



BPSU Study guidance – Data analysis plan for BPSU studies

Dr Simon Lenton on behalf of the Scientific Committee of the British Paediatric Surveillance Unit

Updated 08012019

BPSU parent bodies:



Public Health
England



GREAT ORMOND STREET
INSTITUTE OF CHILD HEALTH

with support from:



Purpose

The purpose of this short paper is to help first time BPSU applicants to write a concise and successful data analysis plan, based on the objectives of the study. The paper is written in general terms, using a hypothetical example (metasyn disease).

This paper should be read in conjunction with epidemiology for BPSU studies and BPSU follow up measurement guidance papers.

Often the data analysis plan is left till last when writing a BPSU study application. This would probably not be the most efficient use of time, the better sequence might be:

1. convert your idea for the study into an answerable question e.g. “what is the incidence of xxx in the UK”
2. check the literature to ensure the incidence is not already known and that no more than 300 cases per year will be generated over the period of the study
3. define the supplementary questions e.g. “how is this condition managed”, “what are the outcomes at 1 year”
4. decide on the data required to answer these questions
5. data should only be requested where it is required to answer the research questions.
6. check this data is available in clinical notes or in the hospital laboratory system (may require a small survey of paediatricians)
7. design the questionnaires in a logical sequence so that the burden of data extraction is kept to a minimum
8. produce your data analysis plan.

The condition

Metasyn disease is a hypothetical newly emerging metabolic syndrome usually presenting in the first year of life associated with moderate learning difficulties, hearing loss and renal dysfunction. Incidence is 1:15k live births. It is known to have a small but significant mortality rate in those most severely affected. Case reports seem to show an association with pre-eclampsia and maternal smoking.

The study

Paediatricians are asked to report **NEW** cases meeting clearly defined criteria. The study seeks to establish the incidence, associations, management, service use and outcomes for this condition.

Data analysis plan

The following are the stages of analysis to address the research objectives.

How many notifications met the analytic case definition?

- Matching for duplicates.
- Removal of clinician reporting error.
- Unconfirmed cases (no questionnaires returned). Chase them up.
- Response rate – cases reported and confirmed (notification questionnaire returned).
- Response rate for follow-up questionnaire.
- Number of eligible cases based on analytic case definition (possibly use of review panel)
- **Cases after exclusion of first month of reports.**

Incidence and demographics

- *What is the incidence of the condition?*
- *What are national/regional incidence rates?*
- *What is the incidence by age group?*
- *What are rates by ethnicity or social group?*

Incidence	Numerator	Denominator
UK	Cases reported	ONS pop data
England	in England	ONS pop data England
Scotland	in Scotland	National Registrations Scotland (NRS). pop data
Wales	in Wales	ONS pop data Wales
NI	in NI	Northern Ireland Statistics and Registration Service (NISRA) pop data
ROI	in ROI	ONS pop data ROI
0-1 years	Aged 0-1 years	ONS pop data 0-1 years
1-5 years	Aged 1-5 years	ONS pop data 1-5 years
IMD quintile	In each quintile	Postcode derived IMD data
ethnicity	In each ethnic group	ONS ethnicity data

Antenatal associations

- *What antenatal factors are associated with a diagnosis of Metasyn disease?*

Associations	Numerator	Denominator
born preterm birth (<37 weeks gestation),	Cases <37 weeks gestation	Total cases
singleton/multiple birth,	Cases multiple birth	Total cases
pregnancy elampsia	Cases pregnancy elampsia	Total cases
prenatal lifestyle factors (e.g. smoking)	Case mothers smoking	Total cases
family history of Metasyn disease	Cases family history of Metasyn disease	Total cases

If there are significant numbers and the frequency of the factors is well documented in the general population comparison and statistical analysis may be possible.

Cause of Metasyn disease

- *What proportion of Metasyn disease cases have an established underlying cause?*

Causes	number (percentage of all cases)
Genetic	
Antenatal exposure	
Mixed	
Unknown	
Total	X (100%)

Comorbidities at 2 years

What conditions are associated with Metasyn disease at 2 years?

	Number (percentage of all cases aged 24 months)
Renal anomalies (define)	
Learning difficulties (define)	
Abnormalities of hearing (define)	

2. clinical management and use of health services

- *How many children with Metasyn disease received best practice investigations?*
- *How many children were referred to specialists?*

	Number (percentage of all cases)
Receiving investigations: Brain MRI Abnormal renal function (define) etc	Number receiving MRI Number CK>x
Referred for specialty opinions Neurology	Numbers referred
Nephrology by one year	Numbers referred by 1 year

Clinical outcomes at one year of age

Confirm which children have survived and have data from 24 months of age and are eligible for inclusion in analyses of outcomes at follow-up times.

- What is the mortality of MetaSym in the first year?
- What are the expected growth parameters at 1 year
- How many surviving children have normal/abnormal renal function at

one year

- What proportion have been referred to specialists

Mortality	Number died in year 1/total born in study period
Growth weight height BMI head circumference	Max, min, mean, mode Distribution curves
Renal function (define)	Report normal/abnormal/live at 1 year
Speech and communication	

Triangulation of results - case ascertainment and incidence comparison
How do the figures from this study compare to other sources?

	Case definition	Number	Population	Incidence
Hospital episode statistics				
Congenital abnormality register				
Laboratory results				
PH register				