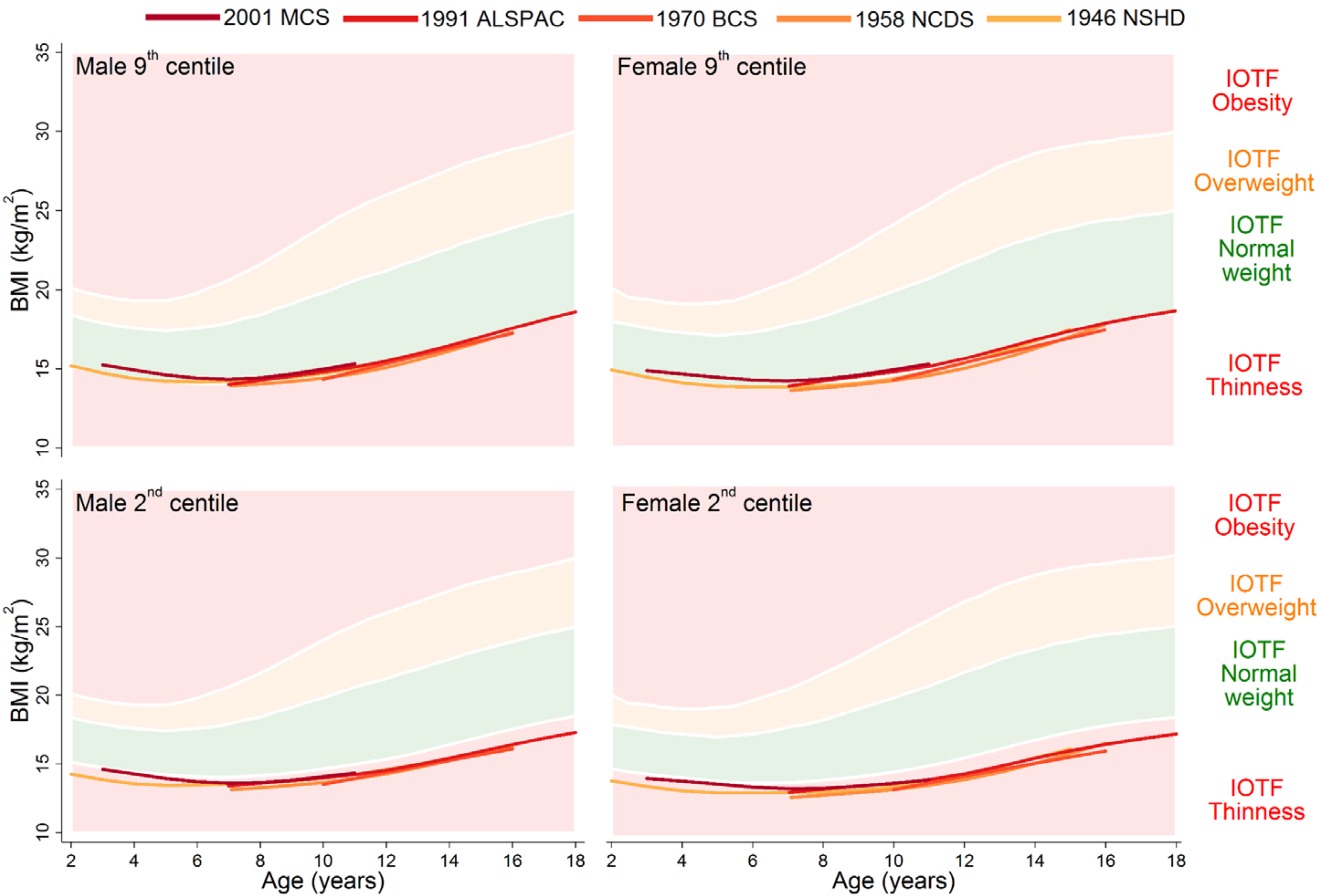
2. Binary logistic multilevel models used to produce trajectories describing the development of overweight or obesity (vs. normal weight) across age



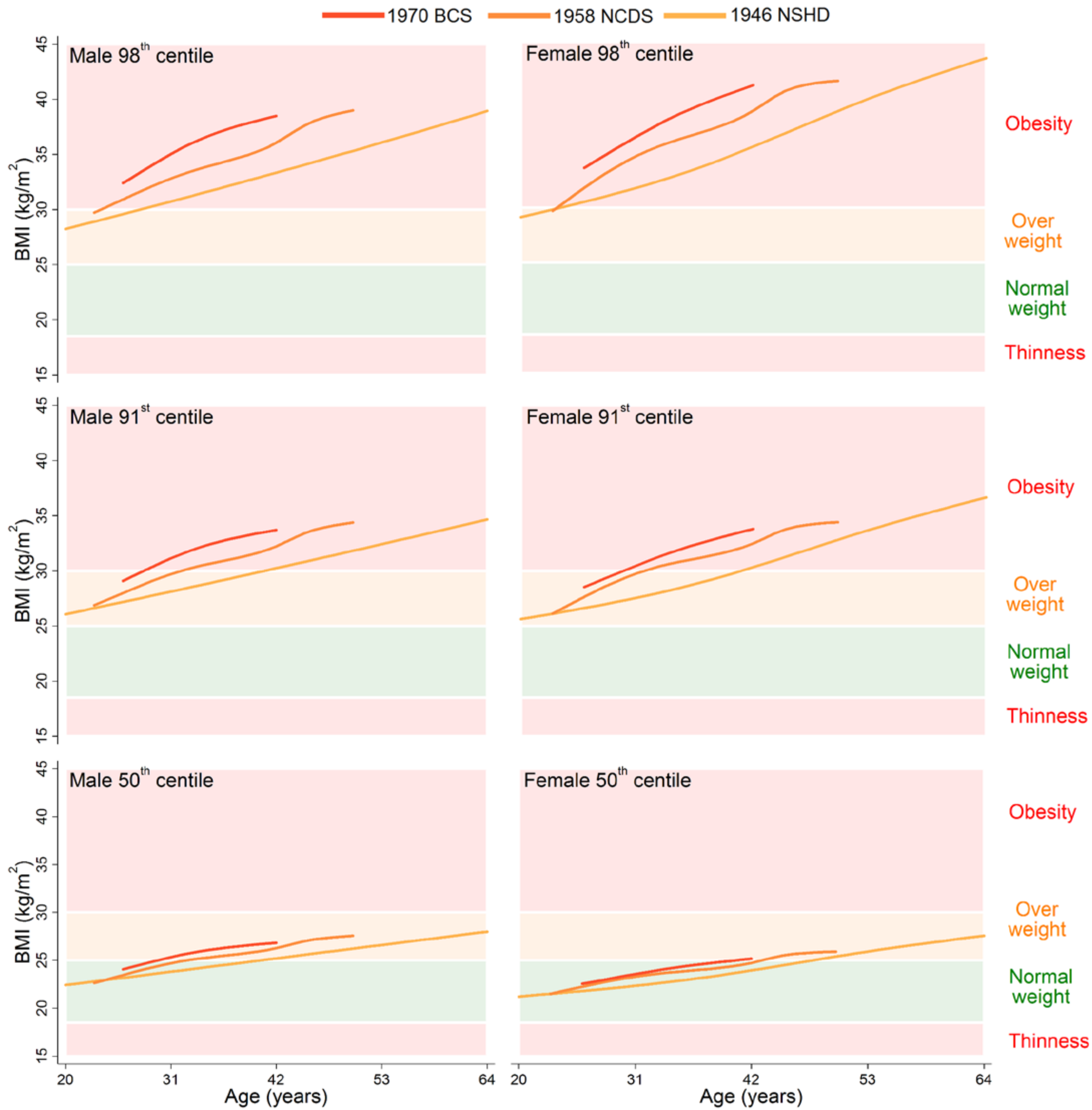
- Thinness, overweight, and obesity defined according to IOTF cut-offs during childhood and cut-offs of 18.5, 25, and 30 kg/m^2 during adulthood

	Male						Female				
	Sweep Target age (date)	N	BMI (kg/m^2) Median (IQR)	Thinness %	Overweight %	Obesity %	N	BMI (kg/m^2) Median (IQR)	Thinness %	Overweight %	Obesity %
1946 NSHD	2 (1948)	2046	17.7 (16.3, 19.2)	7.7	16.8	17.0	1794	17.2 (16.1, 18.8)	7.2	21.8	14.9
2,598 males	4 (1950)	2198	16.2 (15.3, 17.2)	10.5	16.5	6.1	1986	15.9 (14.9, 17.1)	9.8	15.8	4.8
2,359 females	6 (1952)	2050	15.9 (15.0, 16.7)	6.7	9.0	0.8	1841	15.6 (14.8, 16.5)	8.2	10.3	1.3
	7 (1953)	2057	15.8 (14.9, 16.6)	5.9	6.2	0.4	1920	15.5 (14.7, 16.5)	9.6	7.4	1.1
	11 (1957)	2050	16.9 (15.9, 18.1)	9.0	6.6	0.8	1887	17.0 (15.7, 18.7)	12.5	8.5	1.8
	15 (1961)	1881	19.3 (18.0, 20.8)	8.2	7.4	0.8	1700	20.3 (18.6, 22.1)	8.8	11.0	1.7
	20 (1966)	1802	22.5 (20.9, 24.0)	2.3	12.8	1.2	1629	21.4 (19.8, 23.1)	8.0	8.7	1.7
	26 (1972)	1822	23.1 (21.5, 25.1)	1.8	23.0	2.6	1782	21.8 (20.2, 23.8)	5.3	13.9	2.8
	36 (1982)	1631	24.6 (22.7, 26.7)	1.2	37.8	6.2	1618	22.6 (20.9, 25.1)	3.7	18.5	7.1
	43 (1989)	1612	25.3 (23.3, 27.7)	0.6	44.9	10.4	1595	24.0 (22.1, 27.1)	1.6	25.8	13.8
	53 (1999)	1451	27.0 (24.7, 29.7)	0.3	49.2	22.7	1494	26.2 (23.7, 30.1)	0.3	36.3	25.8
	60-64 (2006-2010)	1059	27.6 (25.0, 30.3)	0.3	46.7	28.1	1155	26.9 (24.2, 31.0)	1.0	37.0	30.2
1958 NCDS	7 (1965)	6499	15.8 (15.0, 16.7)	8.9	6.6	1.3	6068	15.6 (14.6, 16.7)	9.7	8.7	2.2
7,927 males	11 (1969)	5931	16.8 (15.8, 18.2)	12.4	6.7	1.3	5687	17.1 (15.8, 18.9)	16.0	9.0	1.4
7,514 females	16 (1974)	5194	19.8 (18.5, 21.4)	10.3	6.8	1.5	4911	20.6 (19.0, 22.5)	9.9	10.3	1.6
	23 (1981)	5680	22.7 (21.2, 24.5)	2.4	17.7	2.4	5732	21.6 (20.1, 23.5)	6.4	11.6	3.1
	33 (1991)	5006	25.1 (23.1, 27.5)	1.0	40.4	10.9	4982	23.4 (21.5, 26.4)	3.4	23.2	11.8
	42 (2000)	5069	26.0 (23.9, 28.5)	0.5	46.4	15.6	5195	24.1 (22.0, 27.5)	1.7	26.6	15.4
	44 (2002)	4249	27.3 (25.0, 30.1)	0.3	49.6	25.6	4305	25.7 (23.1, 30.0)	0.8	32.8	23.5
	50 (2008)	3833	27.4 (24.9, 30.4)	0.3	46.6	27.9	3814	25.7 (22.9, 29.5)	1.3	32.9	22.9
1970 BCS	10 (1980)	5738	16.4 (15.5, 17.7)	10.5	6.3	0.2	5443	16.7 (15.5, 18.3)	12.1	10.3	0.5
7,111 males	16 (1986)	3398	20.4 (19.0, 22.3)	10.1	9.3	1.9	3868	20.9 (19.3, 23.0)	10.7	11.3	1.6
6,781 females	26 (1996)	2322	24.1 (22.1, 26.1)	1.1	29.9	6.4	4324	22.3 (20.7, 24.8)	4.2	16.9	6.6
	30 (2000)	4796	25.1 (23.0, 27.6)	0.9	39.8	11.5	5072	23.2 (21.1, 26.3)	3.2	21.9	11.1
	34 (2004)	4107	26.0 (23.7, 28.7)	0.6	43.2	17.6	4398	24.0 (21.6, 27.4)	2.2	25.4	15.5
	42 (2012)	3907	26.8 (24.4, 29.8)	0.5	44.7	23.8	4037	24.9 (22.3, 28.8)	1.8	29.0	20.3
1991 ALSPAC	7 (1998)	3693	15.7 (14.9, 16.8)	7.8	9.2	2.5	3567	15.9 (14.9, 17.3)	7.2	13.1	4.0
4,461 males	8 (1999)	3048	16.5 (15.5, 17.8)	4.4	12.3	3.4	3017	16.8 (15.5, 18.5)	4.4	17.5	4.6
4,404 females	9 (2000)	3360	16.8 (15.6, 18.7)	7.7	13.5	3.6	3412	17.3 (15.8, 19.4)	8.0	18.0	4.5
	10 (2001)	3298	17.3 (15.9, 19.5)	7.1	14.4	4.1	3338	17.7 (16.1, 20.0)	8.4	17.2	4.9
	11 (2002)	3132	18.0 (16.5, 20.5)	7.6	16.2	4.4	3207	18.6 (16.8, 21.2)	8.9	18.6	4.6
	13 (2004)	2939	18.7 (17.1, 21.1)	7.9	16.0	4.3	3016	19.4 (17.6, 21.9)	9.3	16.9	3.9
	14 (2005)	2699	19.2 (17.7, 21.4)	8.0	13.4	3.9	2761	20.1 (18.3, 22.4)	9.3	15.8	3.8
	15 (2006)	2289	20.4 (18.8, 22.5)	6.5	13.6	3.9	2537	21.1 (19.4, 23.5)	8.2	15.0	4.8
	18 (2009)	1950	21.8 (20.0, 24.3)	8.2	15.9	5.9	2487	22.0 (20.1, 24.8)	7.5	16.7	7.4
2001 MCS	3 (2004)	5726	16.8 (16.0, 17.8)	3.0	18.6	5.1	5625	16.6 (15.7, 17.5)	3.9	19.6	5.1
6,897 males	5 (2006)	6114	16.1 (15.4, 17.1)	3.5	14.5	4.7	5846	16.1 (15.2, 17.1)	3.3	18.3	5.8
6,580 females	7 (2008)	5552	16.2 (15.2, 17.4)	4.7	12.7	4.9	5399	16.3 (15.2, 17.7)	4.8	16.9	6.2
	11 (2012)	5169	18.1 (16.5, 20.6)	5.1	18.7	6.0	5037	18.7 (16.8, 21.5)	6.4	22.6	6.7

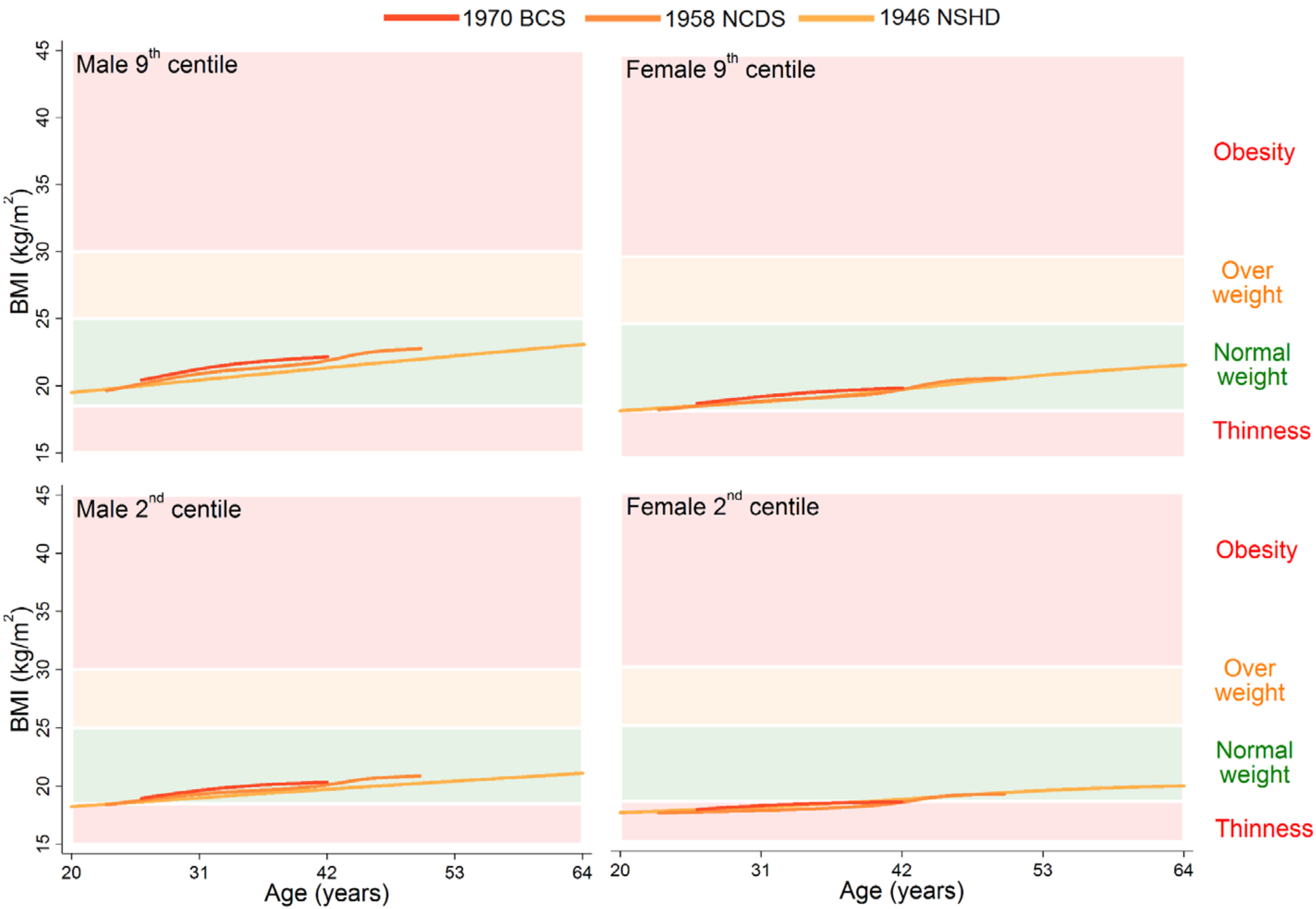

2001 MCS
1991 ALSPAC
1970 BCS
1958 NCDS
1946 NSHD
Male 98th centile
Female 98th centile
Male 91st centile
Female 91st centile
Male 50th centile
Female 50th centile
BMI (kg/m2)
Age (years)
IOTF Obesity
IOTF Overweight
IOTF Normal weight
IOTF Thinness



2001 MCS
1991 ALSPAC
1970 BCS
1958 NCDS
1946 NSHD
Male 9th centile
Female 9th centile
Male 2nd centile
Female 2nd centile
BMI (kg/m2)
Age (years)
IOTF Obesity
IOTF Overweight
IOTF Normal weight
IOTF Thinness



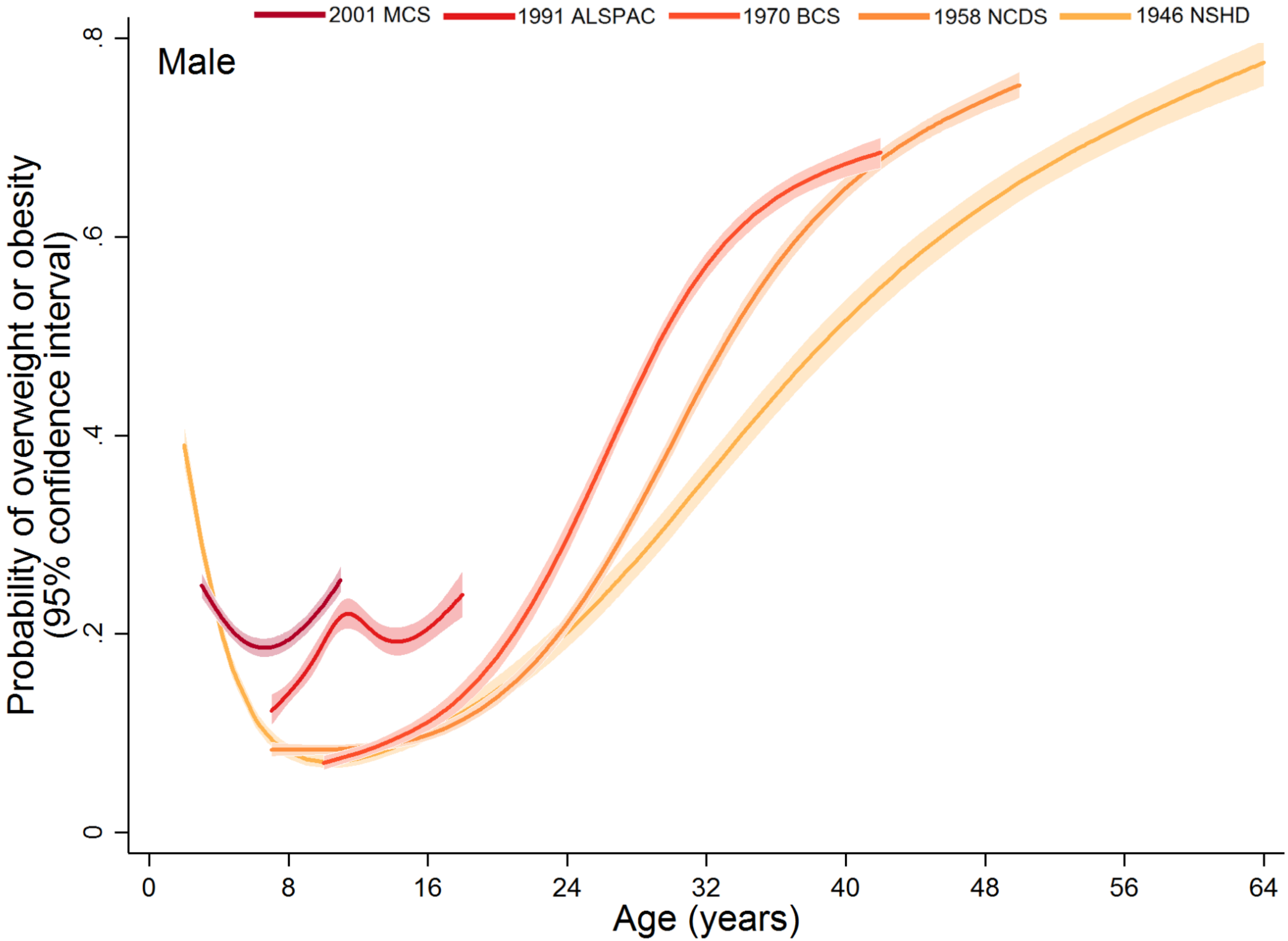
1970 BCS
1958 NCDS
1946 NSHD
Male 98th centile
Female 98th centile
Male 91st centile
Female 91st centile
Male 50th centile
Female 50th centile
BMI (kg/m2)
Age (years)
20
31
42
53
64
15
20
25
30
35
40
45
Obesity
Over weight
Normal weight
Thinness



1970 BCS
1958 NCDS
1946 NSHD
Male 9th centile
Female 9th centile
Male 2nd centile
Female 2nd centile
Obesity
Over weight
Normal weight
Thinness
Obesity
Over weight
Normal weight
Thinness
BMI (kg/m2)
BMI (kg/m2)
15
20
25
30
35
40
45
20
31
42
53
64
Age (years)
Age (years)

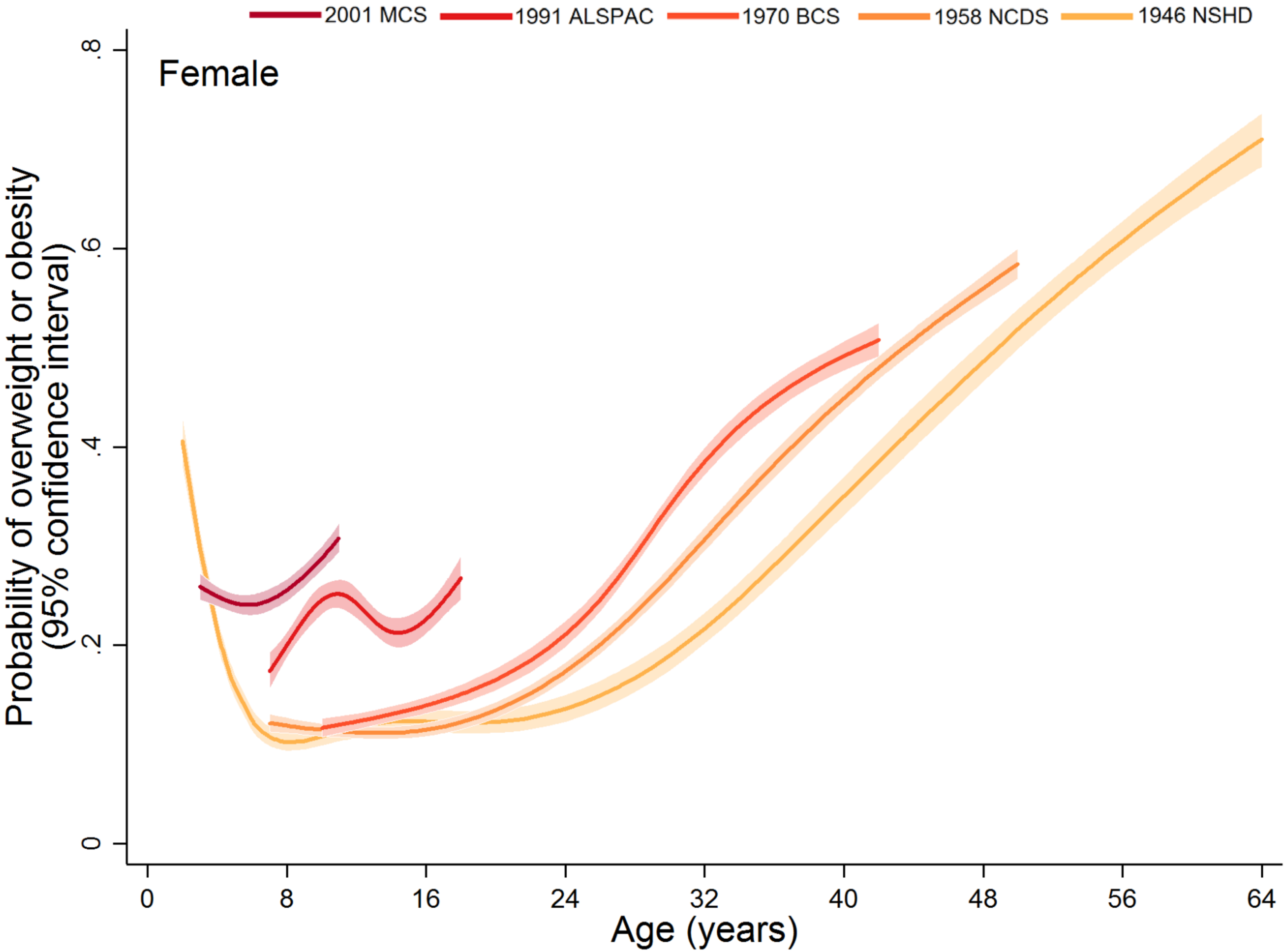
Prevalence at age 11 years has approximately tripled

- Males: 7.1 to 25.8%

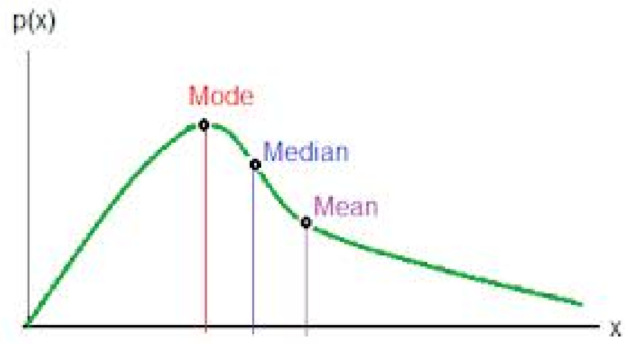


Prevalence at age 11 years has approximately tripled

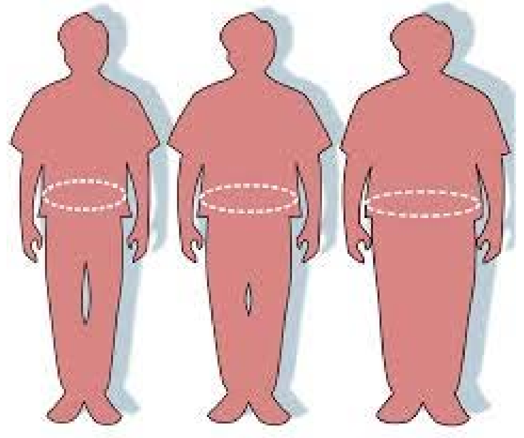
- Females: 11.3 to 31.1%



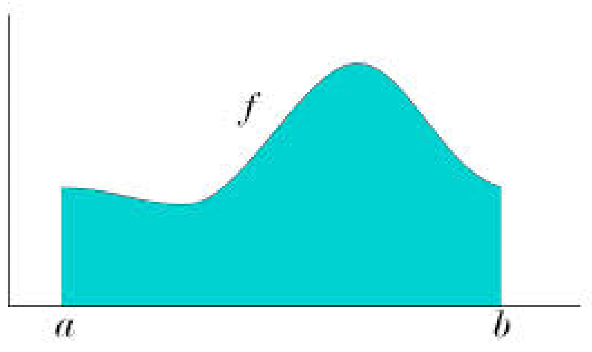
Shifts have occurred at the upper end of the BMI distribution, and in childhood this has contributed to a three-fold increase in overweight or obesity prevalence



Also age-related changes, that have contributed to the median UK adult currently being overweight, and shifts in these trajectories over time



These processes mean that more recently born cohorts are developing
1) Overweight or obesity earlier
2) Accumulating more exposure



Strengths

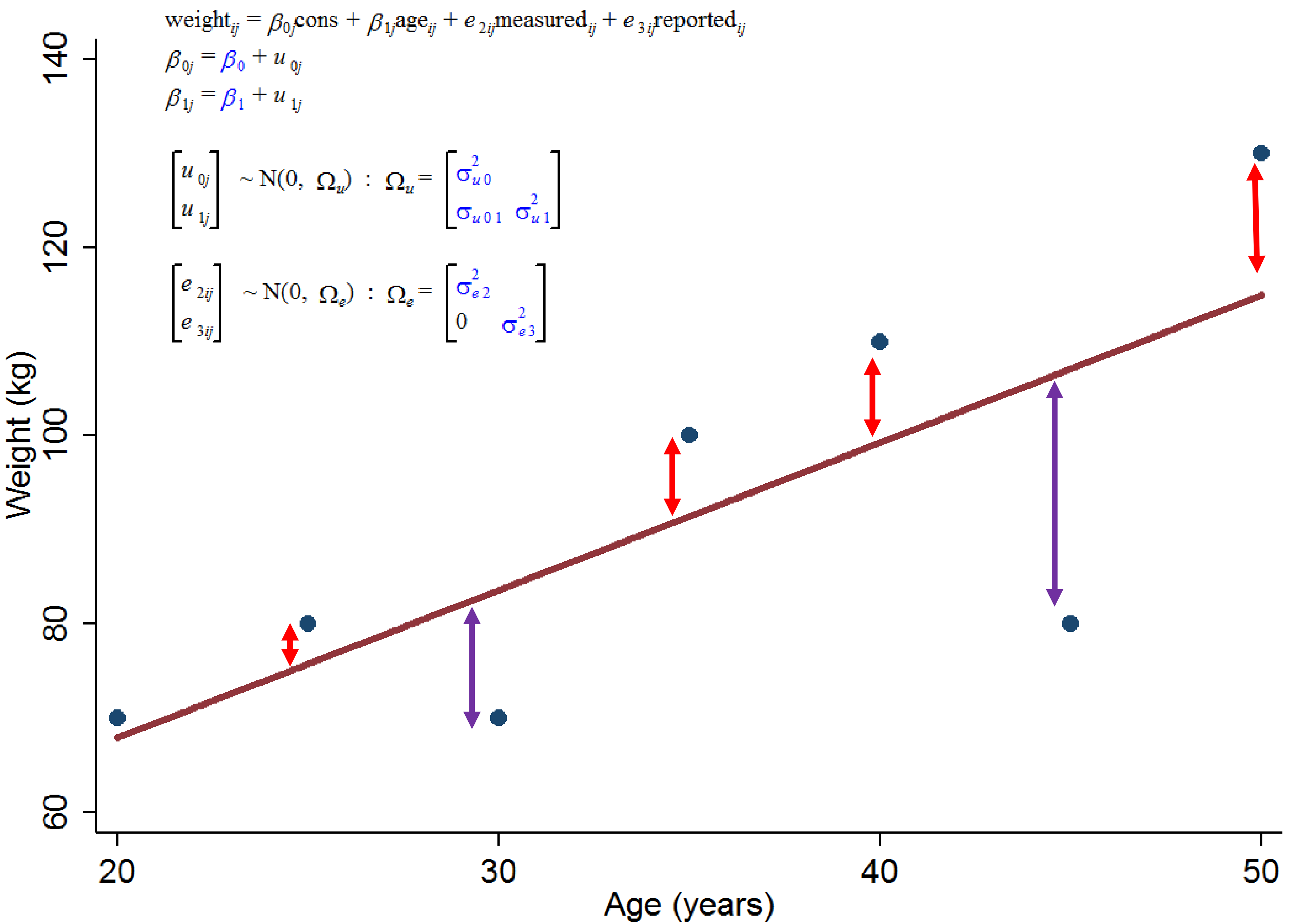
- Extensive serial data; wide range of ages and birth years
- Robust analysis, not focusing on mean BMI

Limitations

- Trajectories smoothed over ages where no data
- Normal limitations of BMI
- Measurement protocols not consistent

Future possbilties

- Determinants and consequences of the secular trends
- Multilevel models that parameterise measurement protocol differences in level one variance



$\text{weight}_{ij} = \beta_{0j}\text{cons} + \beta_{1j}\text{age}_{ij} + e_{2ij}\text{measured}_{ij} + e_{3ij}\text{reported}_{ij}$
$\beta_{0j} = \beta_0 + u_{0j}$
$\beta_{1j} = \beta_1 + u_{1j}$
$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{bmatrix}$
$\begin{bmatrix} e_{2ij} \\ e_{3ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} \sigma^2_{e2} & \\ 0 & \sigma^2_{e3} \end{bmatrix}$
Weight (kg)
140
120
100
80
60
20
30
40
50
Age (years)

Trajectories

- Powerful approach to understand how something changes over age

Cross-cohort comparisons of trajectories

- Different birth year cohorts: powerful approach to investigate how some age-related process has changed over time
- Different geographical cohorts: powerful approach to investigate how some age-related process differs between settings with different confounding structures

Harmonisation and longitudinal methods

- Laborious but necessary
- Degree of harmonisation and which longitudinal method to use are dependent on each other and the research question

size: 35,742,708

variable name	storage type	display format	value label	variable label
sid	float	%9.0g	sidlab	study id
pid	str7	%7s		original particpant id
newpid	float	%9.0g		new participant id
visitnumber	float	%9.0g		visit number
visitage	byte	%9.0g		visit age (years)
sex	double	%8.0g	sexlab	sex
eth	byte	%13.0g	ethlab	ethnicity
core	float	%9.0g	corelab	indicates whether or not participant was part of original cohort
ideath	float	%22.0g	ideathlab	indicates whether or not the partciapnt survived infancy
bmulti	byte	%14.0g	bmultilab	indicates multiple birth (eg twin or triplet)
border	double	%8.0g		birth order
gestdy	double	%8.0g		gestational age at birth (days)
gestwk	double	%9.0g		gestational age at birth (completed weeks)
day	double	%9.0g		day of assessment
month	double	%8.0g		month of assessment
year	double	%8.0g		year of assessment
date	float	%dM_d,_CY		date of assessment
age	float	%9.0g		visit age (years)
xage	float	%9.0g		exact age at assessment (decimal years)
xagedum	float	%9.0g		indicates if xage was orignally missing
nwt	float	%9.0g	*	number of weight observations
wt	double	%10.0g		weight (kg)
wtself	float	%10.0g	selflab	indicates whether weight was measured of self-report
wtimp	float	%9.0g	implab	indicates whether weight was measured in imperial or metric
wtpre	float	%9.0g		indicates precision of weight measurement
meanwt	float	%9.0g		mean weight: sex and visitage stratified
sdwt	float	%9.0g		standard deviation of weight: sex and visitage stratified
flagwt	float	%9.0g		flags weights more than 5*sdwt from meanwt
wtch	float	%9.0g		weight change between age x and age x-1
meanwtch	float	%9.0g		mean wtch: sex and visitage stratified
sdwtch	float	%9.0g		standard deviation of wtch: sex and visitage stratified
flagwtch	float	%9.0g		flags wtch more than 5*sdwtch from meanwtch
nht	float	%9.0g	*	number of height observations
ht	double	%10.0g		height (m)
htself	float	%10.0g	selflab	indicates whether height was measured of self-report
htimp	float	%9.0g	implab	indicates whether height was measured in imperial or metric
htpre	float	%9.0g		indicates precision of height measurement
meanht	float	%9.0g		mean height: sex and visitage stratified
sdht	float	%9.0g		standard deviation of height: sex and visitage stratified
flaght	float	%9.0g		flags heights more than 5*sdht from meanht
htch	float	%9.0g		height change between age x and age x-1
meanhtch	float	%9.0g		mean htch: sex and visitage stratified
sdhtch	float	%9.0g		standard deviation of htch: sex and visitage stratified
flaghtch	float	%9.0g		flags htch more than 5*sdhtch from meanhtch
nbmi	float	%9.0g	*	number of bmi observations
bmi	float	%9.0g		body mass index (kg/m2)

Funders

CLOSER is funded by the Economic and Social Research Council (ESRC) and the Medical Research Council (MRC). It has been awarded a core grant of approximately £5 million for 2012 to 2017. This funding is made possible by a landmark contribution from the Government's Large Facilities Capital Fund.

CLOSER
closer.ac.uk

1946 NSHD
nshd.mrc.ac.uk

1991 ALSPAC
bristol.ac.uk/alspac

1958 NCDS, 1970 BCS, 2001 MCS
cls.ioe.ac.uk
ukdataservice.ac.uk