Monitoring sample size assumptions in a cluster RCT using routine data and adaptations in real time: experience from the GBS3 trial

Lucy Bradshaw[1], Reuben Ogollah[1], Kate Walker[1, 2], Jane Daniels[1]

1 - Nottingham Clinical Trials Unit and 2 – Centre for Perinatal Research, University of Nottingham, UK

Background

- Sample size estimation for cluster randomised trials (cRCTs) requires specification of average cluster sizes as one of the design effect parameters
- When there are long waiting times between trial design and data acquisition, these parameters are likely to change and may affect the study power.
- We present a case study of the GBS3 cRCT (ISRCTN49639731) to outline the internal and external factors that impacted on assumptions relating to cluster size during the trial and how we tackled them.

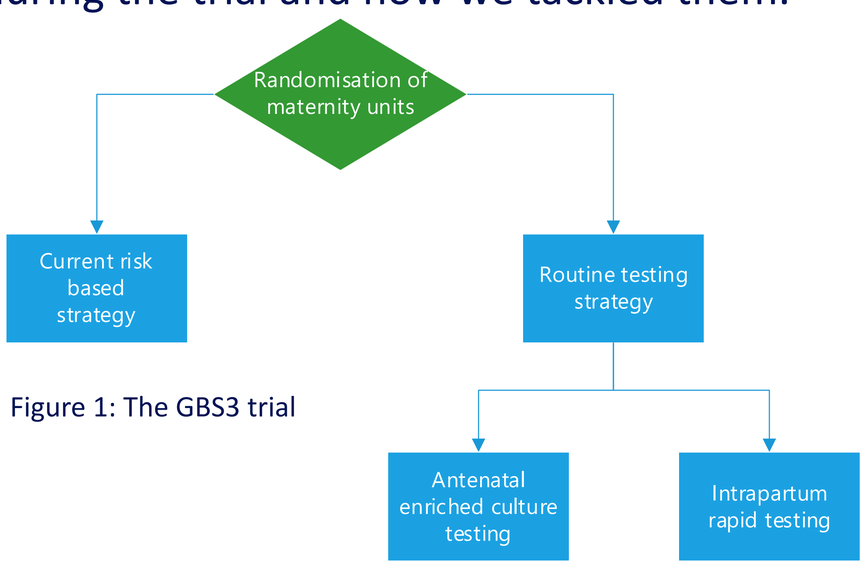

Figure 1: The GBS3 trial

The GBS3 Trial

A cRCT to determine whether routine testing of pregnant women for group B streptococcus reduces the incidence of early-onset neonatal sepsis compared to the current risk based strategy (Figure 1). Outcomes obtained from routine data.

Sample size and assumptions in grant application (2017)

- Designed to detect a 40% relative reduction in all cause early onset neonatal sepsis from 0.0986% to 0.0592% with 90% power
- Required data for 12 months from 72 sites with at least 3000 births a year (Table 1)

Table 1: Original assumptions used in sample size calculation

Sample size required without inflation for clustering	Assumed births per year (NHS maternity statistics[1])		Assumed intra cluster correlation coefficient	Sample size required inflated for clustering
	Mean	Coefficient of variation		
212960	4500	0.31	0.0001	320000

1 – for trusts with a minimum of 3000 deliveries per annum in 2016

Factors impacting on number of clusters and cluster size assumptions & adaptations to maintain the study power

Factor	Impact	Adaptations
Internal – sites needed to be able to commit to implement either testing strategy prior to randomisation	1. Smaller pool of sites than originally anticipated 2. Delayed randomisation of sites	Changed the site eligibility requirement for the minimum number of births per year from 3000 to 2000 to increase the potential pool
External • Reduction in birth rates (Figure 2) • Sharp increase in opt out for use of routine data for research purposes (Table 2)	Smaller cluster sizes than anticipated	Regular review of implications for the number of sites and length of the data collection period required based on up-to-date maternity statistics & opt-out rates
External - anticipated changes to the routine data sources after March 2024	Necessitated a fixed end date for data collection to avoid lengthy delay in obtaining data for analysis	Allowed variable data collection periods for each site from 9 to 16 months, rather than fixed at 12 months

Figure 2: Live births in England and Wales 2016 to 2022

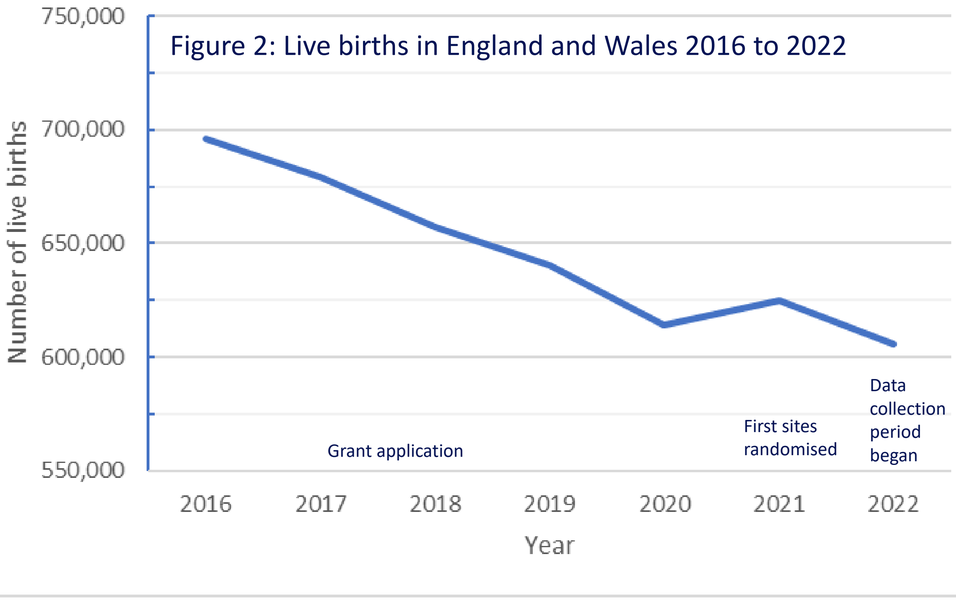

Table 2: National data opt out in England for women by age group (i.e. data cannot be used for research purposes)

Age	2019	2021
10 to 19	1.86%	3.20%
20 to 29	3.43%	7.28%
30 to 39	3.11%	7.89%
40 to 49	2.66%	7.28%

Discussion

It is important to plan how and when sample size assumptions in cRCTs will be monitored to allow changes to be made if needed. This is especially important when there are long waiting times between trial design and actual data acquisition.

University of Nottingham
UK | CHINA | MALAYSIA

FUNDED BY NIHR | National Institute for Health and Care Research

GBS3 ROUTINE TESTING FOR GROUP B STREP

This study is funded by the 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.