Global data for diabetes and obesity research

InterConnect

Global initiative on gene environment interactions in diabetes / obesity in specific populations

Grant agreement no: 602068

Figures to accompany
Final Report, 29 November 2017

Figure 1: Global variation in the prevalence of type 2 diabetes, moving from within-population investigation to the study of between-population differences

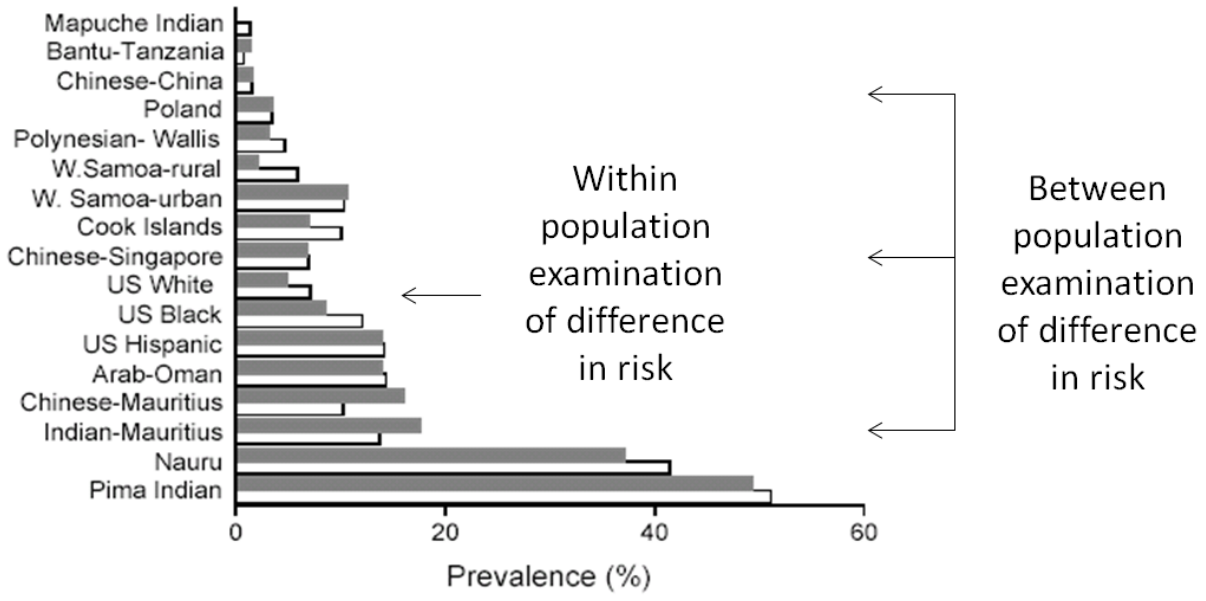


Figure 2: Illustration of the main domains of the InterConnect project

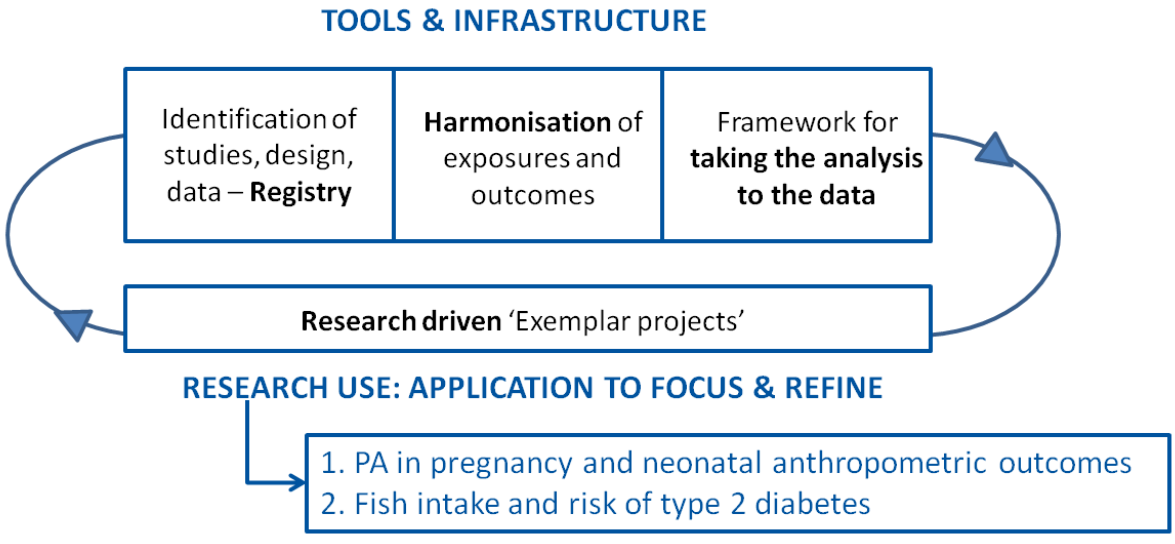


Figure 3: Screen shot of the searchable online registry

Short Name	Name	Study Design	Actual number of participants recruited to the study	Country of residence
MEC	Multiethnic Cohort Study	Prospective cohort study	215 251	United States
SWS	Southampton Women's Survey	Prospective cohort study	12 583	United Kingdom
	Healthy Start study	Prospective cohort study	2 820	United States
ALSPAC	Avon Longitudinal Study of Parents and Children	Prospective cohort study	14 541	United Kingdom
AHS	Agricultural Health Study	Prospective cohort study	89 655	United States
ARIC	Atherosclerosis Risk in Communities Study	Prospective cohort study	15 792	United States
DNBC	Danish National Birth Cohort	Prospective cohort study	101 042	Denmark
EPIC - Turin	European Prospective Investigation into Cancer and Nutrition - Turin	Prospective cohort study	10 604	Italy
NHS I	Nurses Health Study I	Prospective cohort study	121 700	United States
NOMAS	The NOrthern MAnhattan Study	Prospective cohort study	3 298	United States

1 2 3 4 5 6 7 8 9 ... next ›

Figure 4: Geographic spread of studies within the registry; diameter is proportional to the number of studies

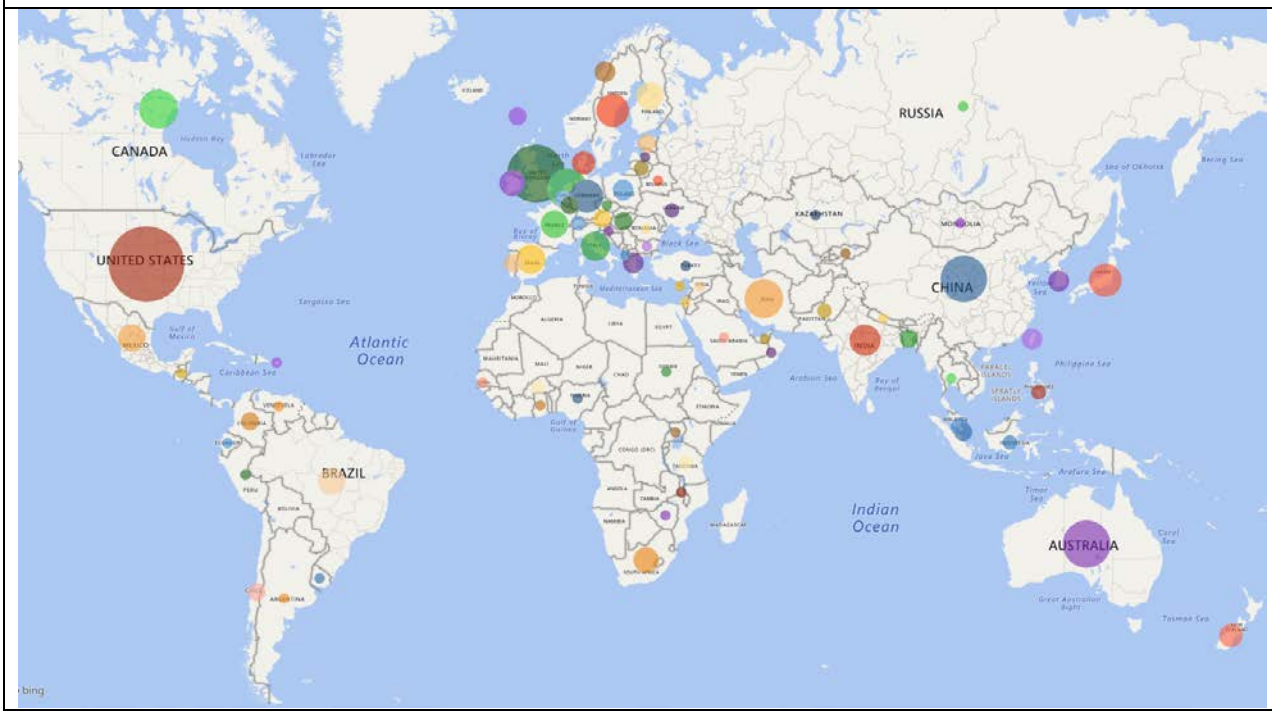
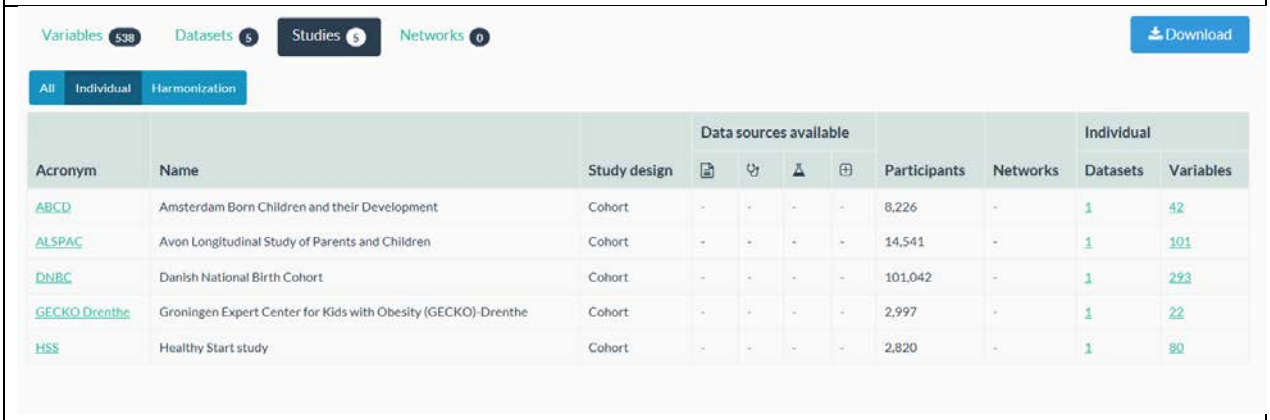


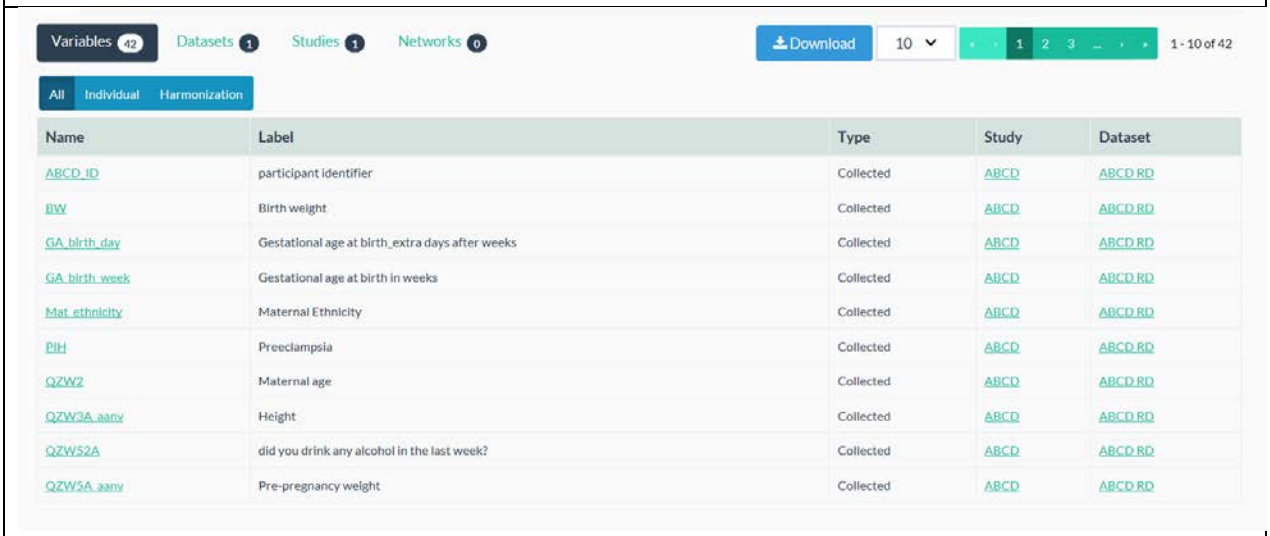
Figure 5: Catalogue the harmonisation algorithms from the exemplar projects so that they are widely available for re-use by others. An overview of the resource is provided below, through screen-shots of the website to illustrate the information therein; these focus on the exemplar that investigated the role of physical activity during pregnancy on neonatal anthropometric outcomes.

a. Studies participating in the exemplar:



Acronym	Name	Study design	Data sources available				Participants	Networks	Individual Datasets	Individual Variables
ABCD	Amsterdam Born Children and their Development	Cohort	-	-	-	-	8,226	-	1	42
ALSPAC	Avon Longitudinal Study of Parents and Children	Cohort	-	-	-	-	14,541	-	1	101
DNBC	Danish National Birth Cohort	Cohort	-	-	-	-	101,042	-	1	293
GECKO Drenthe	Groningen Expert Center for Kids with Obesity (GECKO)-Drenthe	Cohort	-	-	-	-	2,997	-	1	22
HSS	Healthy Start study	Cohort	-	-	-	-	2,820	-	1	80

b. Raw variables within one study:



Name	Label	Type	Study	Dataset
ABCD_ID	participant identifier	Collected	ABCD	ABCD RD
BW	Birth weight	Collected	ABCD	ABCD RD
GA_birth_day	Gestational age at birth_extra days after weeks	Collected	ABCD	ABCD RD
GA_birth_week	Gestational age at birth in weeks	Collected	ABCD	ABCD RD
Mat_ethnicity	Maternal Ethnicity	Collected	ABCD	ABCD RD
PIH	Preeclampsia	Collected	ABCD	ABCD RD
QZW2	Maternal age	Collected	ABCD	ABCD RD
QZW3A_aany	Height	Collected	ABCD	ABCD RD
QZW52A	did you drink any alcohol in the last week?	Collected	ABCD	ABCD RD
QZW5A_aany	Pre-pregnancy weight	Collected	ABCD	ABCD RD

c. Harmonisation potential across participating studies:

Harmonization

Click on each status icon to get more details on the corresponding harmonization results :

Download

Undetermined - the harmonization potential of this variable has not yet been evaluated.
Complete - the study assessment item(s) (e.g. survey question, physical measure, biochemical measure) allow construction of the variable as defined in the dataset.
Impossible - there is no information or insufficient information collected by this study to allow the construction of the variable as defined in the dataset.

Showing 1 to 25 of 29 entries

Variable	ABCD	DNBC	GECKO Drenthe	ALSPAC	HSS
BIRTH_WEIGHT	✔	✔	✔	✔	✔
LTPA_EE_filt	✔	✔	✖	✔	✔
BIRTH_WEIGHT_SGA	✔	✔	✔	✔	✔
VIG_3_filt	✖	✔	✖	✖	✔
MACROSOMIA	✔	✔	✔	✔	✔
ALCOHOL	✔	✔	✔	✔	✔
PREECLAMPSIA	✔	✔	✔	✔	✔
VIG_filt	✔	✔	✖	✔	✔

d. Example harmonisation algorithms for one variable:

Harmonized Variable	Study	Data Collection Event	Status	Comment
BIRTH_WEIGHT -- ABCD	ABCD	Pregnant Women -- ABCD data collection	✔	-
BIRTH_WEIGHT -- ALSPAC	ALSPAC	Population -- ALSPAC data collection	✔	-
BIRTH_WEIGHT -- DNBC	DNBC	1st DNBC cohort -- DNBC data	✔	-
BIRTH_WEIGHT -- GECKO Drenthe	GECKO Drenthe	Sampling frame 1 -- GECKO data collection	✔	-
BIRTH_WEIGHT -- HSS	HSS	Sampling Frame 2 -- HSS DCE	✔	-

Hide Harmonization Algorithms

BIRTH_WEIGHT -- abcd

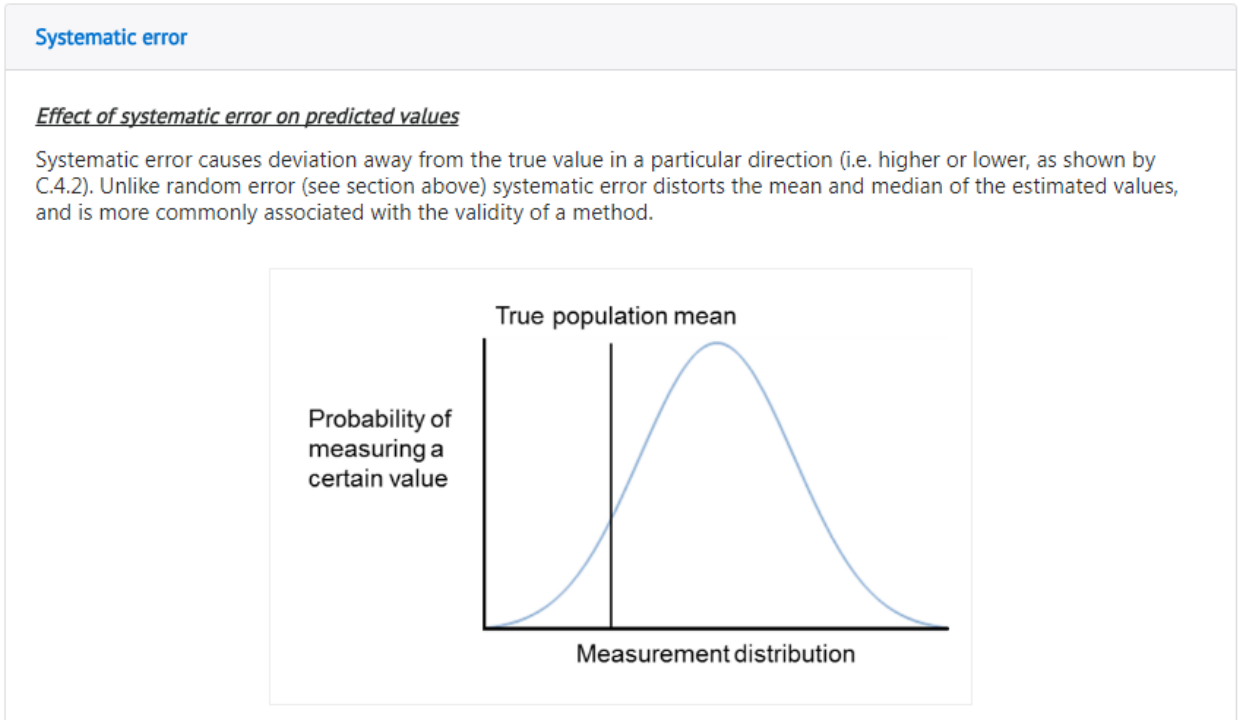
```
birthweight = $('BW').value(); if (birthweight != -1){ if (birthweight < 1000){ output = 1000; } else if (birthweight > 6000){ output = 6000; } else { output = birthweight; } } else { output = -1; } output;
```

BIRTH_WEIGHT -- alspac

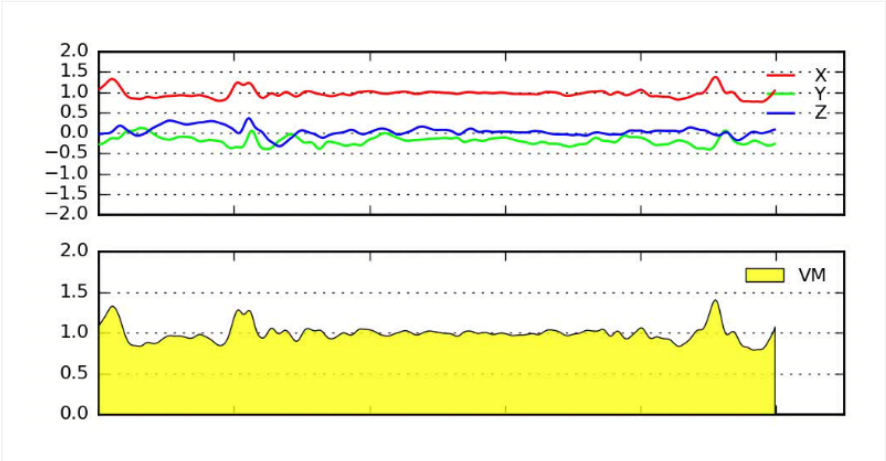
```
if ($('kz030').isNull().value() || $('kz030') == -1 || $('kz030') == -11 || $('kz030') == -10 || $('kz030') == -9999){ weight = -1; } else { weight = $('kz030').value(); }

if (weight != -1){ if (weight < 1000) { output = 1000; } else if (weight > 6000){ output = 6000; } else { output = weight; } } else { output = -1; }
```

Figure 5: Screenshots from the DAPA toolkit to illustrate content

a. Introduction to basic concepts of assessment in population health sciences



b. Inventories of subjective and objective methods; An example for accelerometry methods of assessing physical activity explaining one of the procedures involved when deriving physical activity variables



c. Method selector matrices; A section of the method selector matrix which summarises the suitability of methods to assess different dimensions of physical activity.

Dimension	Questionnaires	Diaries and logs	Simple Pedometer	Time Series Pedometer	Accelerometer	Heart Rate Monitor
Duration	✔	✔		✔	✔	✔
Intensity	✔	✔		✔	✔	✔
Frequency	✔	✔		✔	✔	✔
Volume	✔	✔	✔	✔	✔	✔
Total physical activity energy expenditure	✔	✔			✔	✔
Type	✔	✔			Maybe	
Timing of bouts of activity (i.e. pattern of activity)		✔		✔	✔	✔
Domain	✔	✔				
Contextual information (e.g. location)	✔	✔				
Posture	✔	✔			✔	
Sedentary behaviour	✔	✔		✔	✔	✔

d. Instrument library; A section of the instrument library indicating availability of resources for different instruments for assessing diet and physical activity.

	Access to instrument	Description	Design information	Output information	Links to resources	Information on validity	Examples of use
EPIC Physical Activity Questionnaire (second version)	Link to PDF	✔	✔	✔	✔	✔	✔
EPIC-Norfolk Food Frequency Questionnaire	Link to PDF	✔	✔	✔	✔	✔	✔
24-hour Physical Activity Recall		✔					
One-week Physical Activity Recall		✔					
7-Day Physical Activity Recall	Link to PDF	✔			✔		

e. Overview and case studies describing data harmonisation: Example of a page describing concepts of data harmonisation.

What is data harmonisation?

(Note: British English spelling - 'harmonisation'; North American English spelling - 'harmonization')

There are many definitions of data harmonisation, but a good working definition is provided by Maelstrom Research.

Harmonisation involves achieving or improving comparability of similar measures collected by separate studies or databases for different individuals. Some research programs foster prospective implementation of harmonised measures to collect data across studies, while others turn their efforts to retrospective harmonisation and co-analysis of existing datasets.

In summary, harmonisation seeks to bring together various types, levels and sources of data, which represent measurement of the same latent construct(s), in such a way that they can be made compatible and comparable (see Figure C.8.1 for example on wine consumption).

Harmonisation differs from standardisation in that it does not impose a single methodology or norm, but rather seeks to find ways of integrating or making "an agreeable effect" from information gathered through disparate methodologies [10].

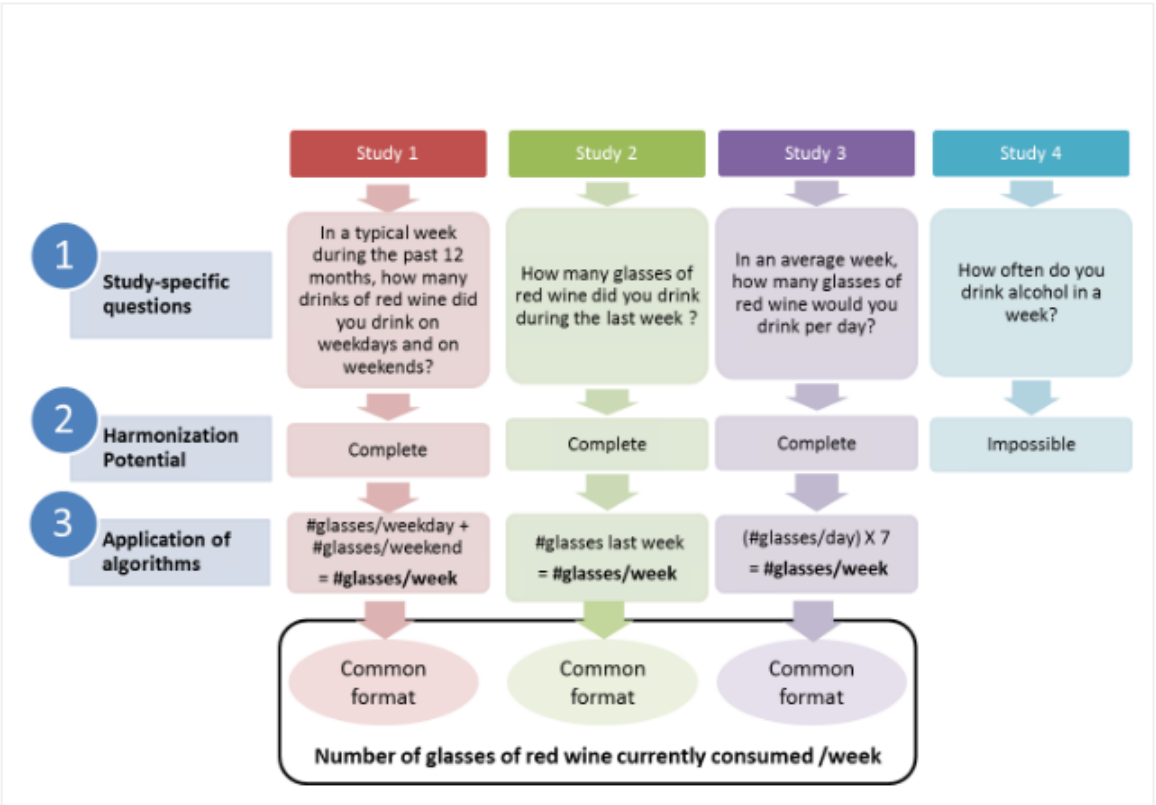


Figure 7: Federated meta-analysis, data stays within the governance structure of the cohort

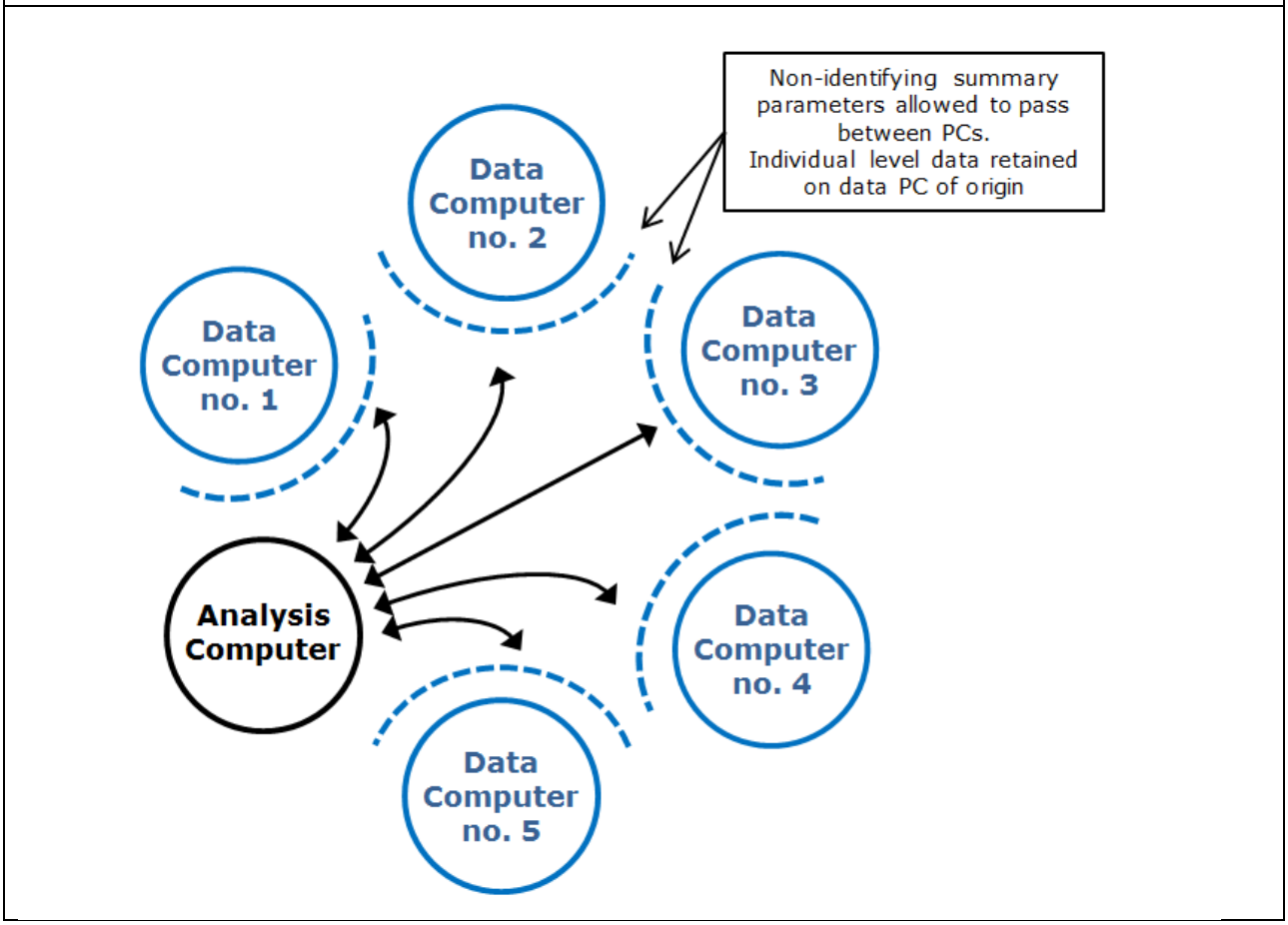


Figure 8: Cohorts participating in the first exemplar question on the association between physical activity in pregnancy and neonatal outcomes

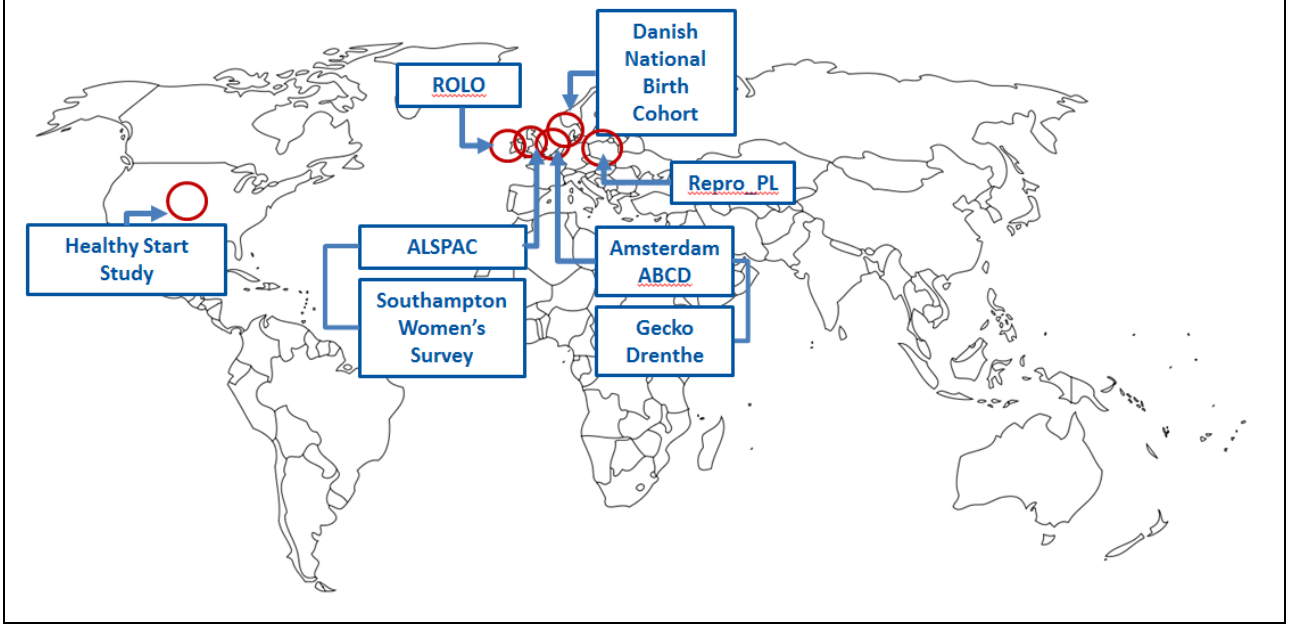


Figure 9: Illustrative results from the first exemplar question, derived from meta-analysis of individual participant data without direct access to the data

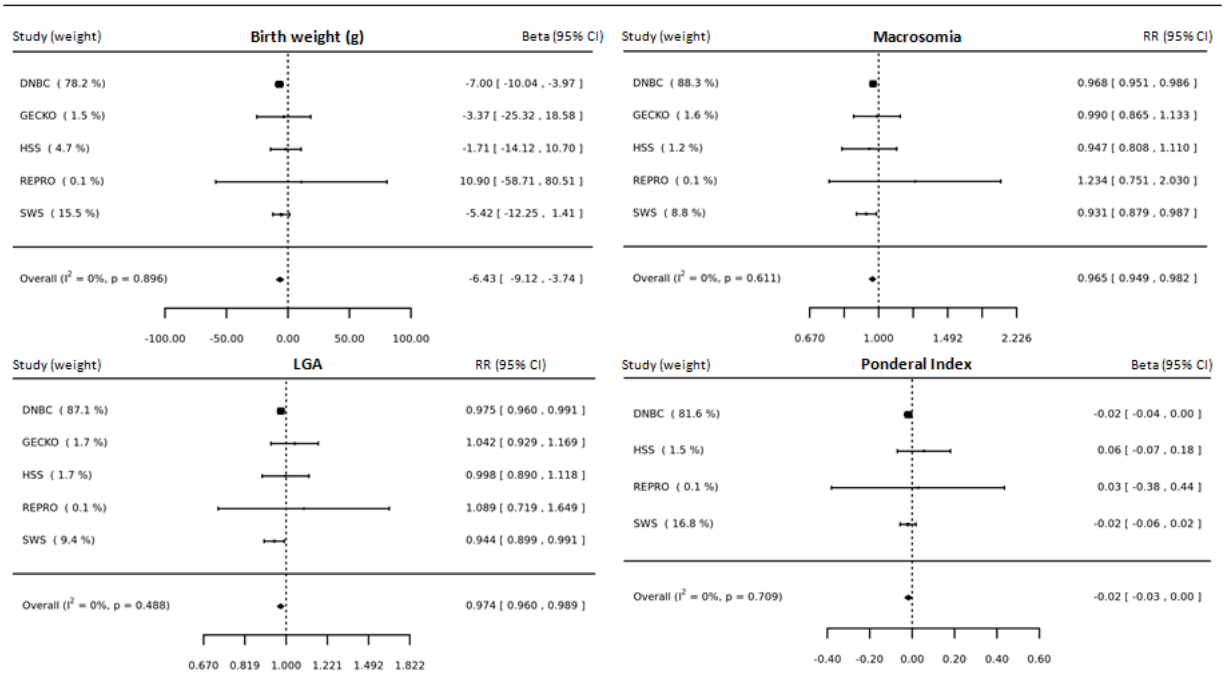


Figure 1. Forest plots for third trimester moderate to vigorous activity associated with birth measurements . All associations were adjusted for gestational age, sex, parity, maternal age, smoking, alcohol, maternal education and ethnicity. N=58,820 except for Ponderal Index (N=57,172)