Routine Data in Clinical Trials – the GBS3 Trial data flow model

Linda Fiaschi, Sarah Craig, Lucy Bradshaw, Reuben Ogollah, Rachel Haines, Kate Walker, Jane Daniels
Nottingham Clinical Trials Unit, University of Nottingham. United Kingdom

Background

The GBS3 Trial (ISRCTN49639731), coordinated by the Nottingham Clinical Trials Unit (NCTU), assesses whether universal testing of pregnant women for Group B Streptococcus colonisation reduces early-onset neonatal sepsis, compared to the current UK risk-factor based strategy. Approximately 320,000 pregnant women from 80 maternity units in England, Scotland and Wales are needed for this cluster randomised trial. The testing policy becomes the standard of care for the duration of the trial and written informed consent for participation is not sought, to avoid selection bias.

Method

Data provider security requirements create a two-step process to limit the release of personal identifiable information. The trial population is defined from maternity datasets through participating hospitals codes where women gave birth within the trial time window. Their babies are then identified by linkage and their identifiers used to request data from the microbiology and neonatal datasets.

All data is directly sent by each data provider to a Trusted Research Environment (TRE) where is safely stored and remotely processed, linked and analysed by designated analysts.

Impact

Our approach, enabling a large study at lower cost, which is becoming increasingly popular in the clinical trial community, has the potential to inform future trials on the strengths and limitations of perinatal RCHD use in clinical trials.

NCTU received approval to conduct the study from a Research Ethics Committee, the Health Research Authority, Confidential Advisory Group (CAG) and the Public Benefit and privacy Panel (PBPP)

Hospitals open to participation in trial

Women to give birth in these hospitals given information about trial, routine data collection and opt out options

Women will be offered test for GBS (or continue with usual care) according to local procedures

Medical notes completed by hospital team

National Databases (Routinely collected Data)

Key data from medical notes is intergrated into natonal databases, as per normal practice

Requests made to national databases providers following approval from the above authorities

National Databases Providers for GBS3 Trial

Maternity, hospital and mortality: NHS Digital, Clevermed (BadgerNET); Digital Health and Care (DHC) Wales; Public Health (PH) Scoltand

Lab test: NHS England, Health Protection (HP) Wales

Paediatric: Paediatric Intensive Care Audit Network (PICANet)

Neonatal: Imperial College (National Neonatal Research Database, NNRD); Clevermed (BadgerNet Neonatal)

National Databases providers verify appropriate approvals in place before any data is sent to a TRE

data safely sent to the TRE by each data provider

Trial Database (Manually collected data)

Some data entered into medical notes is not added to national databases. Therefore...

Local research midwife enters data into a electronic database held by the NCTU at the University of Nottingham

The local research midwife will also enter below identifiers into database:

- NHS/CHI Number
- Date of Birth
- Postcode

The information in the database at NCTU will be exported to the TRE to be linked to the routine datasets

TRE

Process 1 applies to: NHS Digital, DHC Wales and PH Scotland

National Database providers receive request from GBS3 team to provide all data from a hospital/trust during a specific time period

Data provided by these national databases to TRE directly and include below identifiers:

- NHS/CHI Number
- Date of Birth
- Postcode

Identifiers needed from process 1 to start process 2

Process 2 applies to: BadgerNet, NNRD, PICANet, NHS England and HP Wales

National database providers receive the below identifiers from GBS3 team to provide specific individual data.

- NHS/CHI Number
- Date of Birth
- Postcode

Data is sent by national databases providers directly to the TRE

Data from different datasets is combined into one anonymised dataset. Once the data is linked, identifiers are removed

Data is analysed. All reporting of data is done by group only (e.g. by testing type or by hospital)

University of Nottingham
UK | CHINA | MALAYSIA

SUPPORTED BY NIHR | National Institute for Health Research

This study is funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment Programme (HTA), (project reference 17/86/06). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.