An automated protocol for assigning address-level air pollution exposure for longitudinal birth cohort studies

Daniela Fecht, Kevin Garwood, John Gulliver

MRC–PHE Centre for Environment & Health, Imperial College London, UK

Small Area Health Statistics Unit

MRC-PHE
Centre for Environment & Health

ISES 2016, Utrecht

Public Health England

MRC | Medical Research Council

Imperial College London

King's College London

Effects of early life exposure to particulates on respiratory health through childhood and adolescence: ALSPAC

AIM:

To investigate long-term effects of exposure to particulate air pollution (PM_{10}) in the prenatal period and early infancy on:
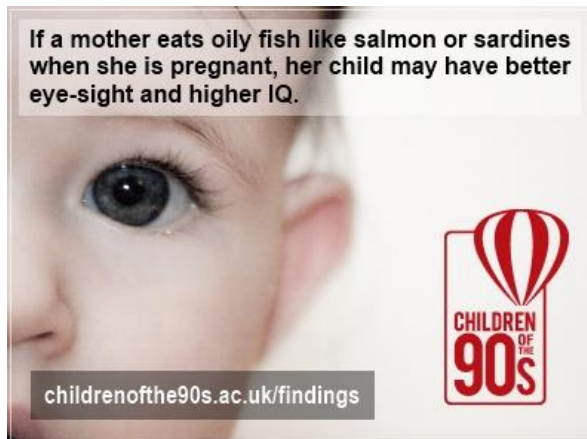
- sensitisation to aeroallergens,
- respiratory symptoms,
- doctor-diagnosed asthma,
- lung function,

in children through to adolescence.

Avon Longitudinal Study of Parents and Children (ALSPAC)

One of the largest and best characterised birth cohort studies in the world

If a mother eats oily fish like salmon or sardines when she is pregnant, her child may have better eye-sight and higher IQ.
childrenofthe90s.ac.uk/findings

Teenage girls who are too thin may be putting their bones at risk.
childrenofthe90s.ac.uk/findings

Just fifteen minutes of daily vigorous exercise can reduce the risk of obesity by half.
childrenofthe90s.ac.uk/findings

www.bristol.ac.uk/alspac

University of BRISTOL
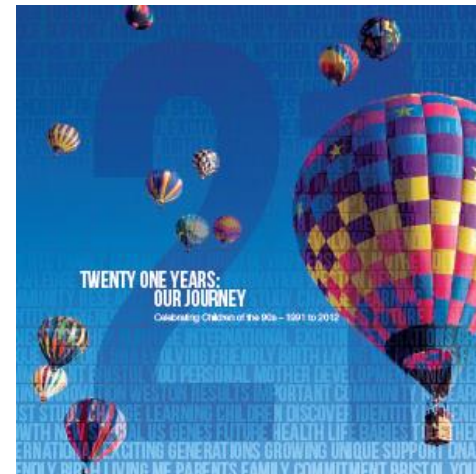
25 YEARS CHILDREN OF THE 90s

Avon Longitudinal Study of Parents and Children (ALSPAC)

One of the largest and best characterised birth cohort studies in the world
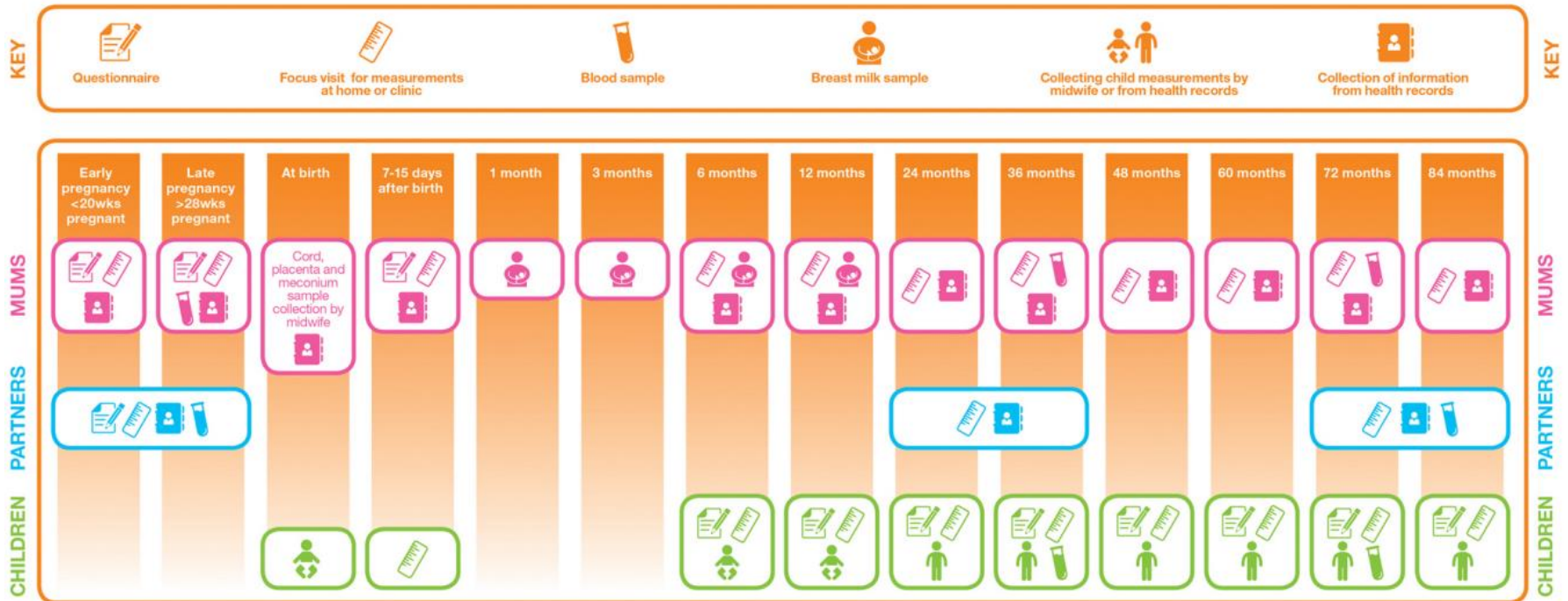
- Longitudinal birth cohort recruited in pregnancy in the early 90s in Avon (Bristol), UK (babies born 1st April 1991 – 31st December 1992)

ALSPAC Media Images

University of BRISTOL

25 YEARS CHILDREN OF THE 90s

ALSPAC

Avon Longitudinal Study of Parents and Children (ALSPAC)

Small Area Health
Statistics Unit

Avon Longitudinal Study of Parents and Children (ALSPAC)

One of the largest and best characterised birth cohort studies in the world

- Longitudinal birth cohort recruited in pregnancy in the early 90s in Avon (Bristol), UK (1st April 1991 – 31st December 1992)

- 14,541 women agreed to take part and gave birth to 14,062 live born infants of which 13, 985 survived to the end of the first year

- Followed intensely with self-completed questionnaires (until the age of 7 years)

- From age 7 years dedicated research clinic, annually for measurements of health, psychological and development outcomes:

    lung function test at
   ~8 years and ~15 years

ALSPAC Media Images

Air pollution exposure

As part of the study on **Effects of early life exposure to particulates on respiratory health through childhood and adolescence** we aimed to:

1) determine PM_{10} (particles ≤ 10 μm in diameter) exposures during the critical periods of lung development:
   - Fetal period (pregnancy trimester specific)
   - Early infancy (0–6 months) and
   - Late infancy (7–12 months)

2) long-term (i.e. periods of 1–year or more) exposures up to age ~8 and ~15 when assessments of asthma, lung function and bronchial responsiveness, and sensitization to allergens were made at follow-up clinics.

ALSPAC Media Images

Air pollution exposure
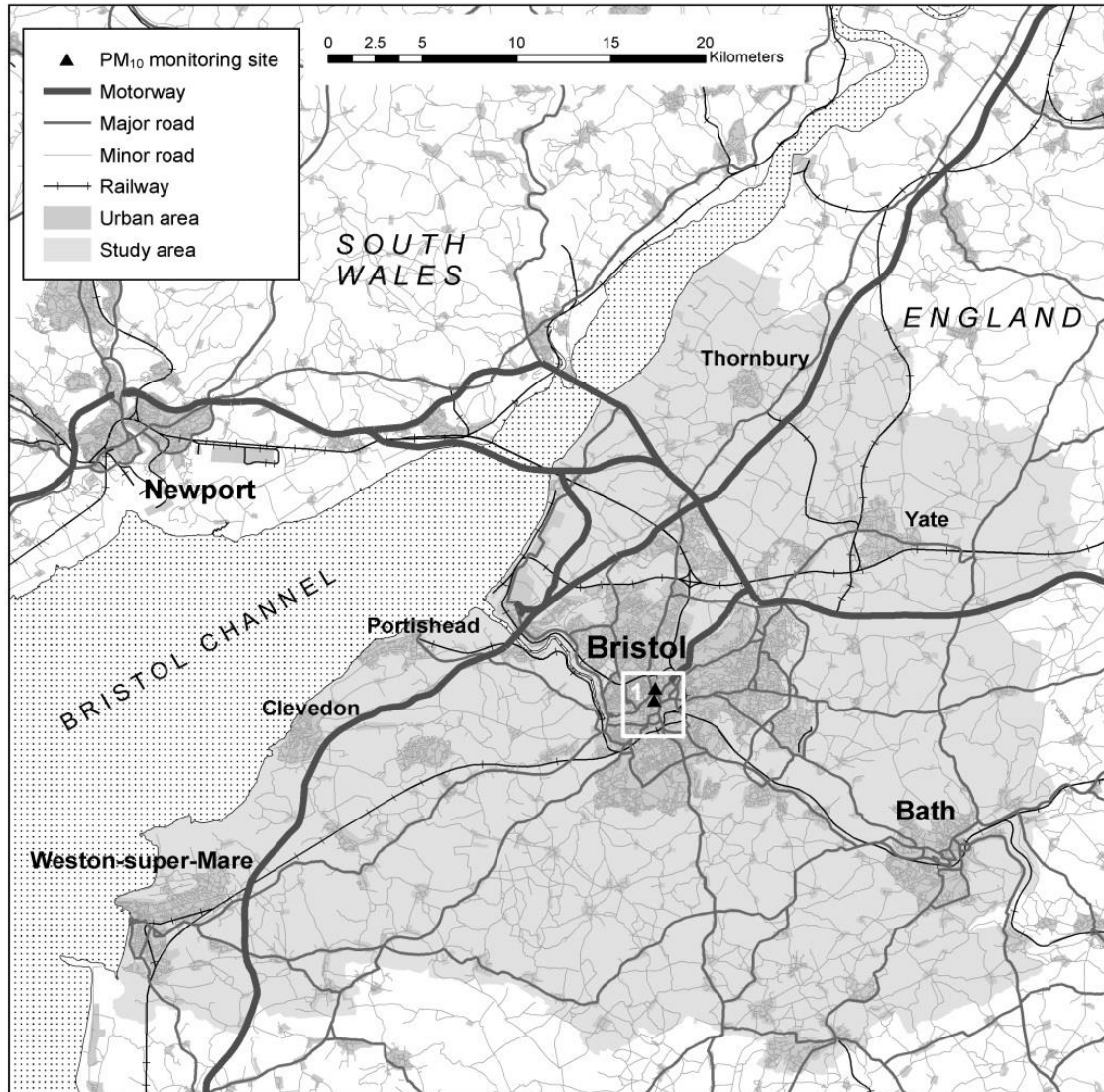
- Only up to 1 monitoring site across study area at any time

- Model daily PM_{10} exposure using a combination of local and regional dispersion model for residential address of each child

- Account for residential mobility across different life stages
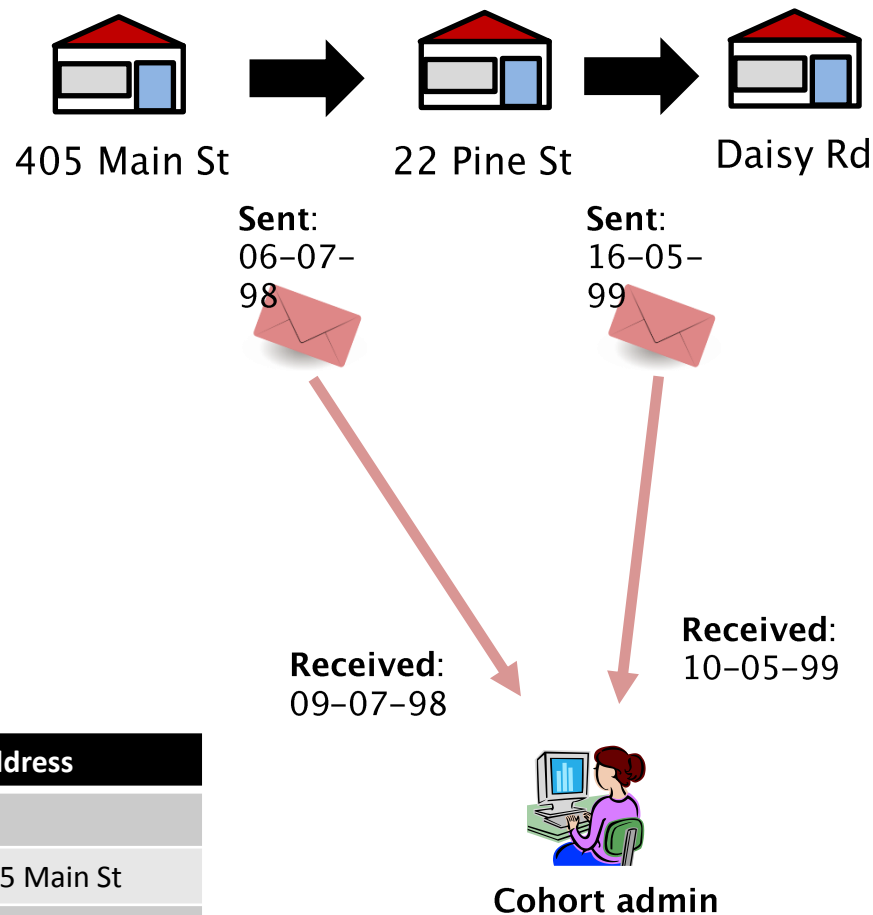
Residential history

- To account for residential mobility during pregnancy, information on residential address history is essential.

- Some countries collect detailed information on residential mobility as part of national registries.

- Cohort studies may need to collect such information retrospectively which might be resource intensive and prone to recall bias.

- An alternative source is the cohort contact database; administrative database designed to audit current addresses, not to track past ones.

- **ALGAE Protocol: AL**gorithms for **G**enerating **A**ddress histories and **E**xposures

Residential history

Reconstructing address history for each cohort study member using the contact database:

- A person needs to occupy exactly one location for each day of their exposure period

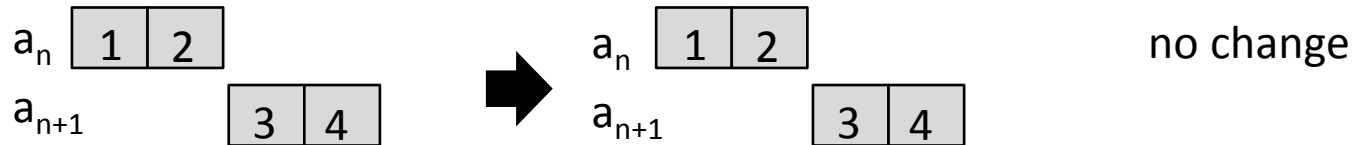- Gaps and overlaps may be significant sources of exposure misclassification

405 Main St → 22 Pine St → Daisy Rd

Sent: 06-07-98
Sent: 16-05-99

Received: 09-07-98
Received: 10-05-99

Cohort admin

Inside the tables of the administrative system

Person ID	Start date	End date	Address
34657	…	…	…
34657	09-07-98	…	405 Main St
34657	10-05-99	18-06-99	22 Pie Street
34657	15-06-99	04-04-01	22 Pine St
34657	05-04-01		Daisy Rd
34657	…	…	…

Address cleaning

- **Start date stronger signal then end data**
- **Prefer to favour start date**

Case 1: Contiguous

a_n [1 | 2] ➡ a_n [1 | 2]   no change

a_{n+1} [3 | 4] a_{n+1} [3 | 4]

Case 2: Gap

a_n [1 | 2 | 3] ➡ a_n [1 | 2]

a_{n+1} [4 | 5] a_{n+1} [3 | 4 | 5] $a_{n+1 start} = a_{n+1 end} + 1$

Case 3: Overlap

a_1 [1 | 2] ➡ a_n [1 | ⬅] $a_{n\ end} = a_{n+1\ start} - 1$

a_2 [2 | 3] a_{n+1} [2 | 3]

Case 4: Duplicate

a_1 [1 | 2] ➡ a_n

a_2 [1 | 2 | 3] a_{n+1} [1 | 2 | 3]
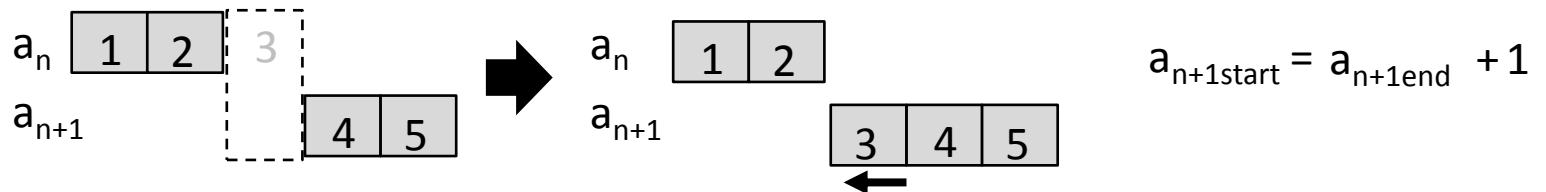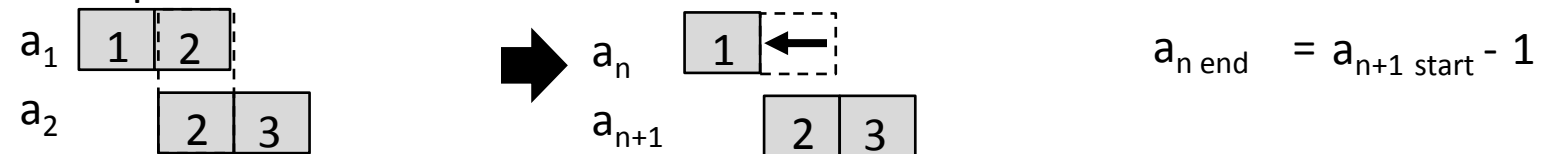
Address cleaning

Assessing exposure measurement error

Gap:

a_n | 10 | 11 |
| 12 | 13 |
a_{n+1} | 14 | 15 | 16 |

→

days of contention

a_n | 10 | 11 | 12 | 13 | ← **opportunity cost exposure**

a_{n+1} | 12 | 13 | 14 | 15 | 16 |

assigned exposure

Overlap:

a_n | 10 | 11 | 12 |
| 12 |
a_{n+1} | 12 | 13 | 14 |

→

a_n | 10 | 11 | 12 | **opportunity cost exposure**

| 12 |

a_{n+1} | 12 | 13 | 14 |
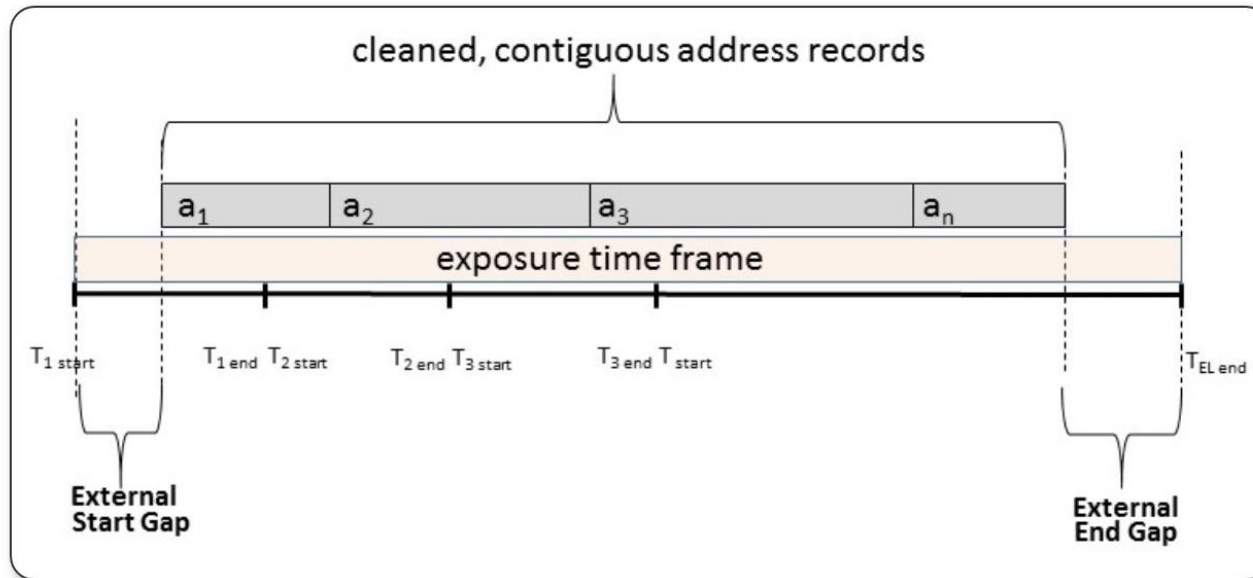
assigned exposure
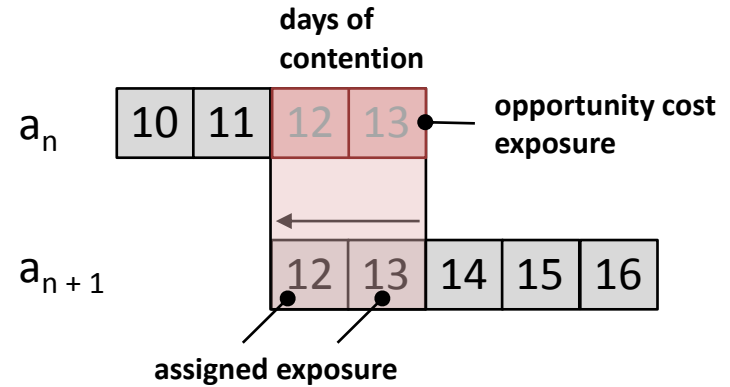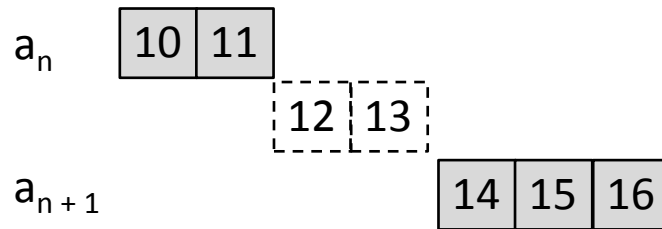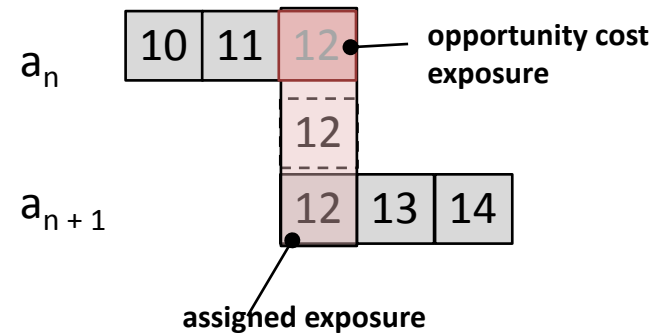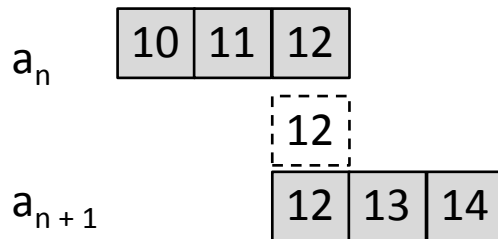
Trimester-weighting

Exposure calculation:
- mean
- median
- cumulative

Pregnancy trimester and early life calculations:

	Start date	End date
Trimester 1	Date of conception	Date of conception + 91 days
Trimester 2	Date of conception + 92 days	Date of conception + 183 days
Trimester 3	Date of conception + 184 days	Date of birth – 1day
Early infancy	Date of birth	Date of birth + 6 months – 1 day
Late infancy	Date of birth + 6 months	Date of birth + 1 year – 1 day

Results

- We reconstructed residential address histories for 14,062 pregnant women based on 45,771 unique addresses

- 4,059 women moved out of the study area and were excluded from the analysis

- We assigned trimester–specific PM_{10} exposures for the remaining 10,003 pregnancies

- 1,447 women (14.5%) changed address during their pregnancy

- 31% of pregnancies had no days changed, 50% had one day changed.
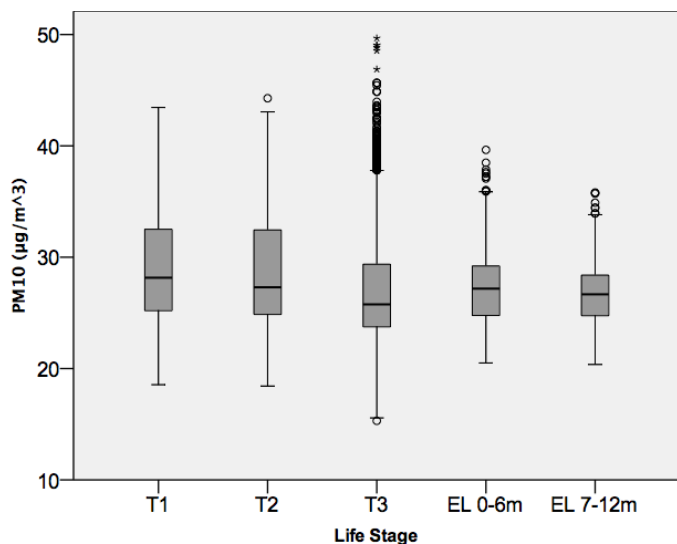
PM_{10} exposure:

	Mean PM_{10} ($\mu g/m^3$)	StDev ($\mu g/m^3$)
T1	28.8	4.56
T2	28.5	4.68
T3	26.9	4.58

ALGAE

The ALGAE Protocol

ALorithms for **G**enerating address histories and **E**xposures

An automated protocol for assessing early life exposures in longitudinal cohorts

Home Methodology Protocol Design Using ALGAE Adapting the Code Testing About Us

Adapt the Code

We expect that there are three main ways that projects may want to adapt the ALGAE protocol to better suit their needs:

- adjust life stage boundaries for early life analyses
- add or remove life stages for the later life analysis
- change one or more of the pollutants

Here we provide some suggestions about how you would do these tasks. Please note however that any changes you make to the code will likely cause the test suites to fail.

How do I change the life stages?

The easiest way to learn how to change the life stages is to look for the tag "#CHANGE_LIFE_STAGES" in the code base. It will take you to the part of the code in the function `early_calc_life_stages` that specifies default stages of T1, T2, T3 and EL.

Definitions of life stages will appear like the following snippet of code:

```
SELECT
    person_id,
    'T2' AS life_stage,
    birth_date,
    (conception_date + INTERVAL '92 days')::date AS start_date,
    (conception_date + INTERVAL '183 days')::date AS end_date
```

In this case, if you wanted to change the name of the life stage `T2` to `Trimester 2`, you would first do a global search-and-replace on the phrase "'T2'" with "'Trimester 2'". If you wanted to change the temporal boundaries, you would adjust the numbers in the phrases `'92 days'` and `'183 days'`. **Remember: however you change the temporal boundaries for a life stage, ensure that they are temporally contiguous with other life stages**.

Please exercise care when you change the life stages in the early life analysis. ALGAE performs checks that fix T2 and T3 periods so that they do not overlap with

ALGAE

SAHSU:

Kevin Garwood

John Gulliver

Anna Hansell

Clare Pearson

Paul Elliott

ALSPAC:

Iain Bickerstaffe

Olly Butters

Tim Morris

Kate Northstone

John Henderson