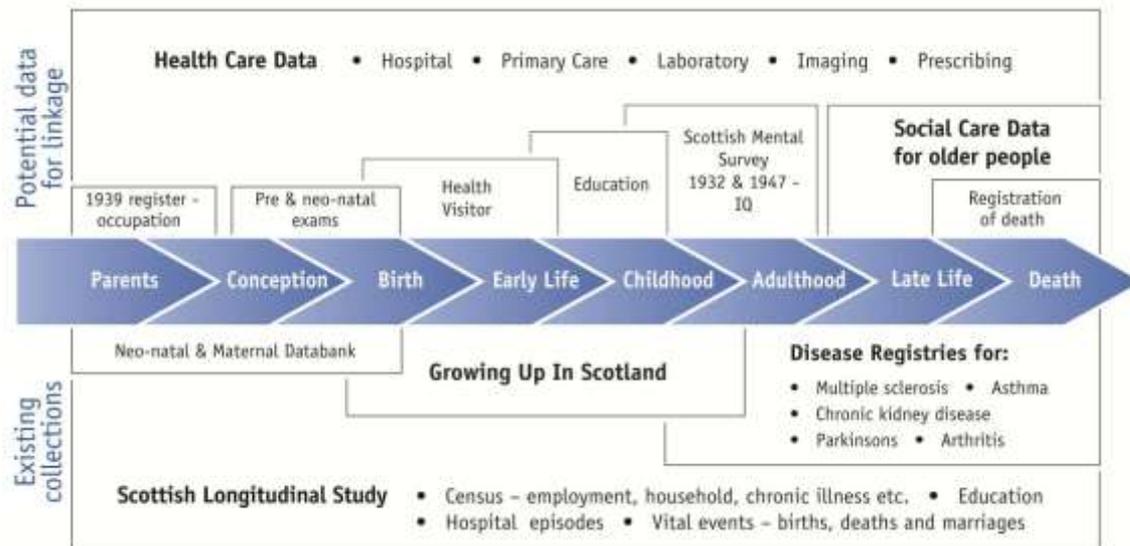


# Big Data and Health Research

# Health sector data

- Routine administrative data
  - Hospital admissions
  - Birth & Death certificates
  - Outpatient attendances
  - GP consultations
- Clinical data
  - lab tests, immunisation, drug prescriptions etc
- Disease / operation registers
- Health surveys



# Research

```
graph TD; Research --> Epidemiology; Research --> InterventionStudies[Intervention Studies];
```

## Epidemiology

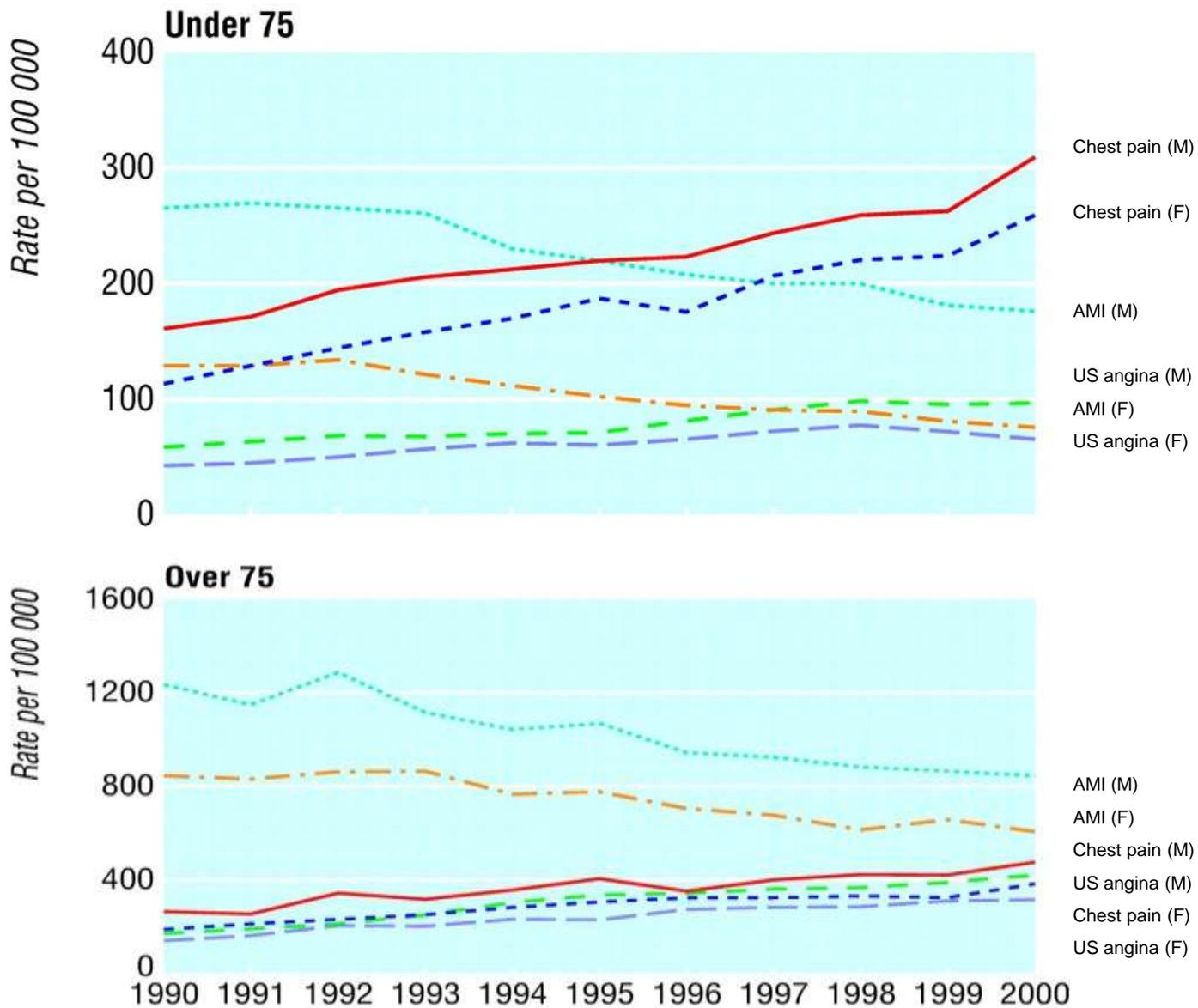
- How much ?
- Who gets it ?
- What is it associated with ?

## Intervention Studies

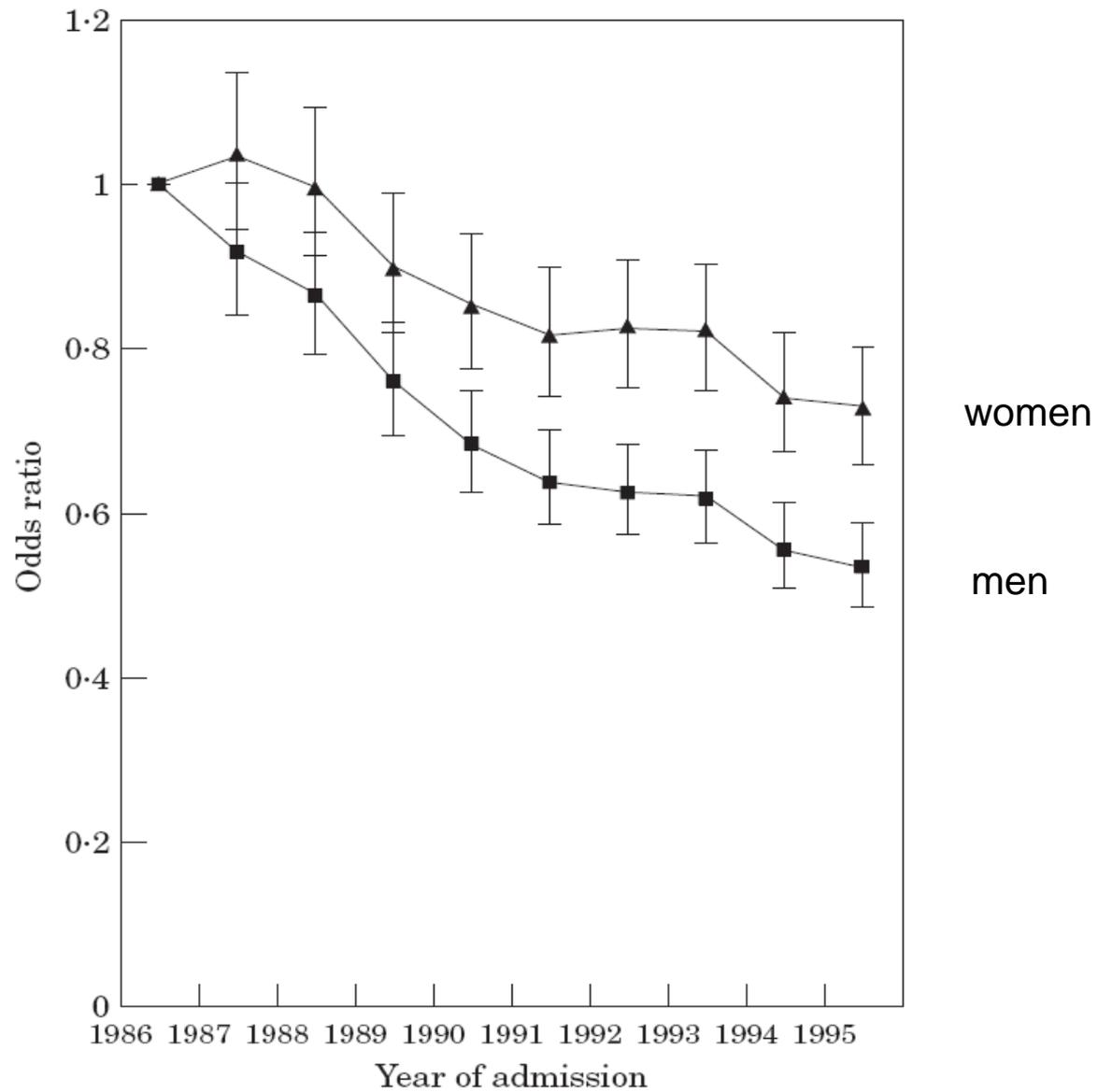
Can you change it ?

**Epidemiology ....**

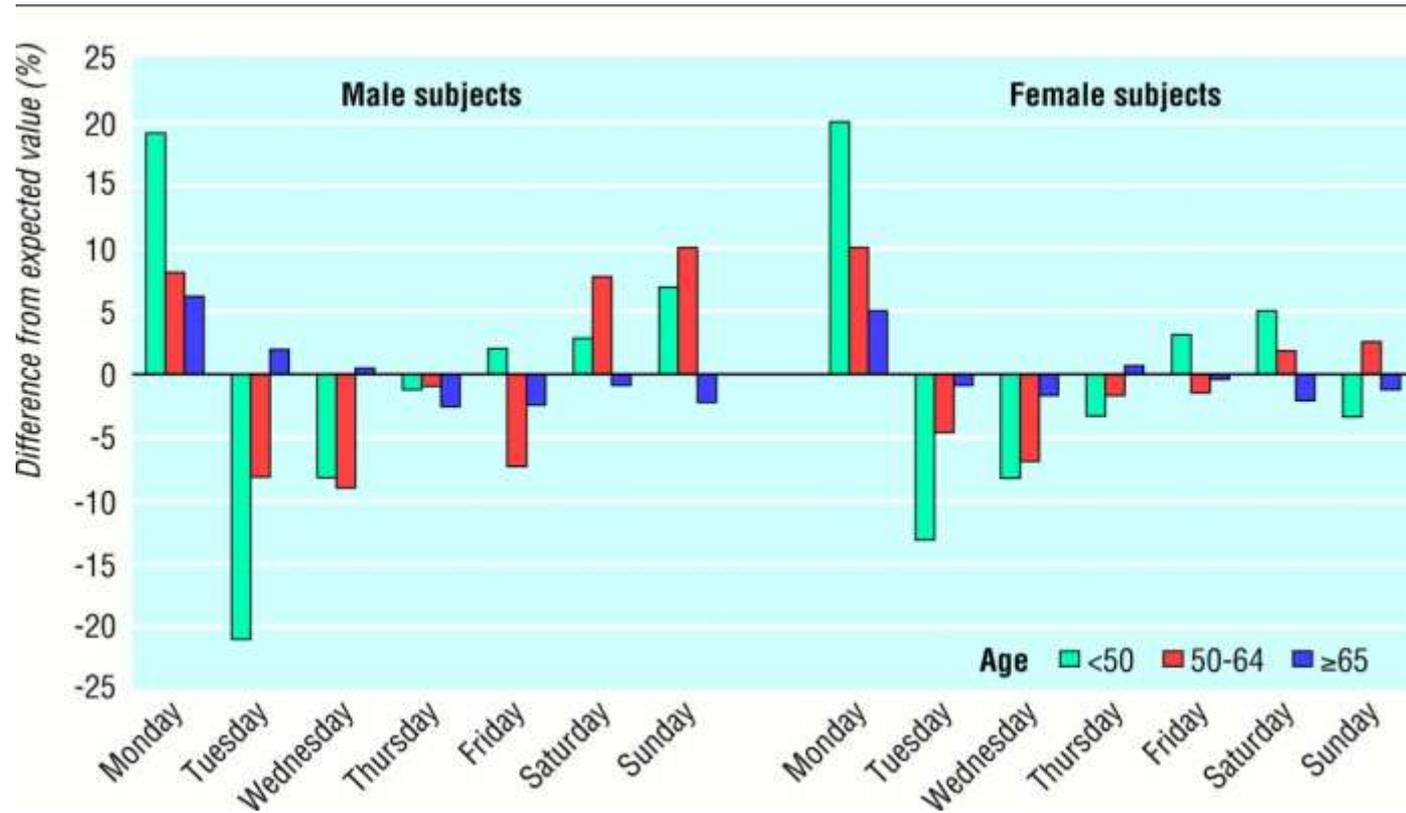
## Time trends in hospitalisations for suspected heart disease



## Time trends in 30 day survival following heart attack

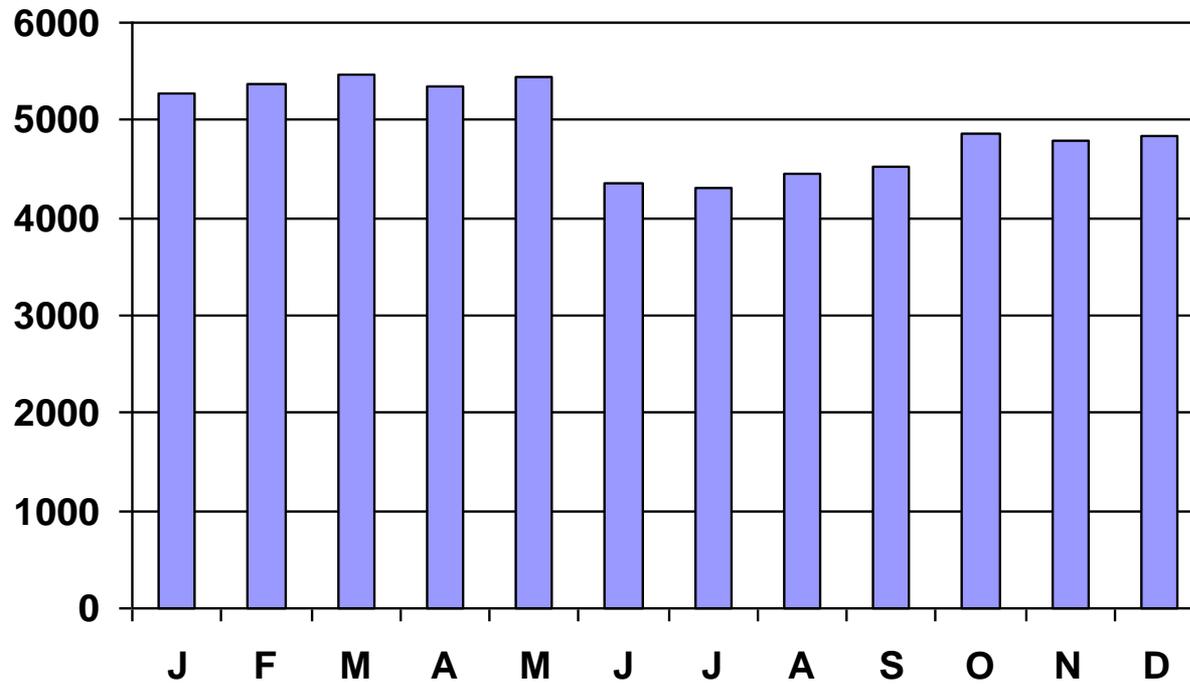


## Heart attacks by day of week

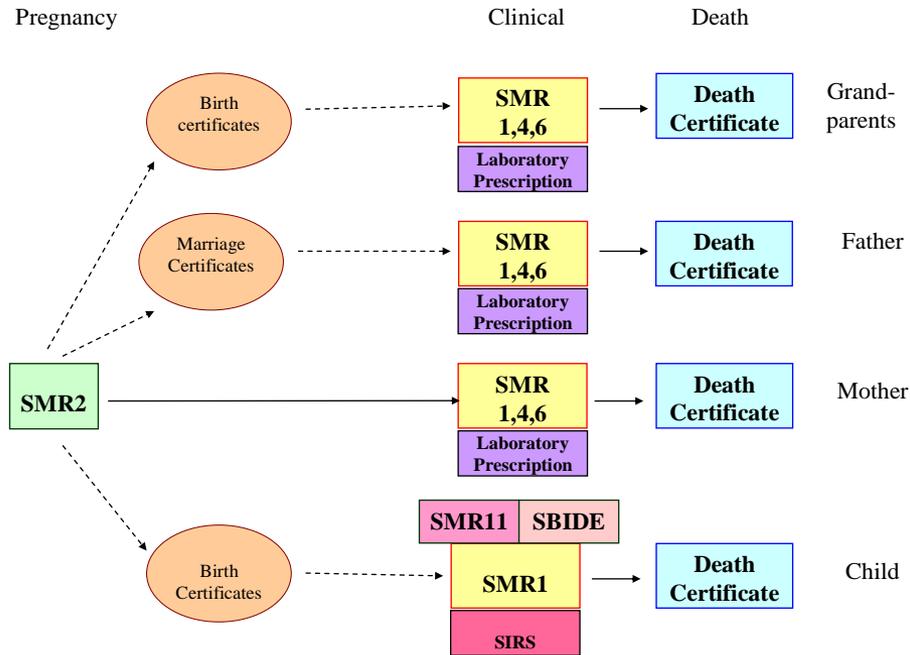


## Children with special educational needs by month of conception

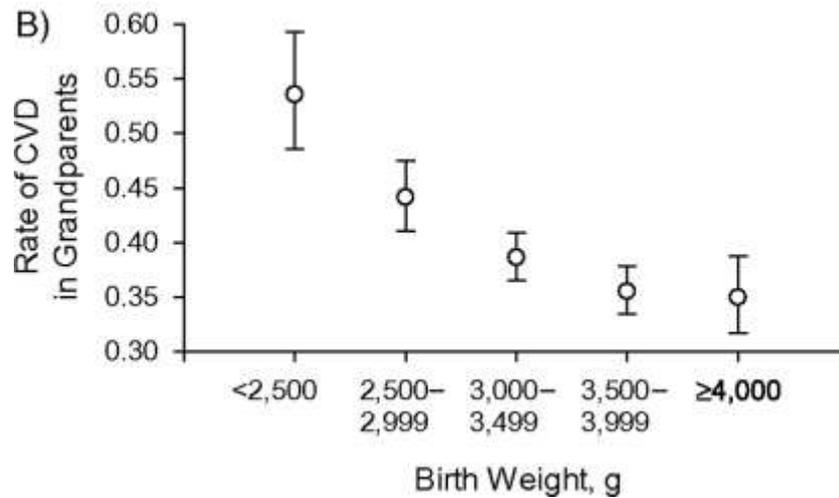
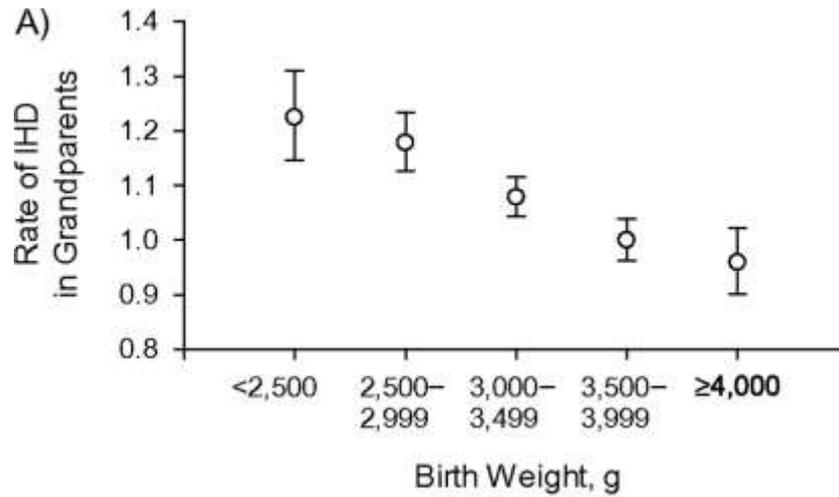
per 1,000 livebirths



# Intergenerational record linkage

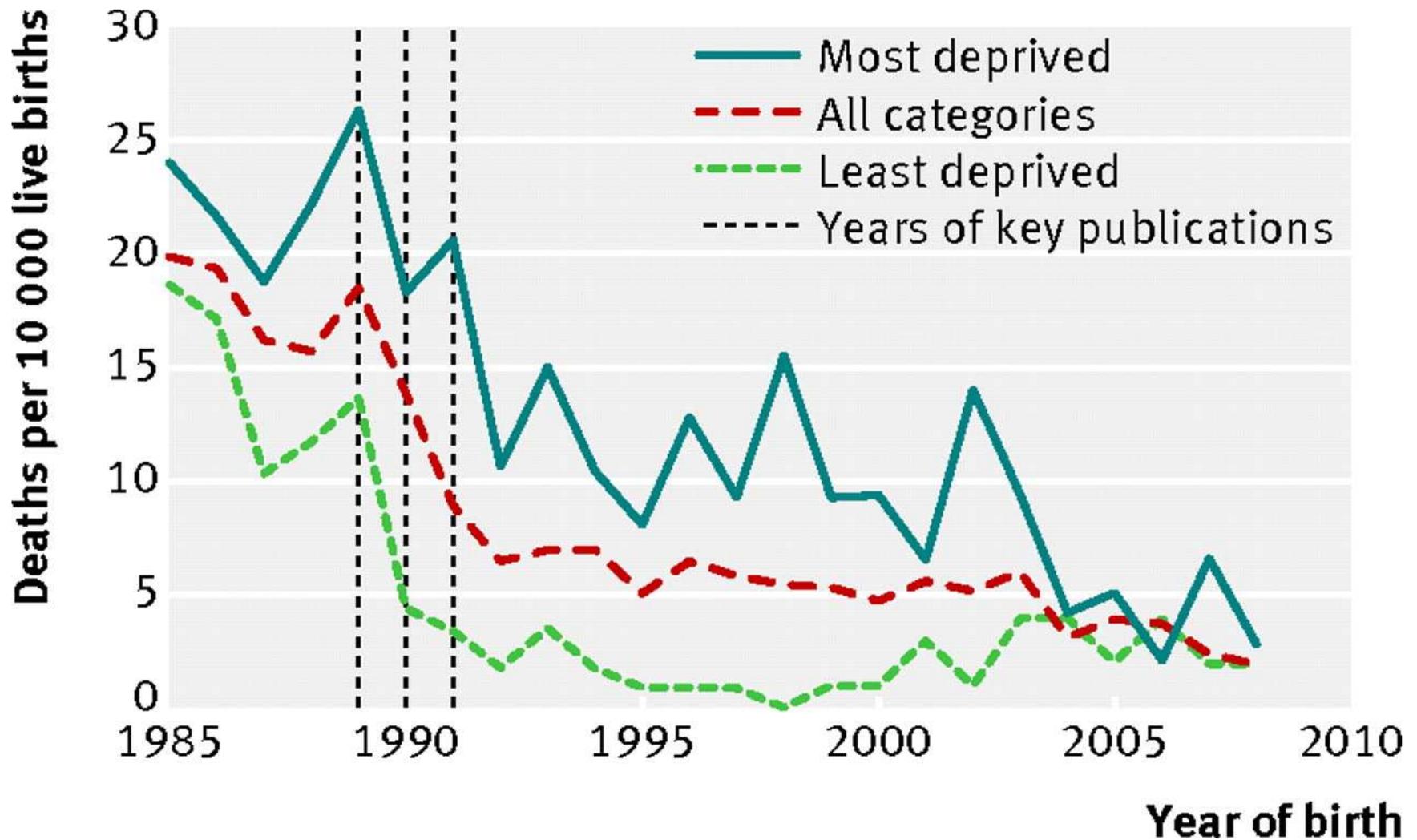


## Grandmaternal heart disease / stroke by infant weight

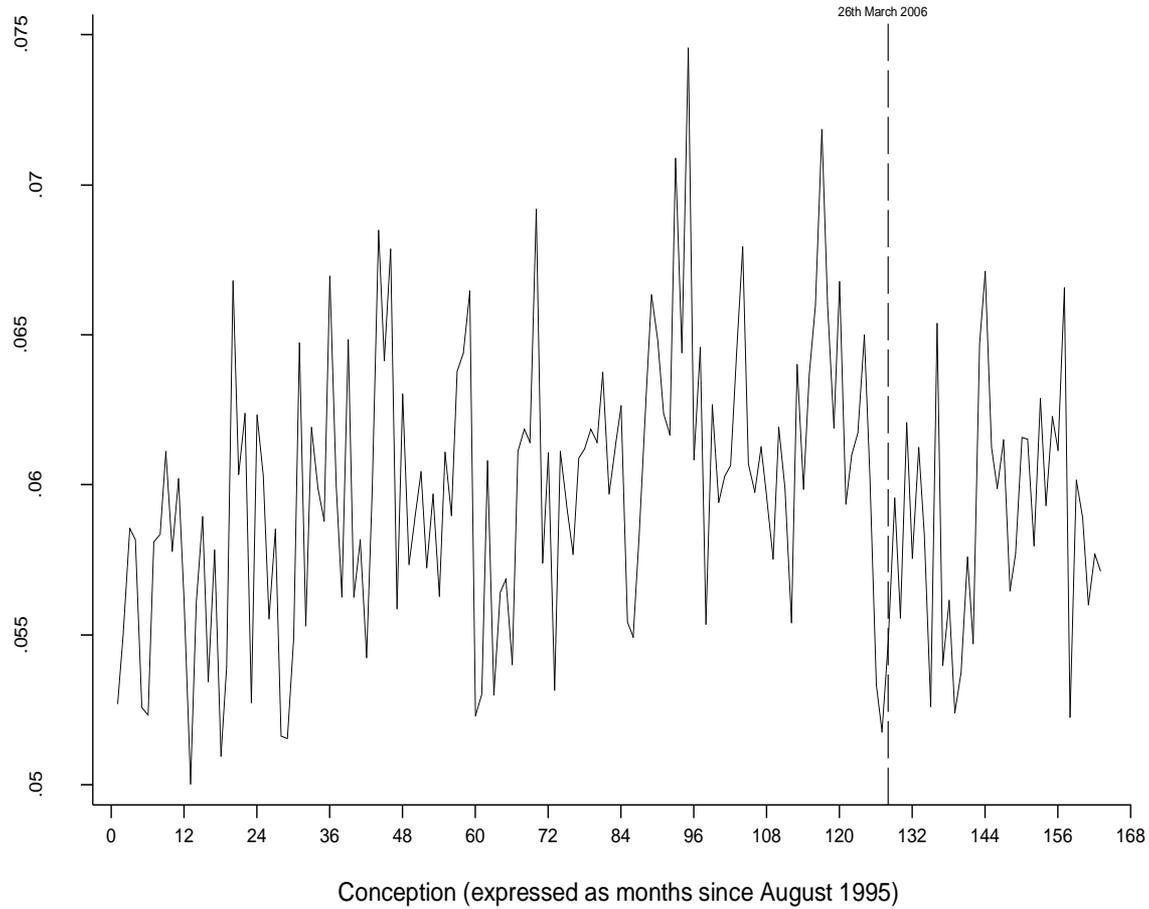


Interventional studies ...

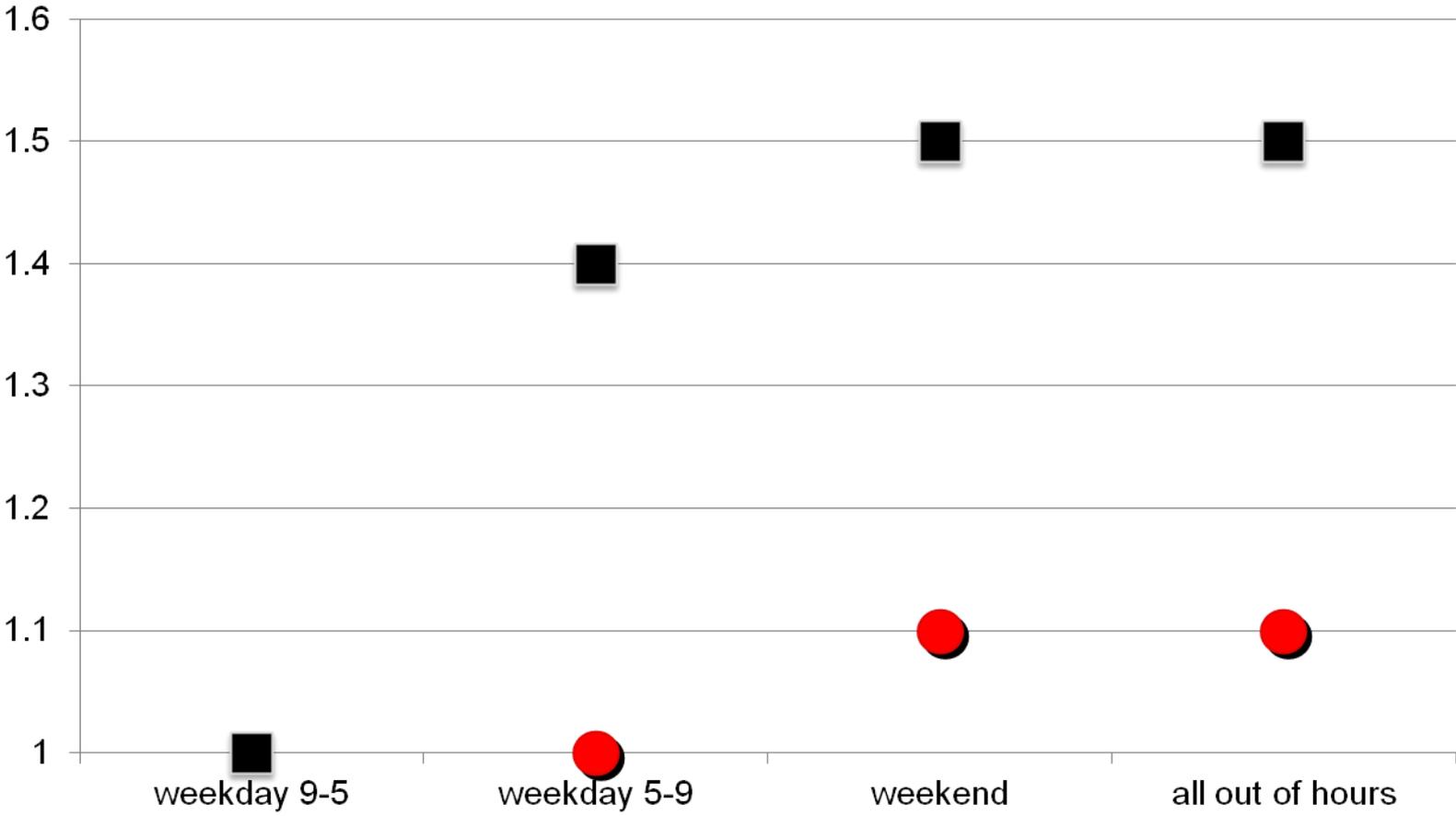
## “back to sleep” campaign for cot death



# Smoke-free legislation and preterm deliveries

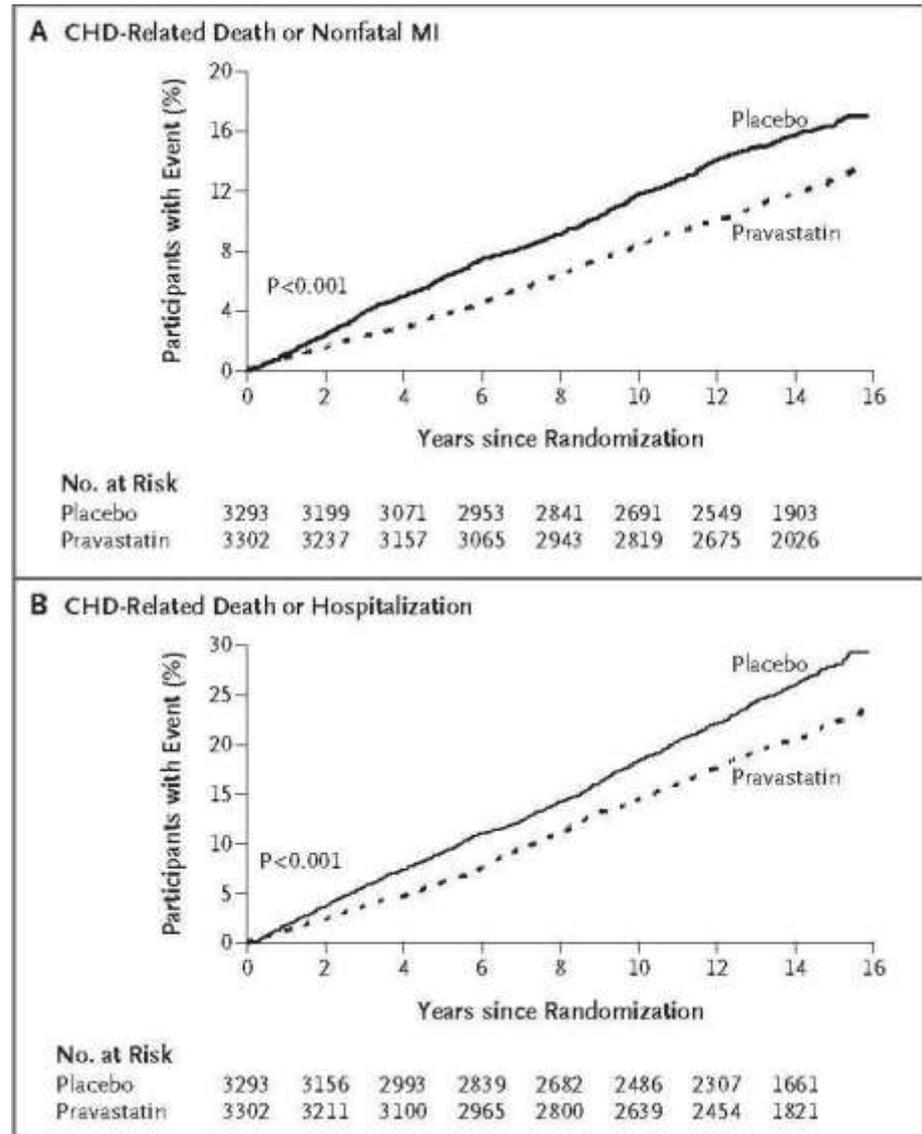


# Neonatal deaths by date and time of birth

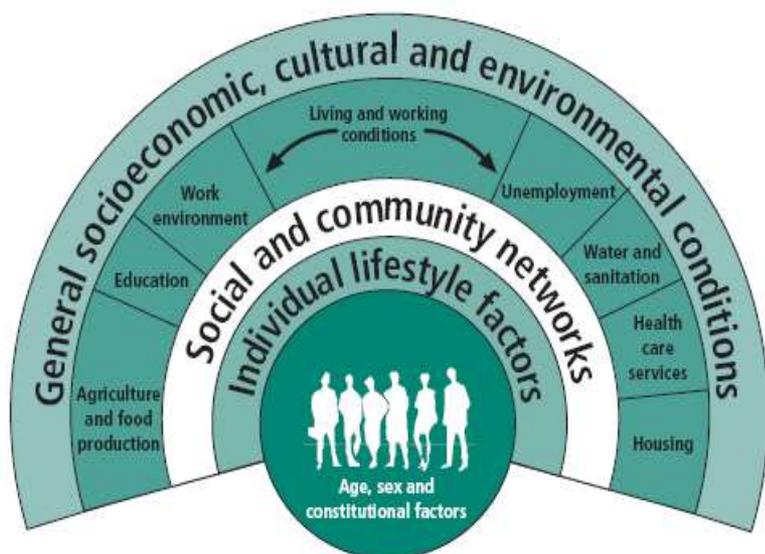


# Randomised Controlled Trials Follow-Up

		Record linkage	
		Yes	No
WOSCOPS	Yes	1043	66
	No	137	



# “Health-related” data from other sectors

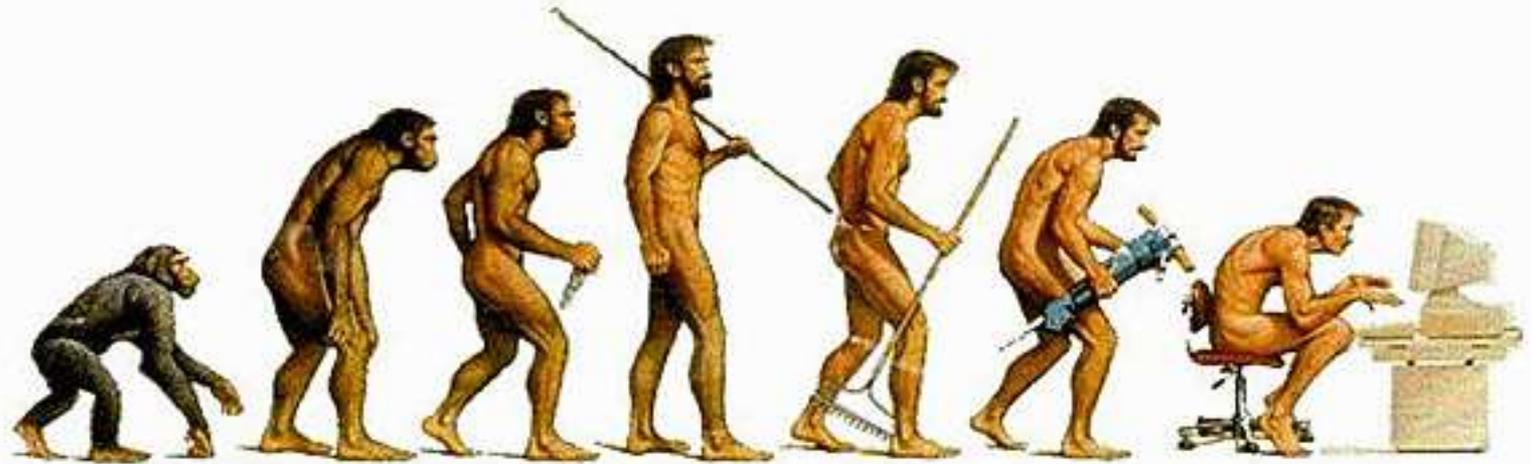


Source: Dahlgren G, Whitehead M. Tackling inequalities in health; what can we learn from what has been tried? Background paper for "The King's Fund International Seminar on Tackling Health Inequalities". Ditchley Park, Oxford: King's Fund; Reproduced with permission of the authors.

WHO 05.111

- Education
- Housing
- Environment
  - pollution
  - weather
- Social services
- Work, pensions & benefits
- Criminal justice

# The evolution of cross-sectoral record linkage



# Maternity-Education Linkage

Health	Maternity record All deliveries	Maternal factors (eg age, SIMD, smoking) Obstetric factors (eg mode/gestation of delivery, urgency) Infant factors (eg sex, birthweight, APGAR)
--------	------------------------------------	---

Education	Annual pupil census 2006 onwards Primary, secondary and special schools	Special educational need (learning/physical disabilities) Need type Truancy/absence Level of English Looked after status Free school meals
	Annual school leavers destination survey	Employed Unemployed Higher education
	Examination grades	All SCQF levels

## The neanderthal period ....

- 2006 requested pupil census data
- 3 years, and 20 meetings later .....
- approval to access centralised data denied
- approvals granted by 19/32 Local Authorities





Education produces two different extracts for ISD and investigators

Pupil ID	Identifiers <i>name, sex, dob, postcode</i>	Education data <i>SEN, type school etc</i>
----------	--	---



Pupil ID	Identifiers <i>name, sex, dob, postcode</i>
----------	--

ISD extract contains no education data only identifiers

Pupil ID	Education data <i>SEN, type school etc</i>
----------	---

Investigators' extract contains no identifier only education data



**ISD**

ISD links extract to health data and then removes the identifiers to produce an extract for the investigators

**INVESTIGATORS**

Investigators merge two extracts to produce one file containing health and education data, Investigators have no access to identifiers

Pupil ID	Health data <i>Pregnancy complications, vaccination status, antenatal markers etc</i>
----------	--

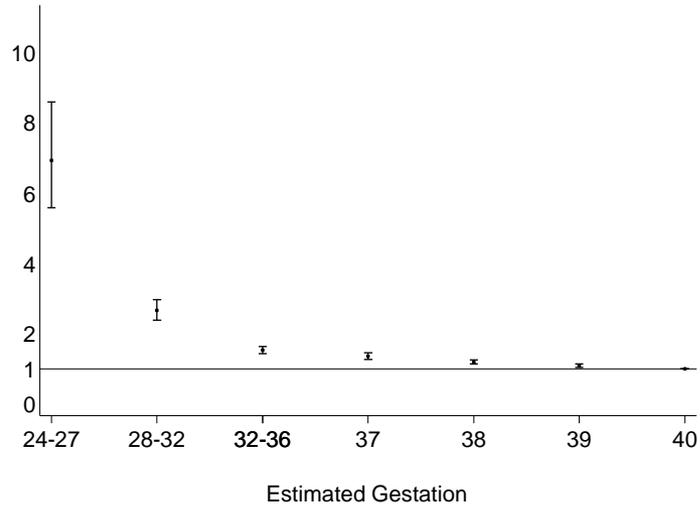


Pupil ID	Health data <i>Pregnancy, vaccination, antenatal markers etc</i>	Education data <i>SEN, type school etc</i>
----------	---	---

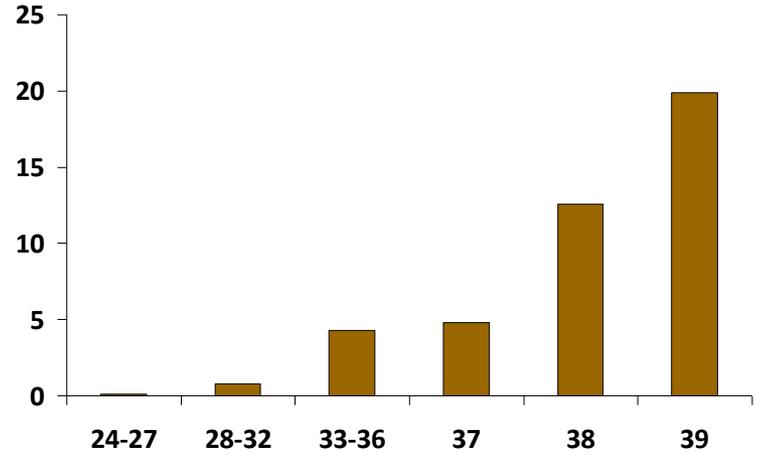
## But finally ....

- 19 Local Authorities
- 75% Scottish population
- N=407,503 pupils
- SEN in 17,784 (4.9%) children
  - 1,565 (8.4%) of preterm
  - 16,219 (4.7%) of term

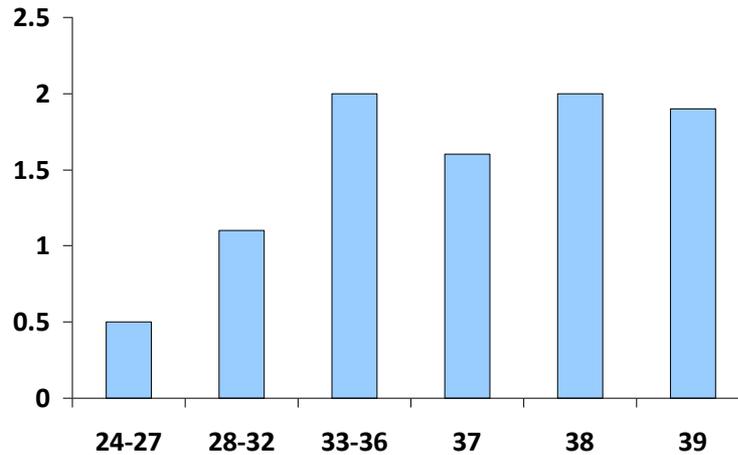
### Odds ratio



### Prevalence



### Population attributable percentage



Gestation overall 10%  
Preterm 3.6%  
Early term 5.5%

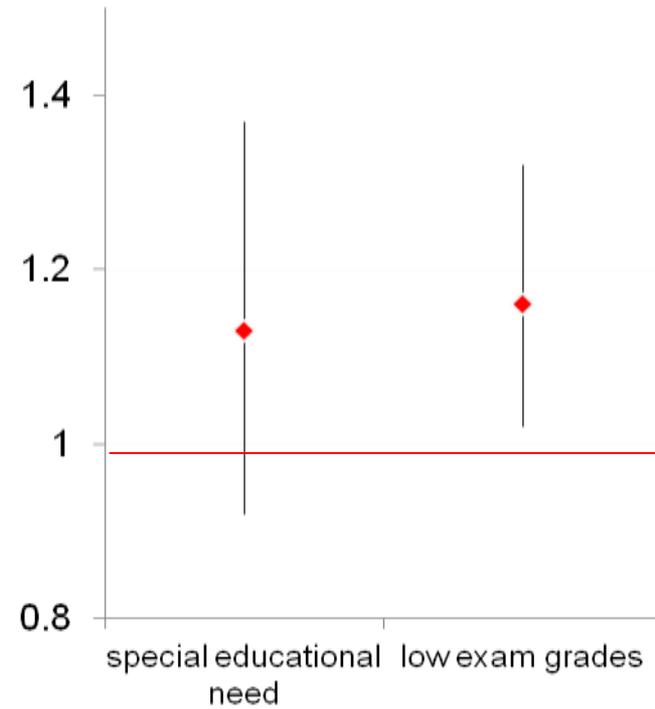
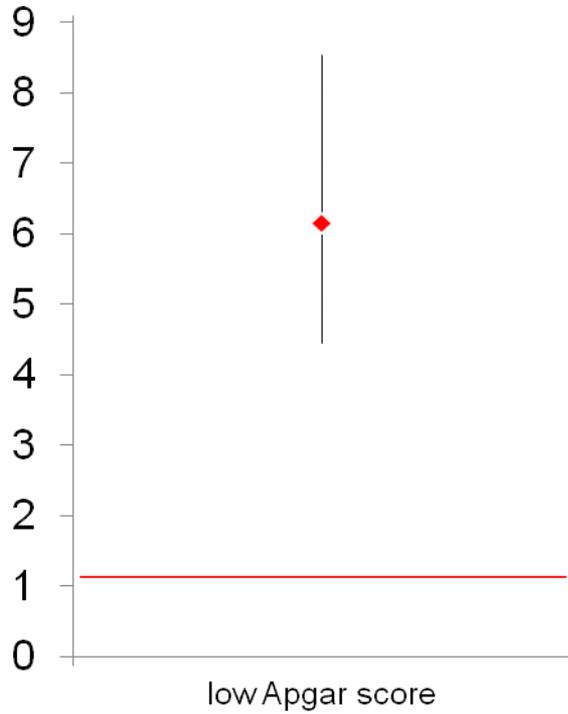
2012 .....

- Cross-sectoral “collaboration”
- Data sharing agreement
- Central data
  
- No permission to release names
- Probability matching on DOB, sex, postcode



- Pupils 2006-2011
- Linkage achieved for
  - 94% 1 census year
  - 97% >1 census year
- N=803,275 pupils
- 674,705 singletons
  - 2,130 (0.3%) breech vaginal
  - 13,054 (1.9%) breech caesarean section
  - 461,571 (68.4%) cephalic vaginal
- % of breech delivered vaginally fell:
  - 23% (2006 pupils) to 7% (2011 pupils)

# Delivery of breech infants Vaginal versus Caesarean section



2014 .....



**Data Sharing and Linkage Service**  
eDRIS

## Safe havens

