Scoping existing dietary data available in CLOSER to support cross-cohort research questions

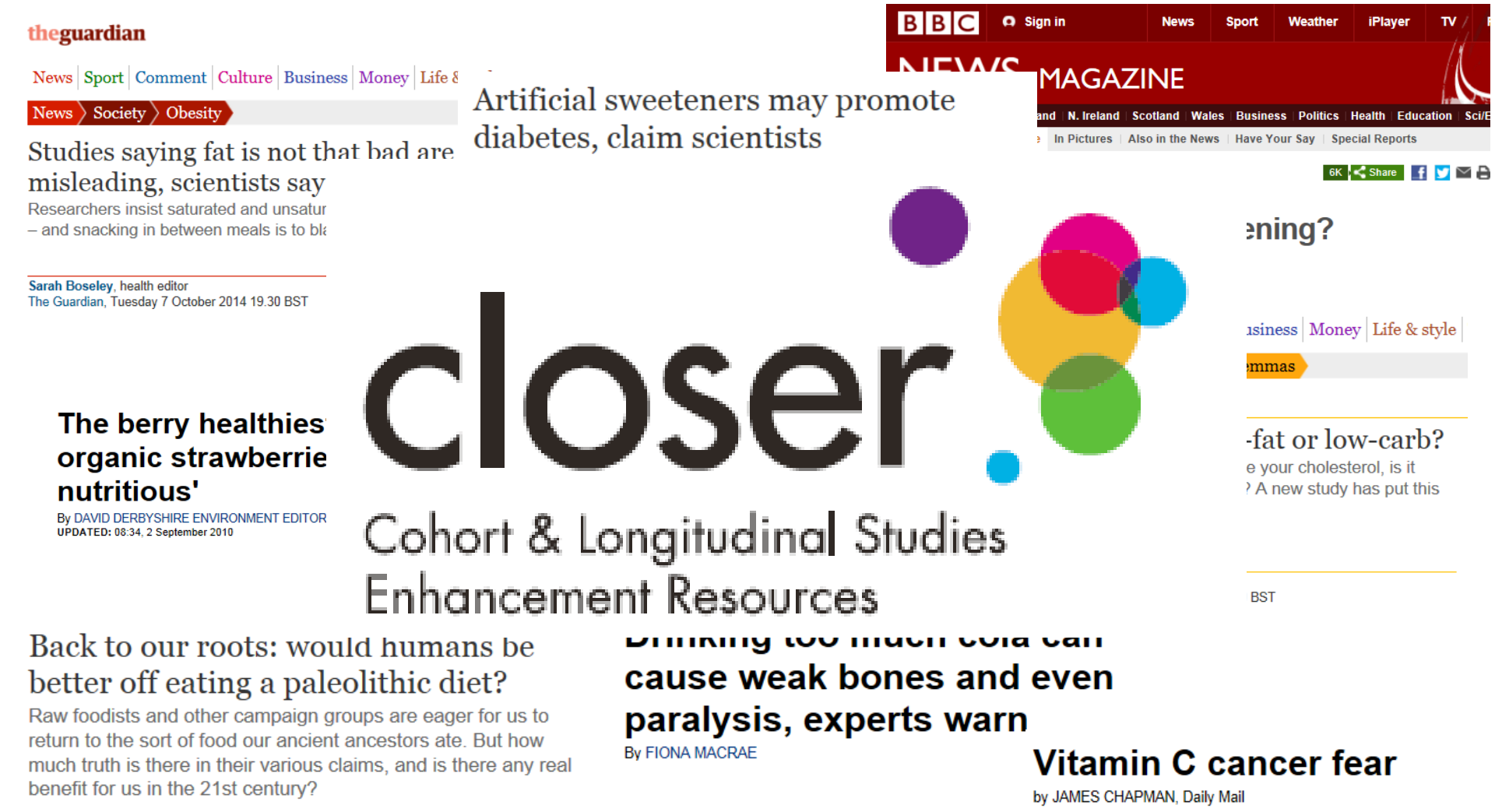
Jane Maddock

Co-investigator: Rebecca Hardy

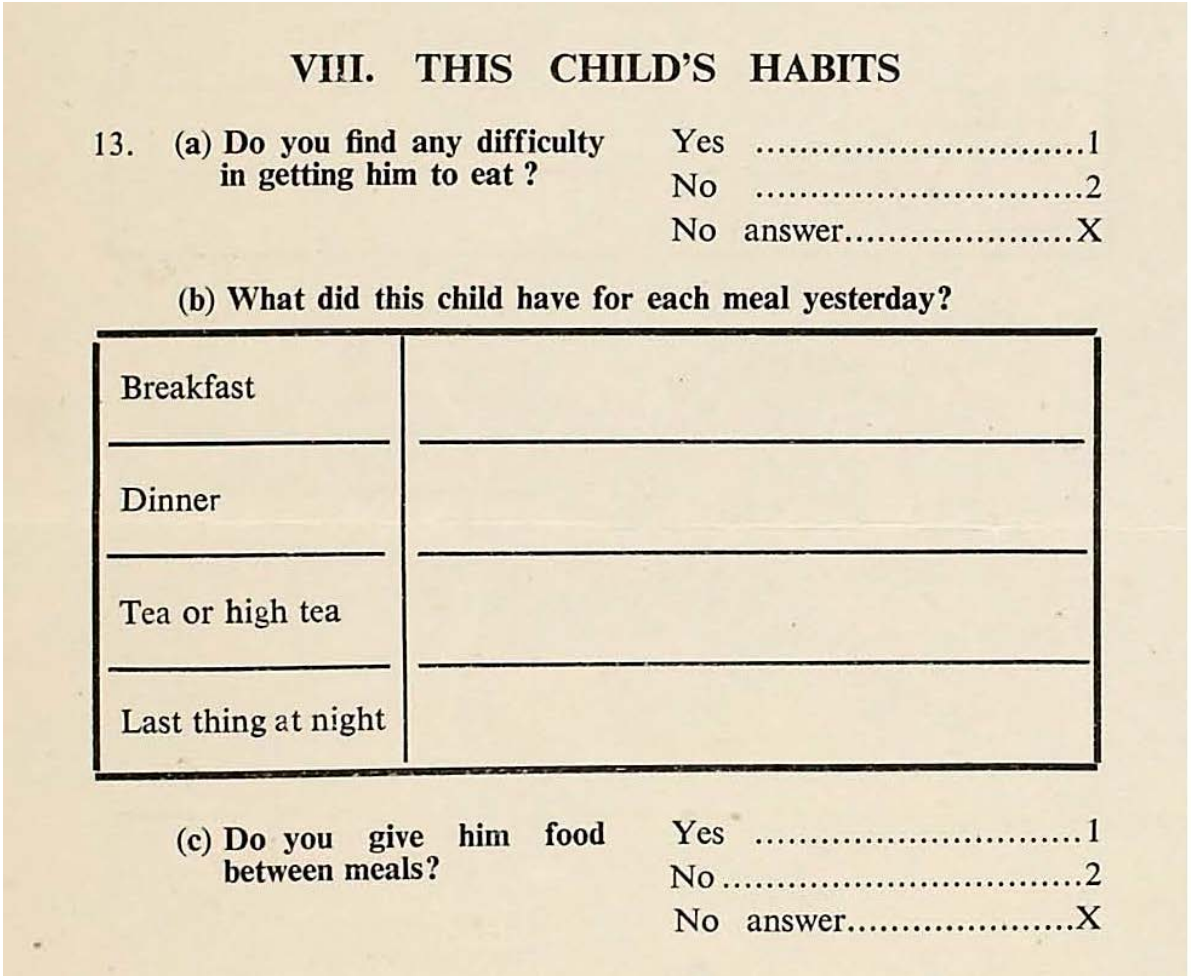
Collaborators: Meena Kumari, Sian Robinson, David Bann, Matthew Suderman

28th March 2019

Why is this project important?



What did you eat yesterday?



VIII. THIS CHILD'S HABITS

13. (a) **Do you find any difficulty in getting him to eat ?**

Yes 1
No 2
No answer.....X

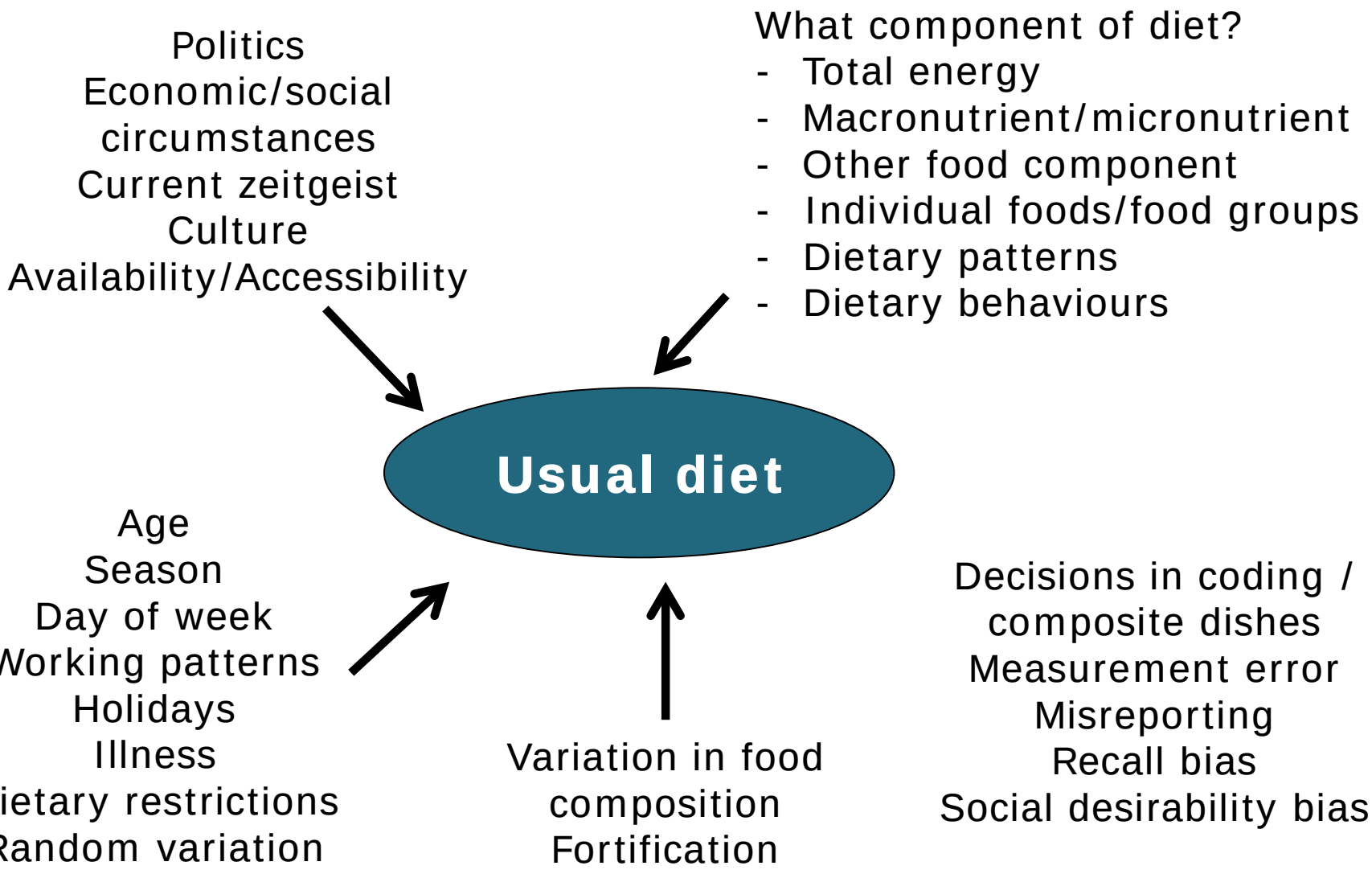
(b) **What did this child have for each meal yesterday?**

Breakfast	
Dinner	
Tea or high tea	
Last thing at night	

(c) **Do you give him food between meals?**

Yes 1
No 2
No answer.....X

Why is diet so difficult to study in cohort studies?



Why bother?


BMJ

Implausible results in human nutrition research

Definitive solutions won't come from another million observational papers or small randomized trials

John P A Ioannidis *professor of medicine, health research and policy, and statistics*

- Observational claims ≠ trials
- Too complex for questionnaire methods
- Too much confounding and bias


Advances in Nutrition

AN INTERNATIONAL REVIEW JOURNAL

Understanding Nutritional Epidemiology and Its Role in Policy[1,2]

Ambika Satija,[3,4] Edward Yu,[3] Walter C Willett,[3–5] and Frank B Hu[3–5]*

[3]Department of Nutrition and [4]Department of Epidemiology, Harvard School of Public Health, Boston, MA; and [5]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA

- Improvements in validity of dietary assessment methods
- Improvements in design/analyses of prospective cohort studies
- Inability to capture long-term diet and endpoints in RCTs

It is not perfect, but with knowledge and cautious interpretation we can maximise the use of diet data in CLOSER to inform policy

Project objective and impact

Main objective

To **document**, **describe** and make comparisons between the available dietary intake information across CLOSER cohorts

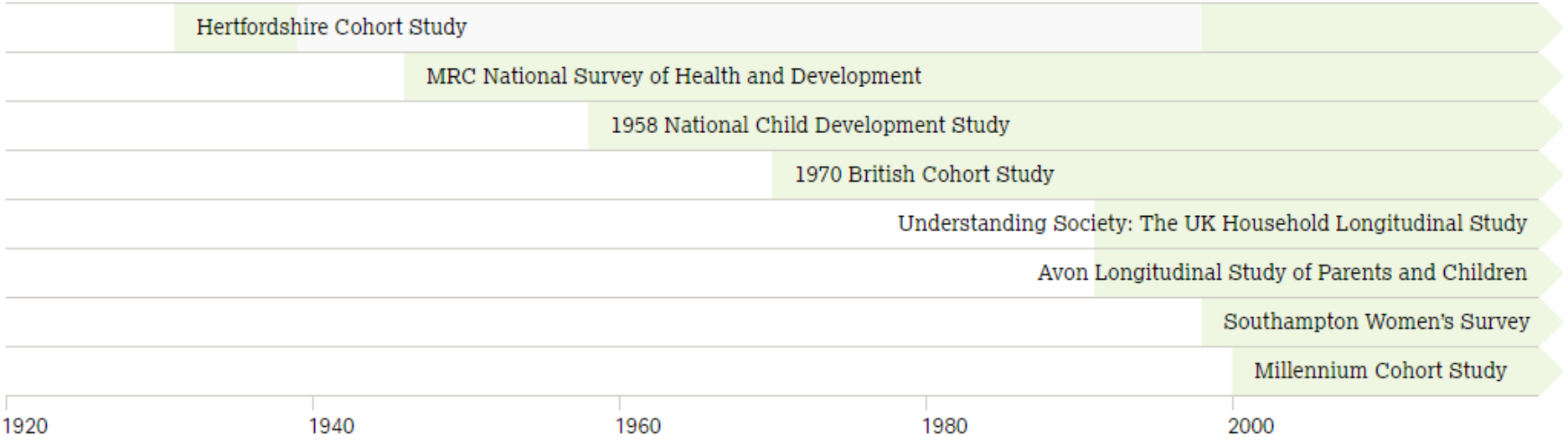
Impact

- Support future researchers in using the dietary data both within and across cohorts
- Prepare the cohorts for next advancement in nutritional epidemiology and dietary assessment

Milestones

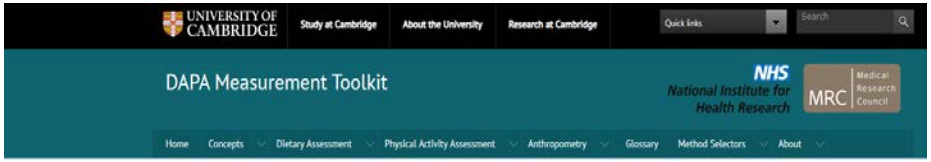
Milestone	
1. Access relevant data and meta-data	
2. Document how diet was collected in each cohort and identify major between-study differences in dietary assessment	
3. Identify relevant variables within each study and document major between-study differences e.g. type of dietary variables (e.g. disaggregated variables, food grouping)	
4. Perform within-study descriptive dietary analysis	
5. Identify if and where diet can be harmonised between the studies	

Overview of CLOSER cohorts



https://www.closer.ac.uk/

Overview of dietary assessment methods



Broadly, dietary assessment asks:

1. What was eaten
2. How much was eaten
3. How often is this eaten

Objective measures example	Subjective measures example
Direct observation Duplicated diets Biomarkers	Estimated or weighed food diaries 24 hour recall Food frequency questionnaire Food checklists

Food diaries

Strengths	Weaknesses
• Detailed • Good estimates of short-term total intake • Can capture contextual situations • Prospective; little reliance on memory • Design can include prompts	• Not suitable for retrospective study • Not suitable irregularly consumed foods • Risk of low completion as number of days increases • High participant burden • Moderate/high researcher burden • Expensive to code • Reliance on individuals ability to describe portion sizes

Adapted from Nutritools.org

Food frequency questionnaires

Strengths	Weaknesses
• Useful for long term usual intakes of foods • Can capture irregularly consumed foods • Low researcher burden • Low participant burden	• Not suitable for cross-country comparisons unless comparable food lists are included • Requires good participant memory • Restricted to items specifically listed in the instrument • Requires specific algorithms to convert frequencies to nutrients

Adapted from Nutritools.org

Milestones 2 & 3:

Document how diet was collected in each cohort & identify relevant variables within each study

Overview of when & how diet was collected in each study

	NSHD	NCDS	BCS70	MCS	HCS	SWS	ALSPAC	UHKLS
Year (age y)	1950 (4) 1982 (36) 1989 (43) 1999 (53) 2006-11(60-64) 2014-2015 (69)	1991(33) 2000(42) 2003(45)	1980 (10) 1986 (16) 2000 (30) 2012 (42) 2016 (46)	2001 (9m) 2004 (3) 2006 (5) 2008 (7) 2012 (11) 2015 (14)	1998-2004 (~65)	1998-(11wk-13y)	1991-2006** (32wks – 13y)	1991-2016***
FFQ					✓	✓	✓	
Diary	✓		✓			✓	✓	
Recall	✓					✓		
Other*	✓	✓	✓	✓				✓

*diet-related questions, non-specific dietary assessment tool
**diet measured at multiple time points from both mothers, partners and children
***some form of dietary data collected throughout waves; some more detailed than others

NSHD dietary data

When	Method
1950 (4y)	24-hour recall (Prynne *et al.* 1999 & 2002 PHN)
1982 (33y)	5-day estimated food diaries & 48-hour recall
1989 (43y)	5-day estimated food diaries & 48-hour recall
1999 (53y)	5-day estimated food diaries
2006-2011 (60-64y)	5-day estimated food diaries
2014-2015	Diet-related questions e.g. How many days do you eat breakfast, what type of bread do you eat?

Extracting nutrients:

- Determine portion size and link to time-appropriate food composition tables: McCance and Widdowson Composition of foods for nutrient content
- MRC HNR In-house programmes "Diet In Data Out (DIDO) and Diet in Nutrients Out (DINO) (Fitt *et al.* 2015 PHN)

BCS70 dietary data

When	Method
1970 (Birth)	Diet-related questions breastfeeding & alternatives
1980 (10y)	• Pupil questionnaire: Frequency of 9 foods, purchasing snack behaviour & free school meals
1986 (16y)	• Pupil questionnaire (B&C): drinks/soft drink consumption & frequency of 18 foods groups • Health behaviour questionnaire (F): "What did you eat and drink yesterday" from pre-defined list • Home (G): Diet-related questions e.g. are you a vegetarian, do you add sugar to drinks • Maternal (P): Diet-related questions e.g. teenager eats cereal & cooking methods & special diet • 4-day food diary*
2006 (30y)	Frequency of consumption of 15 foods
2012 (42y)	Frequency of consumption of home-cooked, take-away, read meals etc.
2016 (46y)	Online dietary data (two 24-hour recalls)

*Not deposited on UK Data Archive (coded by Helen Crawley ~1990s)

SWS dietary data

Time point Age of study child	Mother's 100-item FFQ*	Mother's 24-food diary	Children's FFQ	Child's 24-hour recall	Children's food diary
pre-pregnancy	✓	✓			
11 weeks gestation	✓	✓			
34 weeks gestation	✓				
(non-pregnant sub-sample)	✓				
6 months	✓		✓34-item	✓	✓4-day weighted diary sub-sample
1y	✓		✓78-item		✓4-day weighted diary sub-sample
3y	✓		✓80-item		✓2-day
6-7y	✓		✓		
8-9y	✓		✓		
11-13y	✓		✓		

Extracting nutrients:

- Standard portion sizes allocated to each food item from published sources
- Frequency of a portion*nutrient content from food composition tables (McCance & Widdowson)

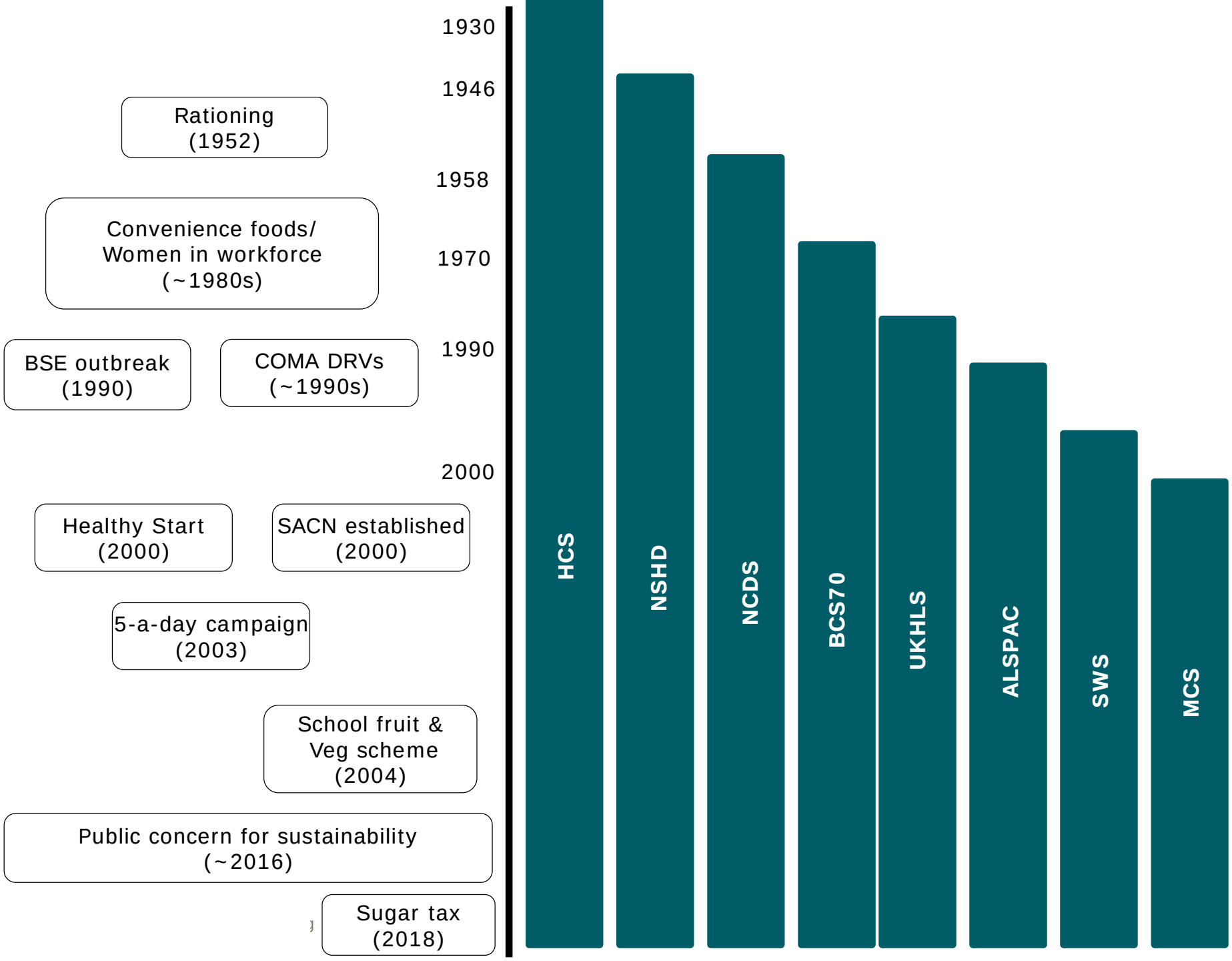
*Robinson 1996 EJCN

UKHLS dietary data

Household-level	Individual-level	Youth	Newborn
Money spent on food purchases	How often do you eat out	Fruit/veg consumption	Infant feeding
	How much do you spend	Crips/fizzy drinks	
	Type of bread/milk	Fast food/takeaways	
	Vegetable/fruit consumption		
	Consumption of ethnic foods (Wave 2 & 5)		

Milestone 4:

Perform within-study descriptive dietary analysis


1930
1946
1958
1970
1990
2000
Rationing (1952)
Convenience foods/ Women in workforce (~1980s)
BSE outbreak (1990)
COMA DRVs (~1990s)
Healthy Start (2000)
SACN established (2000)
5-a-day campaign (2003)
School fruit & Veg scheme (2004)
Public concern for sustainability (~2016)
Sugar tax (2018)
HCS
NSHD
NCDS
BCS70
UKHLS
ALSPAC
SWS
MCS

How has this diet data has been used?

- NSHD: n=~28
- NCDS: n=~10
- BCS70: n=~12
- MCS: n=~13
- SWS: n=~27
- ALSPAC: n=>90
- HCS: n=~16
- UKHLS: n=~5

Example: Prynne et al. 1999 PHN

- 24-hour recall NSHD (1950) vs NDNS (1992/92)

Table 1 Mean* daily weights of food groups consumed, and frequency of consumption by children aged 4 years in national studies in 1950 and 1992/93[10]

	NSHD 1950 (n=4599)		NDNS 1992/93 (n=493)	
Food groups	g day^{-1}	% consuming	g day^{-1}	% consuming
Pasta, rice, etc.	< 1	< 1	31	†
Bread	120	97	48	†
Breakfast cereals	27	59	21	†
Cakes	30	48	28	61
Biscuits	4	28	17	88
Puddings	31	28	21	†
Milk puddings	71	65	9	23
Milk	307	98	247	†
Cream	3	5	4	17
Cheese	2	8	6	57
Yoghurts	0	0	23	33
Eggs	32	55	9	47
Spreading fats	20	86	7	†
Meat (beef, lamb, bacon, etc.)	24	68	21	†
Poultry	≪ 1	< 1	11	†
Offal	1	3	< 1	4
Meat products and dishes	22	23	26	†
Fish and fish products	7	11	10	†
Leafy vegetables	13	25	6	†
Root vegetables	3	7	7	†
Pulses, dry	3	4	11	†
Other vegetables	54	60	35	†
Potatoes	75	79	66	†
Fruit and nuts	36	40	51	†
Preserves, spreads	15	46	2	†
Confectionery	< 1	< 1	25	†
Sugar	6	66	3	57
Tea	194	55	35	38
Soft drinks and juices	13	11	446	†
Sauces, soups, etc.	16	26	33	†

*Over all children.
†Frequency for this food group as a whole was not given.

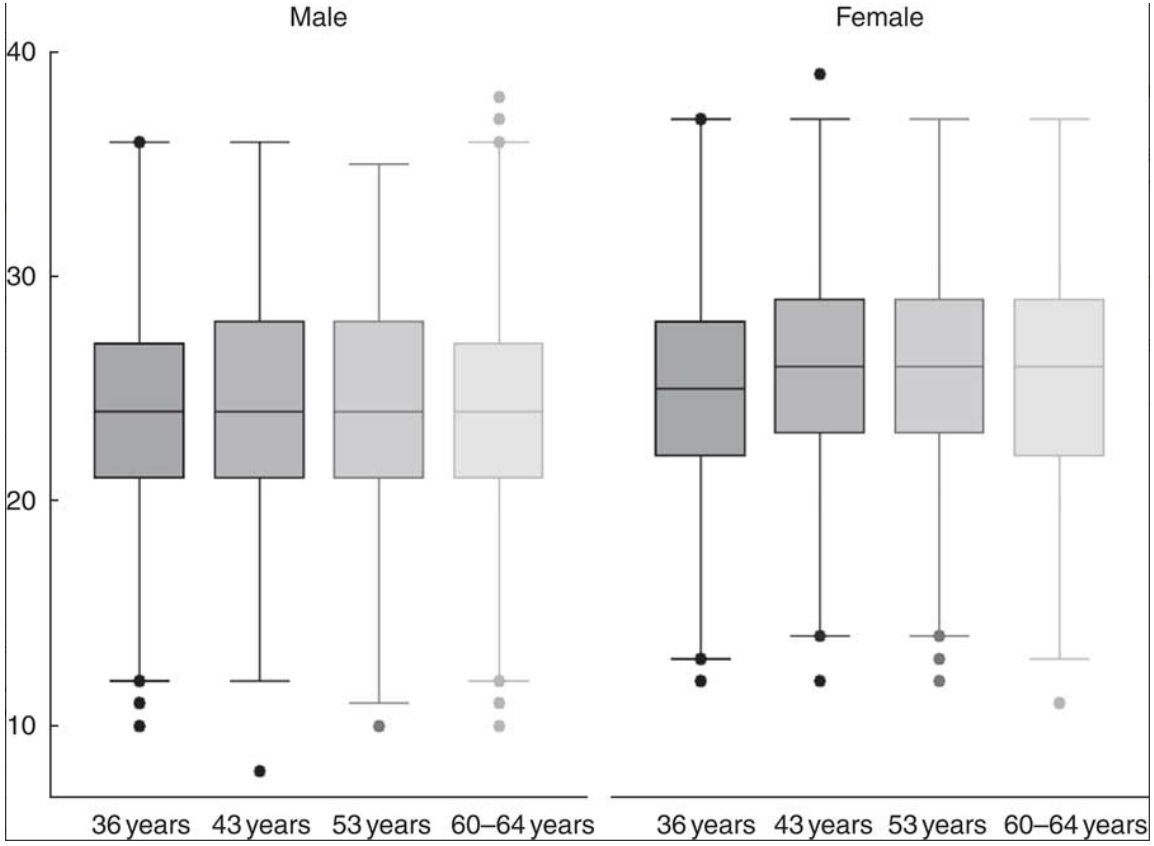
Example: Prynne et al. 1999 PHN

Table 4 Percentage contribution of food groups to micronutrient intake of children aged 4 years in national studies in 1950 (n=4599) and 1992/93 (n=493)[10]

	Calcium		Iron		Carotene		Vitamin A (retinol equivalents)		Vitamin C	
Food group	1950	1992/93	1950	1992/93	1950	1992/93	1950	1992/93	1950	1992/93
Cereals and cereal products	39	23	46	50	2	5	10	10	3	3
of which bread	17	8	29	12	0	NR	0	NR	0	NR
of which breakfast cereals	2	3	7	21	<1	NR	1	NR	0	NR
of which biscuits	<1	NR	1	6	0	NR	0	NR	0	NR
of which milk puddings	13	NR	1	NR	2	NR	5	NR	2	NR
Milk and milk products	50	59	2	5	9	6	24	31	8	6
of which liquid milk	48	46	2	4	9	5	22	23	8	5
Eggs and egg dishes	3	1	8	3	0	0	7	4	0	0
Fat spreads	0	0	<1	0	7	4	20	12	0	0
Meat and meat products	1	3	22	15	17	6	27	14	1	2
Fish and fish products	1	1	1	2	0	<1	0	2	0	0
Vegetables	3	4	12	15	59	61	11	22	59	22
of which potatoes	1	NR	5	6	0	NR	0	1	28	13
of which savoury snacks	0	NR	0	2	0	NR	0	1	0	1
Fruit and nuts	2	1	2	3	4	2	1	1	19	15
Sugar, preserves and confectionery	1	4	3	4	0	1	0	1	3	1
Beverages	1	3	3	2	<1	12	0	4	9	52
Miscellaneous	<1	1	1	2	<1	3	0	1	0	0

NR, not reported.

Example: Maddock et al. 2018 BJN



Dietary Approaches to Stop Hypertension Diet score

Association between long-term Dietary Approaches to Stop Hypertension (DASH) scores (sex-specific quintiles) and classic cardiovascular risk factors[*] – unadjusted (Coefficients and 95 % confidence intervals)

			DASH Q2		DASH Q3		DASH Q4		DASH Q5		
	n	DASH Q1	Coefficient	95 % CI	Coefficient	95 % CI	Coefficient	95 % CI	Coefficient	95 % CI	P_{trend} [†]
Unadjusted											
Antihypertensive medication[‡]	1841	Ref.	0·97	0·65, 1·28	0·68	0·36, 1·00	0·68	0·36, 1·00	0·56	0·23, 0·89	<0·001
Diastolic blood pressure[§]	1825	Ref.	−0·35	−2·21, 1·51	−1·13	−2·93, 0·67	−1·78	−3·57, 0·02	−3·09	−4·89, −1·30	<0·001
Systolic blood pressure[§]	1824	Ref.	−1·98	−5·40, 1·45	−4·60	−7·93, −1·28	−5·39	−8·70, −2·08	−7·70	−11·00, −4·40	<0·001
Lipid-lowering medication[‡]	1883	Ref.	0·93	0·60, 1·25	0·62	0·28, 0·96	0·63	0·29, 0·97	0·45	0·08, 0·81	<0·001
Total cholesterol[§]	1746	Ref.	−0·07	−0·25, 0·12	−0·02	−0·20, 0·15	−0·09	−0·26, 0·09	−0·15	−0·32, 0·03	0·10
LDL-cholesterol[§]	1670	Ref.	−0·08	−0·24, 0·08	−0·02	−0·18, 0·13	−0·06	−0·21, 0·10	−0·07	−0·22, 0·08	0·50
HDL-cholesterol	1746	Ref.	0·05	−0·01, 0·12	0·10	0·04, 0·17	0·11	0·04, 0·17	0·15	0·08, 0·21	<0·001
ln TAG[§]	1679	Ref.	−7·27	−16·56, 2·02	−24·51	−33·50, −15·52	−25·73	−34·77, −16·69	−41·35	−50·30, −32·40	<0·001
Adjusted for SEP, BMI, smoking and physical activity											
Antihypertensive medication[‡]	1774	Ref.	0·98	0·64, 1·32	0·73	0·39, 1·08	0·82	0·48, 1·17	0·78	0·41, 1·15	0·12
Diastolic blood pressure[§]	1759	Ref.	−0·35	−2·19, 1·50	−0·60	−2·41, 1·21	−0·97	−2·80, 0·86	−1·59	−3·49, 0·31	0·08
Systolic blood pressure[§]	1758	Ref.	−1·93	−5·35, 1·49	−3·45	−6·81, −0·08	−3·62	−7·01, −0·23	−4·83	−8·35, −1·31	0·01
Lipid-lowering medication[‡]	1813	Ref.	1·01	0·66, 1·36	0·68	0·32, 1·05	0·76	0·39, 1·12	0·61	0·21, 1·02	0·01
Total cholesterol[§]	1687	Ref.	−0·13	−0·32, 0·06	−0·07	−0·25, 0·12	−0·11	−0·30, 0·08	−0·16	−0·35, 0·04	0·19
LDL-cholesterol[§]	1617	Ref.	−0·13	−0·30, 0·04	−0·05	−0·22, 0·11	−0·07	−0·23, 0·10	−0·06	−0·23, 0·11	0·83
HDL-cholesterol	1687	Ref.	0·03	−0·03, 0·09	0·05	−0·01, 0·11	0·04	−0·03, 0·10	0·05	−0·02, 0·11	<0·001
ln TAG[§]	1626	Ref.	−2·89	−11·88, 6·09	−13·84	−22·64, −5·04	−13·35	−22·31, −4·39	−22·59	−31·85, −13·33	<0·001

Q, sex-specific quintile; Ref., referent values.

[*] *N*'s not restricted to those with carotid intima–media thickness or pulse wave velocity measures.

[†] Linear trend test, that is DASH quintiles fitted as continuous exposure in regression model. No evidence for deviation from linear tend using log likelihood ratio test, that is testing DASH quintiles fitted as continuous exposure *v.* DASH quintiles fitted as categorical exposure, $P \geq 0{\cdot}17$ for all models.

[‡] Logistic regression, OR presented.

[§] Censored regression to account for medication use.

Long-term Dietary Approaches to Stop Hypertension (DASH) score and vascular function (Coefficients and 95 % confidence intervals)

	Standardised carotid intima–media thickness*		Standardised pulse wave velocity†	
	Coefficient	**95 % CI**	**Coefficient**	**95 % CI**
DASH score sex-specific quintiles (Q)				
Model 1‖				
Q1	Ref	Ref	Ref	Ref
Q2	−0·08	−0·28, 0·12	−0·13	−0·35, 0·09
Q3	−0·15	−0·34, 0·04	−0·07	−0·28, 0·14
Q4	−0·18	0·37, 0·01	−0·18	−0·39, 0·02
Q5	−0·35	−0·54, −0·16	−0·30	−0·51, −0·10
	P trend‡ =<0·001		P trend‡=0·003	
	P deviation from trend§=0·72		P deviation from trend§=0·52	
Model 2‖				
Q1	Ref.	Ref.	Ref.	Ref.
Q2	−0·06	−0·26, 0·14	−0·13	−0·35, 0·09
Q3	−0·10	−0·29, 0·09	−0·06	−0·27, 0·15
Q4	−0·10	−0·29, 0·09	−0·17	−0·38, 0·04
Q5	−0·24	−0·44, −0·04	−0·28	−0·50, −0·07
	P trend‡ =0·02		P trend‡ =0·01	
	P deviation from trend§=0·74		P deviation from trend§=0·50	

Ref., referent values.

*cIMT model 1: *n* 1309 model 2: *n* 1298.

†PWV model 1: *n* 1061 model 2: *n* 1051.

‡Linear trend test, that is DASH quintiles fitted as continuous exposure in regression model.

§Log likelihood ratio test, that is testing DASH quintiles fitted as continuous exposure *v.* DASH quintiles fitted as categorical exposure.

‖Model 1 adjusted for socioeconomic position; model 2 additionally adjusted for BMI, smoking and physical activity.

Milestones 5:

Identify if and where diet can be harmonised between the studies

Harmonisation difficulties

- Sources of heterogeneity in dietary data between cohorts:
  - Instrument (Established dietary assessment instruments [5 cohorts only] vs. diet-related questions)
  - Instrument details (e.g. food lists used, prompts for assessment)
  - Differences in coding

Potential for harmonisation

- Dietary component to be harmonised depends on the research question e.g. nutrient vs. food group vs. dietary pattern?
- Collapsing more complicated variables to largest denomination between studies e.g. fruit and vegetable consumption?
- The first step: require relevant meta-data ✓
  - Determine specific instrument & observation period
  - Assess format of raw data & subcomponents measures
  - Determine assumptions made during processing
  - *E.g InterConnect consortium for fish consumption (12 studies)

*DAPA Measurement Toolkit harmonisation

Johnson 2015
PLOSMed

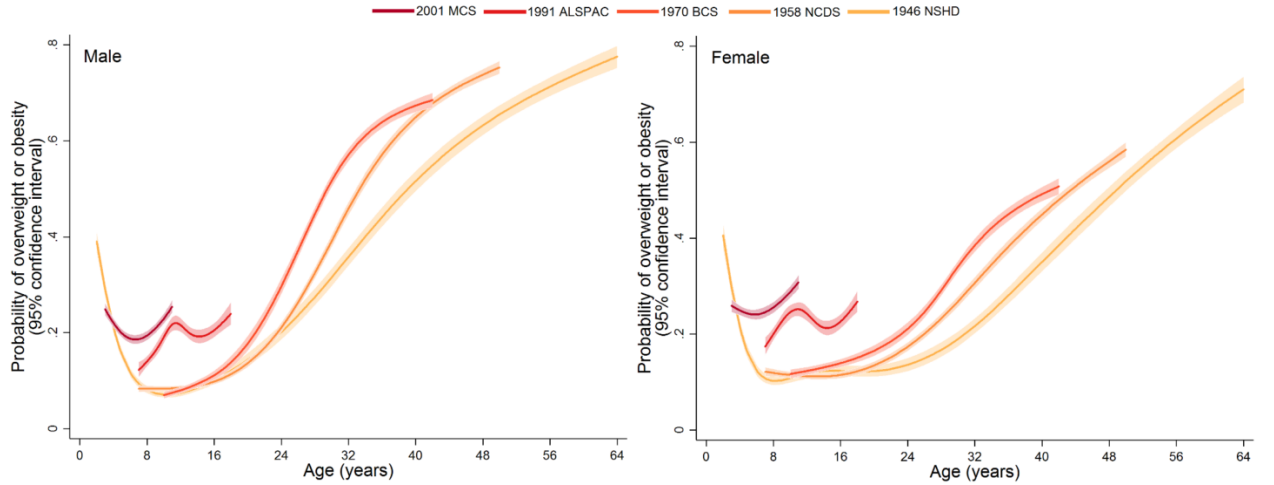


Fig 3. Trajectories of the probability of overweight or obesity (versus normal weight) from sex- and study-stratified multilevel logistic regression models. NSHD: Medical Research Council National Survey of Health and Development, NCDS National Child Development Study, BCS: British Cohort Study, ALSPAC: Avon Longitudinal Study of Parents and Children, MCS: Millennium Cohort Study.

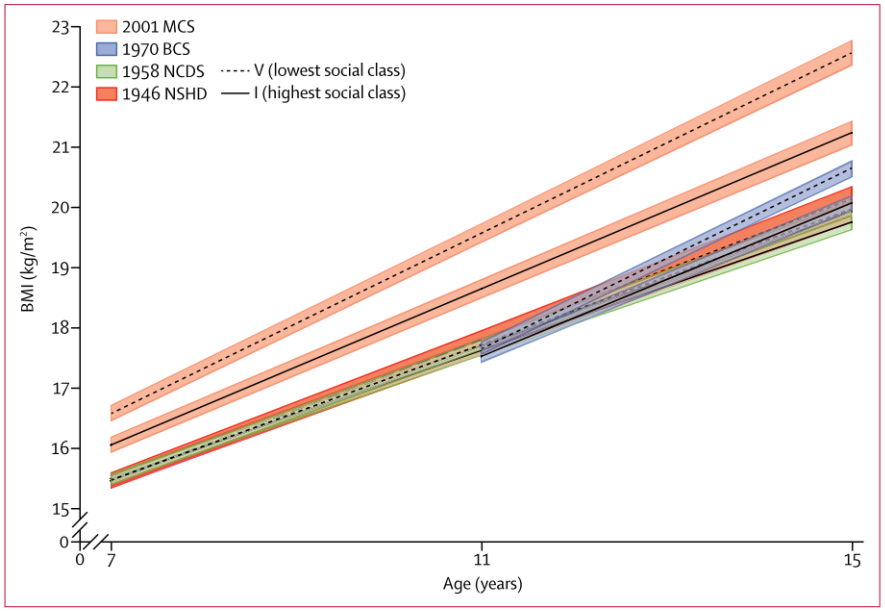


Figure 3: **BMI across childhood to adolescence by social class* in four British birth cohort studies**
Lines are estimated BMI and widths of the shaded area are 95% CIs at each age among women, estimated with multilevel general linear regression models (the appendix shows the full model estimates). BMI=body-mass index. NSHD=MRC National Survey of Health and Development. NCDS=National Child Development Study. BCS=British Cohort Study. MCS=Millennium Cohort Study. *Social class characterised by father's occupation.

Bann 2018
Lancet Public Health

Thank you,

Questions/suggestions?

Extra slides

NCDS dietary data

When	Method
1991 (33y)	Frequency of consumption of six food groups & type of bread
2000 (42y)	Frequency of consumption of 15 food groups
2003 (45y)	Frequency of consumption of 13 food groups & type of milk & supplement use

Unable to extract nutrients

MCS dietary data

When	Method
2001 (9 months)	Infant-feeding related questions
2004 (3y)	Infant-feeding related questions
2006 (5y)	• Infant-feeding related questions • Diet-related questions e.g. usual snacking foods, breakfast consumption, usual drinks, portions of fruit, meal regularity (answered by caregiver)
2008 (7y)	Diet-related questions e.g. usual snacking foods, breakfast consumption, usual drinks, portions of fruit, meal regularity (answered by caregiver)
2012 (11y)	Diet-related questions e.g. usual snacking foods, breakfast consumption, usual drinks, portions of fruit, meal regularity (answered by caregiver)
2015 (14y)	Diet-related questions e.g. breakfast consumption, portions of fruit and vegetables, type of bread & milk, soft drinks, (answered by child)

HCS dietary data

When	Method
1998-2003	129-item FFQ (modified from EPIC) covering previous 3 months (Robinson et al 2017; Robinson 2009)
2010-2011(East Hertfordshire sub-group follow-up)	FFQ

Extracting nutrients:

- Standard portion sizes allocated to each food item from published sources
- Frequency of a portion*nutrient content from food composition tables (McCance & Widdowson)

ALSPAC dietary data (Emmett 2009 EJCN)

Table 1 Nutrition data available in ALSPAC[a] whole cohort and CIF[b] subsample

Time point / *Age of study child (year of data sweep)*	*Mother's FFQ data*		*Partner's FFQ data*		*Child's data*		
	Sample	*Response rate, n (%)*	*Sample*	*Response rate, n (%)*	*Sample*	*Data type*	*Response rate, n (%)*
32 weeks gestation (1991/1992)	ALSPAC	12 423 (85)	ALSPAC (selected questions)	9960 (68)			
4 weeks (1991/1992)					ALSPAC	Infant feeding	12 353 (88)
4 months (1992/1993)					CIF	1-day diet record	964 (67)
6 months (1991/1993)					ALSPAC	Infant FFQ	11490 (82)
8 months (1992/1993)					CIF	3-day diet record	1131 (79)
15 months (1992/1994)					ALSPAC	Infant FFQ	11077 (79)
18 months (1994)					CIF	3-day diet record	1026 (72)
2 years (1994/1995)					ALSPAC	FFQ	10432 (75)
3 years (1995/1996)					ALSPAC	FFQ	10145 (73)
$3\frac{1}{2}$ years (1996)					CIF	3-day diet record	863 (60)
4 years (1996/1997)	ALSPAC	9504 (65)	ALSPAC	5102 (35)	ALSPAC	FFQ	9722 (70)
5 years (1997)					CIF	3-day diet record	772 (54)
6/7 years (1997/99)					ALSPAC	FFQ	8512 (61)
7 years (1998/2000)					ALSPAC	3-day diet record	7309 (54)
8/9 years (2000/02)	ALSPAC	7661 (53)	ALSPAC	3638 (25)	ALSPAC	FFQ	7965 (56)
10 years (2002/2003)					ALSPAC	3-day diet record	7474 (55)
12/13 years (2004/2006)	ALSPAC	6819 (47)	ALSPAC	3340 (23)	ALSPAC	FFQ part parent	6781 (48)
						Part child	6780 (48)
13 years (2004/2006)					ALSPAC	3-day diet record	6113 (45)

Abbreviations: ALSPAC, Avon Longitudinal Study of Parents and Children; CIF, Children in Focus; FFQ, food frequency questionnaire.
[a]ALSPAC—mothers and partners ($n = 14\,541$ pregnancies), children ($n = 14\,062$ live births, 13 988 alive at 1 year), clinic visits for whole cohort ($n = 13\,602$ available at 7 years).
[b]CIF—randomly selected subsample ($n = 1432$ attended at least once).