Assessing quality of NHS Numbers for Babies data and providing gestational age statistics

Kath Moser
Office for National Statistics
and Lisa Hilder
Department of Midwifery, City University

This paper investigates the quality of information in the National Health Service (NHS) Numbers for Babies (NN4B) dataset and provides statistics on gestational age at birth in relation to live/stillbirth status, multiple birth status, age of mother, babies' sex, and birthweight. Gestational age information is not recorded at the registration of live births and the NN4B system provides the opportunity to access this information for all live and stillbirths in England and Wales. All NN4B records for babies born in England and Wales in 2005 were used in this analysis. Data quality was generally good although some aspects need further investigation. The gestational age data are credible, consistent with other UK data sources, and the statistics by maternal age, multiplicity and sex are as expected. These data, previously unavailable for England and Wales as a whole, provide a reference against which to monitor trends in preterm births and can inform the planning of neonatal care provision. NN4B is a powerful new data source which can extend information on births in England and Wales.

Introduction

Low gestational age at birth is the principal factor associated with mortality in the perinatal period.[1] For example, in Scotland preterm births accounted for 6.2 per cent of singleton live births in 2004 but 66 per cent of stillbirths and 64 per cent of neonatal deaths.[2] Babies born very preterm are at particular risk of subsequent cognitive, sensory and motor impairments.[3,4] Babies who are small, or large, for their gestational age are more likely than others to have adverse outcomes.[5,6] Substantial effort is made antenatally to identify abnormal fetal size for gestational age and poor fetal growth may prompt a decision to induce labour. Induction of labour is recommended routinely in England for pregnancies that continue beyond 41 completed weeks[7] and recent calls have been made to reduce the gestation at which induction is offered routinely in certain groups.[8,9] Gestational age at birth is both an important obstetric outcome and a risk factor.

Given its importance it is surprising that, until now, comprehensive gestational age data have not been readily available for England.[10,11] Gestational age is not recorded at registration of live births in England and Wales and so is not present in registration data, the routine source of information on births. Although gestational age is included in birth notification records, it had not been feasible to aggregate these data routinely for all births in England, although this had been possible for Wales.[12] Estimates of number of deliveries by gestational age are available from the Hospital Episode Statistics for England[13] but, because some hospitals fail to contribute full maternity data, they are incomplete. In Scotland gestational age data are available from maternity hospital records[14] and in Northern Ireland from child health systems. Wales, Northern Ireland and Scotland each routinely collate gestational age data with information on mortality.[2,15] This left England as the only country in the UK where, until now, it had not been possible to derive levels and

trends in preterm birth or gestation specific stillbirth, perinatal and infant mortality rates. To date, therefore, England has not been included in international comparisons of gestational age statistics such as the Peristat indicators for describing perinatal health across Europe.[16,17]

In 2002 a system for allocating NHS numbers at birth to babies in England, Wales and the Isle of Man was introduced.[18] Including the NHS number in medical records from birth, rather than from birth registration which can take place up to six weeks after birth, is essential for monitoring the care given to new babies, especially those admitted for neonatal care. As a by-product of this change, a subset of birth notification details, including gestational age, is now collected centrally on a single computer system. Thus, for the first time, it is feasible to collate this information for England and Wales as a whole. Moreover, because NHS numbers are now allocated at birth, it is possible to link these notification details with other datasets containing NHS numbers.

The Office for National Statistics (ONS) has received daily downloads of a subset of the variables contained in the data received by the NHS Numbers for Babies (NN4B) Service, for all NHS numbers issued from 1 January 2005. Using these data, statistics on the distribution of live births in England and Wales by multiplicity and gestational age were released for the first time in May 2007.[19] Subsequently, using NN4B data linked with birth registration records and registration records for deaths in the first year of life, statistics on gestation-specific infant mortality in England and Wales were published for the first time in August 2007.[20]

In this paper we describe the NN4B dataset, assess its quality and use it to derive statistics on gestational age at birth in relation to live/stillbirth status, multiple birth status, age of mother, babies' sex, and birthweight.

Methods

The process of allocating NHS numbers to babies is initiated very shortly after delivery when the birth is notified to the NN4B Service via a computer link. An NHS number is promptly returned and incorporated in the baby's record. Of the data submitted to the NN4B Service, the subset of variables received by ONS is shown in Table 1. Approval for the use of NN4B data for linkage and the production of statistical data was given by the North East London Ethics Committee and the Patient Information Advisory Group.

NN4B records relating to all babies born from 1 January to 31 December 2005 inclusive were used in this analysis. A check for duplicate NHS numbers found seven pairs of records identical on all variables including NHS number. One of each pair was deleted leaving 649,740 NN4B records for births occurring in England, Wales and the Isle of Man in 2005. As the Isle of Man is not part of England and Wales, and has separate arrangements for birth registration, the 566 births occurring in hospital in the Isle of Man were excluded from further analysis, leaving a total of 649,174 live and stillbirths.

Assessing NN4B data quality and validity

1. Data quality was assessed by considering missing values and outliers on six key variables: live/stillbirth status, multiple birth status, sex, mothers' date of birth, birthweight and gestational age. These results are presented in Section A and Table 2.

2. In addition, gestational age and birthweight data were examined in combination to further assess the quality of recording. This examination was restricted to live singletons and is covered in Section B (Table 5) alongside the presentation of the detailed gestational age statistics.

3. It is important when considering statistics from a new data source to assess their validity by comparison with data from other sources, and this is covered in Section C. The primary source against which we compared NN4B data was birth registration data. These form the routine source of birth data in England and Wales and are considered complete, as registration is a legal requirement and the data undergo extensive quality checks.[21,22] As registration data include sex, live/stillbirth status, birthweight, gestational age of stillbirths, multiple birth status and mother's age, we were able to compare the distributions of these variables in the two datasets (Table 6). Since birth registration data do not include gestational age of live births we could not validate NN4B gestational age data for **all** births against this source. However, we could make comparisons with data from three other UK datasets, one for each of England, Wales, and Scotland, as described below. These comparisons include all live and stillbirth records with a gestational age recorded as 20 to 45 weeks inclusive (Figures 1 and 2). In an additional check on data quality we compared NN4B birthweight for gestational age data for live singletons with the Scottish data, again including only those with gestational age 20 to 45 weeks (Figure 3).

Table 1 NHS Numbers for Babies variables available to the research project

Variables	Status of variable (M=mandatory C=conditional, O=optional)	Used in this analysis?
Baby and birth details		
NHS number	M	–
Date of birth	M	Y
Delivery time	M	–
Sex	M	Y
Live or still birth	M	Y
Birthweight	M	Y
Gestation	M	Y
Number of births in this confinement	M	Y
Birth order in this confinement (if multiple)	C	–
Suspected congenital anomaly	M	–
Ethnic category of baby	M	–
Postcode of usual address of baby	O	–
Place of birth details		
Organisation name	C	–
Organisation code	C	–
Delivery place type	M	–
Mother details		
Mothers' NHS number	C	–
Mothers' date of birth	C	Y
Health care professional details		
National GP code	C	–
Child health organisation code	M	–

Table 2 Completeness and quality of select NN4B variables, 2005

Variable	Records with missing values percentage (number)	Records with outlying or implausible values percentage (number)
Live/stillbirth status	0% (0)	0.04% (231) recorded as 'dead' [1]
Multiple birth status	0% (0)	0.002% (11) >=5 births in the confinement
Sex	0% (0)	0% (0)
Mothers' date of birth	0.07% (471)	0.01% (83)
Birthweight	0.4% (2,782) [2]	0.01% (52) birthweight >=6,000g
Gestational age	0.7% (4,319) [3]	0.02% (121) <20 weeks, 0.03% (216) 46+ weeks

1 The option 'dead' on NN4B records allows the clinician to indicate that a baby born alive had died before the notification was made. These births were counted as live born in the analysis presented in this paper.
2 Live singletons 0.4% (2,536), live multiples 1.0% (195), stillbirths 1.5% (51).
3 Live singletons 0.7% (4,186), live multiples 0.7% (133), stillbirths 0% (0).

Datasets used for comparisons:

1. *Birth registration data for England and Wales, 2005*.[21,22]
2. *Maternity Hospital Episode Statistics (HES) data for England, 2004/05*.[13] Since some maternity units fail to contribute data, information is missing for 25 per cent of deliveries in hospitals and most home births. The HES estimates of the number of deliveries at each gestational age are obtained by grossing up the percentage distribution obtained from records with data to match the known number of deliveries in the year.
3. *The National Community Child Health Database for Wales (NCCHD)*.[12] This uses data derived from Community Child Health System databases (CCH2000) held by NHS Trusts and merged to create an all-Wales database. Since 2002 CCH2000 records have been initiated from NN4B records but may subsequently be amended or added to by Trusts. As there were only 32,768 births in Wales in 2005,[22] data for the five years 2001–05 have been used.[23]
4. *Scottish Morbidity Record (SMR02), maternity hospital data*.[14] Data are available for 98 per cent of all births in Scotland. Most of the shortfall is due to incomplete data submissions from maternity units, the remainder being due to home births and births in non-NHS hospitals. As there were only 54,678 births in Scotland in 2005,[22] data for the five years 2001–05 have been used.[24]

In comparing NN4B data with these datasets the following points should be considered. The birth registration dataset refers to all births occurring in England and Wales in 2005, exactly as does the NN4B dataset. However, the HES data refer to England, and the NCCHD data refer to Wales. Unfortunately, as we found postcode to be poorly recorded in the NN4B data, it was not easy to separate out the NN4B records for England from those for Wales; consequently all comparisons used NN4B data for England and Wales combined. In contrast to the comparison datasets for England and Wales, the Scottish data are an external source that refer to a different country and, therefore, comparisons with NN4B data may reflect real differences in health and clinical policy and practice between Scotland and England and Wales.

Gestational age statistics

We derived: i) the gestational age distribution, ii) the mean gestational age, and iii) the percentage preterm, separately for live and stillbirths subdivided by multiple birth status in each case (Table 3). In line with WHO definitions,[25] preterm birth was defined as before 37 completed weeks of gestation, term as 37–41 completed weeks, and post term as 42 or more completed weeks (Box One). Births with unstated gestational age were omitted from the calculations of means, and from the denominator used in calculating the percentage preterm. More detailed analyses of singleton live births tabulated gestational age by mother's age at birth, by babies' sex (Table 4) and by birthweight (Table 5). Only births of stated birthweight were included in the calculations of the mean birthweight and the percentage of babies of low birthweight (those weighing under 2,500g).

These statistics are presented in Section B.

Box one

Definitions used

Gestational age

Preterm	Less than 37 completed weeks
Term	37 to 41 completed weeks
Post term	42 or more completed weeks

Birthweight

Low birthweight	Less than 2,500g

Results

A. Internal quality of NN4B data

The completeness and quality of NN4B data on live/stillbirth status, multiple birth status, sex, mothers' date of birth, birthweight and gestational age is shown in Table 2. Data completeness is generally good for these variables, although 0.4 per cent of births have unstated birthweight and 0.7 per cent unstated gestational age. On some of the variables a small percentage of records have outlying or implausible values. The gestational age and birthweight data are examined further towards the end of the next section.

B. Gestational age statistics from NN4B data, and examining consistency of birthweight and gestational age information

The percentage of all births in England and Wales in 2005 born preterm was 7.9 per cent (Table 3). Although most preterm births were live singletons, the proportion of live multiples and proportion of stillbirths born preterm were much higher (53.3 per cent and 63.8 per cent respectively) than the proportion of live singletons born preterm (6.2 per cent). These differences are reflected in the mean gestational age, which was 39.25 weeks for live singletons but four weeks less for live multiples and six weeks less for stillbirths. Multiple births accounted for 2.9 per cent of all births but 20 per cent of preterm births. More than one in 25 births were born post term.

Singleton live births

The percentage of singleton live births born preterm showed a U-shaped association with the age of mother (Table 4). The highest percentage born preterm occurred in babies of mothers aged under 20 and those aged 40 and over, groups which between them account for 10 per cent of all live singletons. The percentage of live singletons born at under 28 weeks was also highest in these maternal age groups. The lowest percentage born preterm was in babies of mothers aged 30–34. The percentage born post term decreased with mothers' age, from 4.8 per cent in the youngest to 3.8 per cent in the oldest age group. The mean gestational age at birth increased slightly from 39.23 weeks for babies of mothers aged under 20 to a maximum of 39.29 for mothers aged 30–34; it then decreased with mothers' age to 39.02 weeks at ages 40 and over.

A higher percentage of boys were born preterm (6.7 per cent) than girls (5.8 per cent); the difference between the sexes was largely confined to gestational ages 32–36 weeks. The percentage born post term was slightly higher for boys than girls.

Table 3 **Gestational age at birth by live/stillbirth and multiple birth status, all births: NN4B data, 2005**

England and Wales

	Live births			Stillbirths			All births	Number of all births
	Singleton	Multiple	All live births	Singleton	Multiple	All stillbirths		
Gestational age distribution (percentage)								
<22 wks	0.05	0.27	0.05	:	:	:	0.05	330
22 wks	0.02	0.18	0.02	:	:	:	0.02	159
23 wks	0.03	0.44	0.05	:	:	:	0.05	294
24 wks	0.06	0.69	0.08	8.01	11.15	8.26	0.12	785
25 wks	0.06	0.64	0.08	6.29	7.19	6.36	0.11	735
26 wks	0.10	0.67	0.11	5.73	5.76	5.73	0.14	921
27 wks	0.10	0.85	0.12	4.82	6.47	4.95	0.14	940
28 wks	0.14	1.28	0.17	4.82	3.24	4.69	0.19	1,251
29 wks	0.15	1.64	0.19	3.85	3.60	3.83	0.21	1,361
30 wks	0.20	1.93	0.25	3.88	4.68	3.94	0.27	1,755
31 wks	0.23	2.69	0.30	3.35	5.04	3.48	0.32	2,069
32 wks	0.32	4.02	0.43	3.79	7.19	4.06	0.45	2,898
33 wks	0.47	5.09	0.61	3.75	7.91	4.09	0.62	4,051
34 wks	0.77	7.75	0.97	3.85	8.99	4.26	0.99	6,437
35 wks	1.19	10.06	1.45	4.54	7.19	4.75	1.46	9,509
36 wks	2.32	14.69	2.68	5.26	6.47	5.35	2.70	17,519
37 wks	5.26	20.39	5.70	6.13	5.40	6.07	5.70	36,990
38 wks	13.40	19.11	13.57	7.45	4.68	7.23	13.53	87,846
39 wks	22.12	4.78	21.61	6.85	3.24	6.56	21.53	139,776
40 wks	27.88	1.54	27.11	9.17	1.80	8.58	27.01	175,364
41 wks	20.06	0.38	19.49	6.91	0.00	6.36	19.42	126,045
42 wks	4.00	0.18	3.89	1.28	0.00	1.18	3.88	25,179
43 wks	0.26	0.03	0.26	0.13	0.00	0.12	0.26	1,663
44 wks	0.09	0.01	0.09	0.09	0.00	0.09	0.09	552
45 wks	0.03	0.00	0.03	0.06	0.00	0.06	0.03	210
46+ wks	0.03	0.00	0.03	0.00	0.00	0.00	0.03	216
not stated	0.67	0.71	0.67	0.00	0.00	0.00	0.67	4,319
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	
Number	626,924	18,776	645,700	3,196	278	3,474	649,174	649,174
Percentage preterm (<37 wks)	*6.2*	*53.3*	*7.6*	*61.9*	*84.9*	*63.8*	*7.9*	
Mean gestational age, weeks	39.25	35.36	39.14	33.13	31.21	32.97	39.10	

The distribution of live singletons by gestational age and birthweight is shown in Table 5, along with the birthweight distribution, percentage low birthweight and mean birthweight for gestational age groups. Over 90 per cent of live singletons are neither preterm nor of low birthweight, while 3.5 per cent are both preterm and of low birthweight. Over 100 live singletons were recorded as being born at under 20 weeks gestational age of which half (52) were recorded as having a birthweight of 2,500g or above. This combination is implausible and suggests errors in the recording of birthweight, gestational age or both. Similarly, nine live singletons were recorded as of 20–21 weeks gestational age and birthweight of at least 2,500g. Furthermore, 944 term/post term births were recorded with a birthweight of under 500g (that is, 1–499g), again suggesting recording errors in birthweight, gestational age or both. Although these 944 births account for under 0.2 per cent of term/post term live singletons, the numbers involved are not insubstantial, and give rise to only 34.2 per cent recorded as preterm amongst live singletons weighing 1–499g at birth. In comparison, 97.6 per cent of those weighing 500–1,499g at birth are born preterm. Almost 5 per cent of babies of under 32 weeks gestational age have unstated birthweight compared to less than 0.5 per cent of term and post term births.

C. Assessing quality of NN4B gestational age data through comparison with other data sources

NN4B data and birth registration data for England and Wales are compared in Table 6. There is almost total agreement between the datasets regarding the percentage of births that are born live, the percentage that are multiples, the sex ratio and the maternal age distribution. The birthweight distributions of live births agree closely between the datasets except that the proportion of births weighing 1–499g is higher in the NN4B than the registration data. The birthweight distributions of stillbirths broadly agree although less closely than for live births. The gestational age distributions of stillbirths are similar; however the percentage of stillbirths that are preterm is lower according to NN4B than registration information, while the percentage term/post term is higher. There are 205 fewer NN4B than registration records for births in 2005.

Figure 1 compares the gestational age distribution of all births of gestational age 32–45 weeks in the NN4B data with distributions in each of Scotland, Wales and England. Equivalent comparisons are shown on a larger scale in Figure 2 for births of 20–31 weeks gestation. The

Table 4 **Gestational age at birth by mothers' age and by babies' sex, live singletons: NN4B data, 2005**

England and Wales

	Mothers' age						Babies' sex[1]		All live singletons
	under 20	20-24	25-29	30-34	35-39	40 and over	Male	Female	
Gestational age distribution (percentage)									
<22 wks	0.1	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0
22-23 wks	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1
24-27 wks	0.5	0.3	0.3	0.3	0.3	0.4	0.3	0.3	0.3
28-31 wks	1.0	0.8	0.7	0.6	0.7	0.8	0.8	0.7	0.7
32-34 wks	2.0	1.5	1.6	1.4	1.6	1.8	1.7	1.4	1.6
35-36 wks	4.0	3.7	3.5	3.3	3.5	3.9	3.8	3.3	3.5
37-38 wks	16.4	17.7	18.0	18.7	20.7	23.8	19.1	18.2	18.7
39-41 wks	70.2	70.5	70.8	70.6	68.5	64.9	69.1	71.0	70.1
42+ wks	4.8	4.6	4.4	4.5	4.1	3.8	4.5	4.3	4.4
not stated	1.1	0.9	0.7	0.5	0.5	0.4	0.7	0.7	0.7
Number of live singletons	44,192	119,893	159,585	181,628	99,927	21,215	321,030	305,348	626,924
Percentage preterm (<37 wks)	*7.6*	*6.5*	*6.2*	*5.8*	*6.2*	*7.1*	*6.7*	*5.8*	*6.2*
Mean gestational age, weeks	39.23	39.27	39.28	39.29	39.18	39.02	39.21	39.29	39.25

1 There were 546 live singleton births recorded as unknown sex.

Table 5 **Gestational age by birthweight, live singletons: NN4B data, 2005**
(shaded areas: dark grey = preterm & low birthweight; light grey = preterm or low birthweight but not both)

England and Wales

Gestational age (numbers)	Birthweight (grams)											Total
	1–	500–	1,000–	1,500–	2,000–	2,500–	3,000–	3,500–	4,000–	4,500+	not stated	
<20 wks	47	0	0	0	1	10	22	13	6	1	9	109
20-21 wks	139	11	2	0	1	2	2	3	2	0	8	170
22-23 wks	113	200	4	0	0	2	1	3	0	0	15	338
24-27 wks	64	1,319	394	14	7	17	24	13	9	5	103	1,969
28-31 wks	30	490	2,064	1,310	143	81	76	49	13	5	205	4,466
32-34 wks	46	33	724	3,111	3,934	1,282	303	94	17	5	241	9,790
35-36 wks	52	6	103	1,467	6,328	9,324	3,578	850	168	36	118	22,030
37-38 wks	183	9	42	777	8,783	36,158	45,951	19,617	4,335	742	359	116,956
39-41 wks	706	38	30	190	5,095	54,101	168,123	150,734	50,054	8,784	1,325	439,180
42-45 wks	54	4	9	12	173	1,682	7,934	11,073	5,202	1,252	119	27,514
46+ wks	1	0	0	2	6	34	77	65	29	1	1	216
not stated	5	15	31	52	208	737	1,483	1,144	399	79	33	4,186
Total	1,440	2,125	3,403	6,935	24,679	103,430	227,574	183,658	60,234	10,910	2,536	626,924

Gestational age	Birthweight distribution (per cent)[1]											Within gestational age group			
	1–	500–	1,000–	1,500–	2,000–	2,500–	3,000–	3,500–	4,000–	4,500+	Total[1]	Low birthweight (per cent)	Mean birthweight (grams)	Lower 95% confidence limit	Upper 95% confidence limit
<20 wks	*47.0*	*0.0*	*0.0*	*0.0*	*1.0*	*10.0*	*22.0*	*13.0*	*6.0*	*1.0*	*100.0*	*48.0*	1,925	1,603	2,248
20–21 wks	*85.8*	*6.8*	*1.2*	*0.0*	*0.6*	*1.2*	*1.2*	*1.9*	*1.2*	*0.0*	*100.0*	*94.4*	570	457	684
22–23 wks	*35.0*	*61.9*	*1.2*	*0.0*	*0.0*	*0.6*	*0.3*	*0.9*	*0.0*	*0.0*	*100.0*	*98.1*	587	544	631
24–27 wks	*3.4*	*70.7*	*21.1*	*0.8*	*0.4*	*0.9*	*1.3*	*0.7*	*0.5*	*0.3*	*100.0*	*96.4*	952	923	980
28–31 wks	*0.7*	*11.5*	*48.4*	*30.7*	*3.4*	*1.9*	*1.8*	*1.1*	*0.3*	*0.1*	*100.0*	*94.7*	1,472	1,455	1,489
32–34 wks	*0.5*	*0.3*	*7.6*	*32.6*	*41.2*	*13.4*	*3.2*	*1.0*	*0.2*	*0.1*	*100.0*	*82.2*	2,105	2,095	2,115
35–36 wks	*0.2*	*0.0*	*0.5*	*6.7*	*28.9*	*42.6*	*16.3*	*3.9*	*0.8*	*0.2*	*100.0*	*36.3*	2,654	2,647	2,661
37–38 wks	*0.2*	*0.0*	*0.0*	*0.7*	*7.5*	*31.0*	*39.4*	*16.8*	*3.7*	*0.6*	*100.0*	*8.4*	3,129	3,126	3,131
39–41 wks	*0.2*	*0.0*	*0.0*	*0.0*	*1.2*	*12.4*	*38.4*	*34.4*	*11.4*	*2.0*	*100.0*	*1.4*	3,482	3,480	3,483
42–45 wks	*0.2*	*0.0*	*0.0*	*0.0*	*0.6*	*6.1*	*29.0*	*40.4*	*19.0*	*4.6*	*100.0*	*0.9*	3,661	3,655	3,666
46+ wks	*0.5*	*0.0*	*0.0*	*0.9*	*2.8*	*15.8*	*35.8*	*30.2*	*13.5*	*0.5*	*100.0*	*4.2*	3,415	3,340	3,489
not stated	*0.1*	*0.4*	*0.7*	*1.3*	*5.0*	*17.7*	*35.7*	*27.5*	*9.6*	*1.9*	*100.0*	*7.5*	3,323	3,304	3,342
Total	*0.2*	*0.3*	*0.5*	*1.1*	*4.0*	*16.6*	*36.4*	*29.4*	*9.6*	*1.7*	*100.0*	*6.2*	3,349	3,347	3,350
Within birthweight group:															
Percentage preterm	*34.2*	*97.6*	*97.6*	*85.8*	*42.6*	*10.4*	*1.8*	*0.6*	*0.4*	*0.5*					

1 Distribution of births with stated birthweights only.

Figure 1 Gestational age distribution, all births, 32–45 weeks gestation: comparing NN4B data with data from England, Wales & Scotland

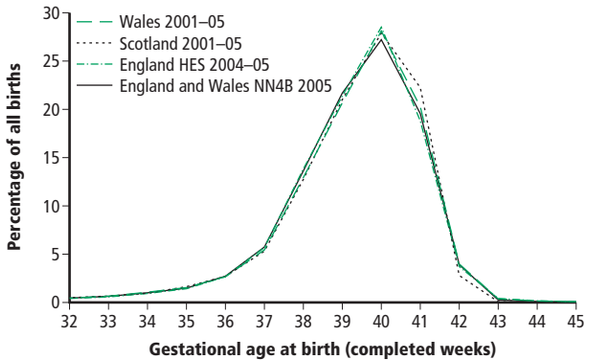


Figure 2 Gestational age distribution, all births, 20-31 weeks gestation: comparing NN4B data with data from England, Wales & Scotland

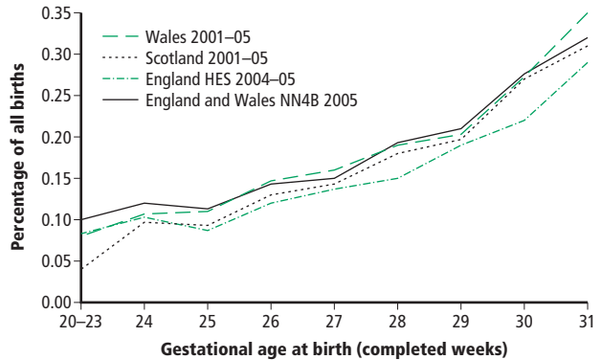
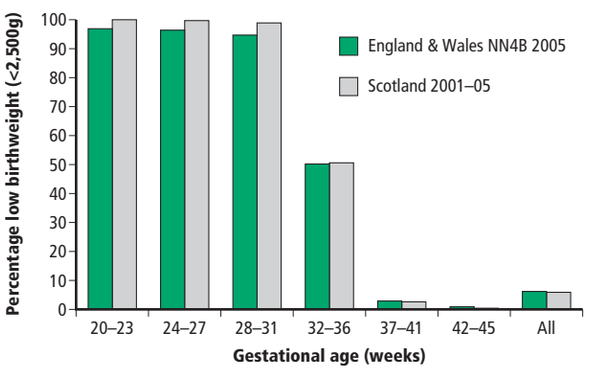
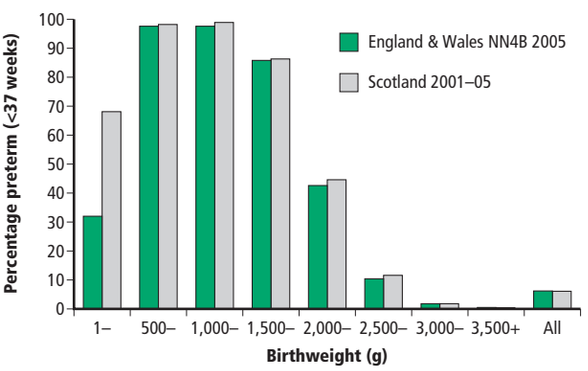


Figure 3 Birthweight and gestational age data, live singletons: comparing NN4B data with Scottish data

a) Percentage low birthweight within gestational age group



b) Percentage preterm within birthweight group



NN4B data have a very similar gestational age distribution to the other datasets except that they indicate a higher percentage of births at very low gestational ages.

The NN4B birthweight for gestational age data for live singletons are compared with the Scottish data in Figure 3. The data on the percentage low birthweight within gestational age groups show close agreement between the two datasets (Figure 3a). The percentage preterm within birthweight groups also shows close agreement, with the exception of the birthweight group 1–499g where the NN4B data indicate a much lower percentage born preterm (Figure 3b). This difference results from the NN4B birthweight and/or gestation recording errors identified earlier.

Discussion

This paper has presented and validated a new dataset, NN4B, that has the potential to enhance what is known about births in England and Wales. We have demonstrated the data to be essentially of good quality. In particular gestational age data are credible, consistent with other data sources, and the statistics by maternal age, multiplicity and sex are as expected. The gestational age statistics presented, for all births in England and Wales in 2005, provide a reference against which to monitor trends in the percentage born preterm, the absolute number of preterm births, and changes at the upper and lower end of the gestational age distribution. This information can inform the planning of neonatal care provision. It can also be used by clinicians to compare their local populations against the national distribution and offers the potential to monitor size for gestational age. The information can furthermore be used to investigate differentials across population subgroups and in making international comparisons.

The consistency and validation checks, both internally and in comparison with other data sources, have indicated some problems. Firstly, there were slightly fewer NN4B records for births in 2005 than there were birth registration records. This discrepancy will be investigated. However, it is unlikely that the statistics presented in this paper will be affected materially. Secondly, the opportunity to relate birthweight to gestational age information has highlighted some implausible combinations of the two characteristics. These are most evident amongst the lowest weight and gestational age groups. In our study 109 live singletons were under 20 weeks gestational age, of which half did not have low birthweight. Live birth before 20 weeks gestation is itself unusual and birthweights of 2,500g or more in this group are highly improbable. Our recently published analysis of gestation-specific infant mortality[20] gave rise to serious doubt as to whether the births recorded as being of under 22 weeks gestational age with birthweights 1,000g and over were truly of extremely low gestational age. Focussing on all babies recorded as born before 24 weeks gestational age, we plan to investigate the inconsistencies in the recording of gestational age and birthweight information by cross-validating the information against that collected by the Confidential Enquiry into Maternal and Child Health (CEMACH).[26] However, going back to the original records is not always feasible in large routine datasets.

Table 6 **Comparison of NN4B and birth registration data, 2005**

England and Wales

	NN4B data		Birth registration data	
Live births as percentage of all births	99.5		99.5	
Multiple births as percentage of all births	2.9		3.0	
Sex ratio (male/female births)	1.05		1.05	
Maternal age distribution of all births (per cent)[1]				
under 20	6.9		6.9	
20-24	18.9		18.9	
25-29	25.4		25.4	
30-34	29.1		29.1	
35-39	16.2		16.1	
40 and over	3.5		3.5	
Birthweight distribution of births (per cent)	live births	stillbirths	live births	stillbirths
1g–	0.3	9.1	0.1	8.4
500g–	0.4	22.5	0.4	24.0
1,000g–	0.7	11.8	0.7	12.6
1,500g–	1.5	10.0	1.5	10.7
2,000g–	4.7	10.9	4.8	11.0
2,500g–	16.9	12.4	17.0	12.0
3,000g–	35.6	12.8	35.6	11.2
3,500g–	28.5	6.3	28.5	5.7
4,000g–	9.3	2.1	9.3	1.8
4,500g+	1.7	0.8	1.7	1.0
not stated	0.4	1.5	0.3	1.6
Gestational age distribution of stillbirths (per cent)				
24–27 wks	25.3		26.7	
28–31 wks	16.0		16.6	
32–34 wks	12.4		12.9	
35–36 wks	10.1		10.2	
37–38 wks	13.3		12.5	
39–41 wks	21.5		18.5	
42+ wks	1.4		0.9	
not stated	0.0		1.6	
Number of birth records[2]	649,174		649,379	

1 Births with stated mother's age.
2 Extract of registration records taken August 2006.

In Scotland unusual combinations of birthweight and gestation are queried at source as part of the data processing.[27] Different criteria for excluding records with implausible weights for gestational age tested using routine Canadian data found all methods excluded relatively more larger babies and more babies who subsequently died.[28] Therefore, a priori exclusion is inadvisable as it runs the risk of introducing bias. While it is important to try and resolve these issues, they do only affect a small number of records and, importantly, the percentage of live singletons born preterm (6.2 per cent) remains unchanged regardless of whether births of under 22 weeks gestational age are included or not.

It is useful here to draw on the findings (described in full elsewhere[29]) from the pilot linkage of NN4B and registration records for births in the first quarter of 2005. This identified some records with discrepancies between the NN4B and registration birthweight information. Although 98.5 per cent of linked records had identical birthweights, the groups most affected by the discrepancies were those with weights recorded in NN4B data as missing or less than 500g. A substantial number of weights recorded in NN4B as 1–499g were recorded in the registration information with one or more additional zeros. Over 95 per cent of the linked records with unstated birthweight in the NN4B data had a value in the registration data. These differences are probably due to subsequent updating or correction of records by maternity staff, child health staff, registrars or ONS.[22]

The pilot linkage contributed other information regarding NN4B data quality. The gestational age at birth of the stillbirth registrations that linked with an NN4B record agreed in 81.8 per cent of records; in 13.5 per cent there was a difference of one week, and in 4.7 per cent a difference of two or more weeks. Linkage of the NN4B and registration records provides the opportunity to improve the data held in either source, where a data item is held in both sources. Birthweight and mother's date of birth from an NN4B record was available for almost all of the birth registrations missing this information. The linkage also found that live/stillbirth status and multiple birth status were each differently classified in approximately one in 1,000 records.

The NN4B data specification asks for gestation length in weeks 'calculated from relevant menstrual data held within the maternity system'. This is in accordance with the internationally accepted definition of gestational age as the number of completed days or weeks since the first day of the last normal menstrual period (LMP).[25] However it is now widely accepted that ultrasound, first introduced into clinical obstetric practice in the 1970s, gives a better estimate of gestational age than LMP does.[30,31] While the convention of calculating the duration of gestation from the LMP still persists, there is an increasing tendency to use the gestational age assessed by ultrasound now that second trimester scans are a routine part of antenatal assessment in Britain. A study of births of 27/28 weeks gestation in England, Wales and Northern Ireland between 1998 and 2000 found that 79 per cent of the mothers had had an ultrasound before 20 weeks gestation, and 85 per cent had had menstrual history recorded.[10] In a study of maternity units in north west London, 95 per cent of all mothers between 1988 and 2000 had received a midtrimester ultrasound.[9] Gestational age distributions have been shown to differ depending on the method used to assess gestational age. Studies have shown that if second trimester ultrasound is used in place of LMP, the mean gestational age at birth is one week earlier, the proportion of preterm births is slightly higher, and the proportion of post term births lower.[32,33,34]

No information about the method of gestational age assessment is recorded on NN4B records. We may assume that common practice is to submit the estimate in use at the time of labour but it is possible that in some cases gestational age based on LMP is submitted, as requested, even if it has been revised in the light of an ultrasound scan. Either way, gestational age in completed weeks is a truncated index, with the actual gestational age being 0–6 days longer. The distribution of births on each of these days is not equal; before 40 weeks the distribution of births at a given gestational age is skewed to later in the week, after 40 weeks to earlier in the week. A further issue concerning the measurement of gestational age is the fact that the full gestational age distribution, while broadly normal, is highly skewed to the left and mean gestation statistics must therefore be interpreted with caution.

Aspects of clinical obstetric practice, such as the use of ultrasound to assess gestational age, induction of labour, and elective Caesarean section, are likely to affect the distribution of gestational ages. As such they may contribute to any differences in the gestational age distributions observed between countries, and also by maternal age groups.

Conclusion

This paper has described the new NN4B data and shown their overall quality to be good, although more exploration is needed of certain aspects of the data. The gestational age statistics presented for England and Wales, and previously unavailable routinely, can inform the planning of neonatal care provision, and provide a reference against which to monitor trends in preterm

births, and investigate differentials across population subgroups. From now onwards, information available from the NN4B data that is not collected at birth registration, such as gestational age, will be analysed by ONS and will form part of National Statistics on births in England and Wales.

Key findings

- NHS Numbers for Babies (NN4B) data are generally of good quality
- The gestational age data are credible, consistent with other data sources, and the statistics by maternal age, multiplicity and sex are as expected
- The percentage of singleton live births born preterm is lowest in babies of mothers aged 30–34 as compared to all other maternal age groups
- The gestational age statistics, previously unavailable for England and Wales as a whole, provide a reference against which to monitor trends in preterm births, and can inform the provision of neonatal care
- The NN4B dataset is a powerful new source which can extend what is known about births in England and Wales

Acknowledgements

This study would not have been possible without the help and support from the NHS Numbers for Babies team during the initial stages of obtaining the data. We are also grateful to Louise James, Health Solutions Wales and Gwyneth Thomas, Health Statistics and Analysis Unit, Welsh Assembly Government for supplying us with the NCCHD data for Wales; and to Morven Ballard, Etta Shanks and James Chalmers from Information Services Division NHS National Services, Scotland for supplying us with SMR02 data for Scotland. We thank Alison Macfarlane, City University for invaluable advice on earlier drafts.

References

1. Brocklehurst P (1999) 'Infection and preterm delivery', *British Medical Journal* **318**, 548–549.
2. NHS Scotland, Information Services Division (2006) *Scottish perinatal and infant mortality and morbidity report 2005*, ISD: Edinburgh.
3. Marlow N, Wolke D, Bracewell M A *et al* (2005) 'Neurologic and developmental disability at six years of age after extremely preterm birth', *New England Journal of Medicine* **352(1)**, 9–19.
4. Cooke R W (2005) 'Perinatal and postnatal factors in very preterm infants and subsequent cognitive and motor abilities', *Archives of Disease in Childhood Fetal & Neonatal Edition* **90(1)**, F60-F63.
5. Gardosi J, Mul T, Mongelli M *et al* (1998) 'Analysis of birthweight and gestational age in antepartum stillbirths', *British Journal of Obstetrics and Gynaecology* **105(5)**, 524–530.
6. Hutton J L, Pharoah P O, Cooke R W *et al* (1997) 'Differential effects of preterm birth and small gestational age on cognitive and motor development', *Archives of Disease in Childhood Fetal & Neonatal Edition* **76(2)**, F75-F81.
7. Royal College of Obstetricians and Gynaecologists Clinical Effectiveness Support Unit (2001) *Induction of Labour. National Evidence-based clinical guideline Number 9*, RCOG Press: London.
8. Sairam S, Costeloe K, Thilaganathan B (2002) 'Prospective risk of stillbirth in multiple-gestation pregnancies: a population-based analysis', *Obstetrics & Gynecology* **100(4)**, 638–641.
9. Balchin I, Whittaker J C, Patel R *et al* (2007) 'Racial variation in the association between gestational age and perinatal mortality: prospective study', *British Medical Journal* **334**, 833–837.
10. Confidential Enquiry into Stillbirths and Deaths in Infancy (2001) *8th Annual Report*, Maternal and Child Health Research Consortium: London.
11. Macfarlane A, Mugford M (2000) *Birth Counts. Statistics of pregnancy and childbirth. Volume 1*, The Stationery Office: London.
12. Statistical Directorate, National Assembly for Wales, National Community Child Health Database (NCCHD): First data, 2004, accessed 14 Jan 2008, available at: www.new.wales.gov.uk/topics/statistics/headlines/health–2007/health–2006/health–2005/hdw20050712/?lang=en
13. The Information Centre for Health and Social Care (2006) *NHS Maternity Statistics, England: 2004–05*, The Information Centre, Community Health Statistics: Leeds.
14. NHS Scotland, Information Services Division, Hospital based maternity and birth data (SMR02), accessed 14 Jan 2008, available at: www.datadictionaryadmin.scot.nhs.uk/isddd/9066.html
15. All Wales Perinatal Survey (2007) *Annual Report 2005*, Cardiff University: Cardiff, accessed 14 Jan 2008, available at: www.cardiff.ac.uk/medic/aboutus/departments/childhealth/ourresearch/neonatalmedicine/allwalesperinatalsurvey/index.html
16. Zeitlin J, Wildman K, Breart G *et al* (2003) 'PERISTAT: Indicators for monitoring and evaluating perinatal health in Europe', *European Journal of Public Health* **13**(3 supplement), 29–37.
17. Blondel B, Macfarlane A, Gissler M *et al* and the PERISTAT Study Group (2006) 'Preterm birth and multiple pregnancy in European countries participating in the PERISTAT project', *British Journal of Obstetrics and Gynaecology* **113**, 528–535.
18. Connecting for Health, NHS Numbers for Babies, accessed 14 Jan 2008, available at: www.connectingforhealth.nhs.uk/systemsandservices/nhsnumber/nn4b
19. Office for National Statistics (2007) Preterm births, England and Wales, 2005. Available on the National Statistics website at: www.statistics.gov.uk/statbase/Product.asp?vlnk=14882&Pos=&ColRank=1&Rank=272
20. Moser K, Macfarlane A, Chow Y H *et al* (2007) 'Introducing new data on gestation-specific infant mortality among babies born in 2005 in England and Wales', *Health Statistics Quarterly* **35**, 13–27. Available on the National Statistics website at: www.statistics.gov.uk/downloads/theme_health/HSQ35.pdf Supplementary tables are available on the National Statistics website at: www.statistics.gov.uk/downloads/theme_health/Gestation-HSQ35/gest_spec_infantmortality.pdf
21. Office for National Statistics (2006) *Birth statistics, England and Wales, 2005*, Series FM1 No. 34, ONS: London.
22. Office for National Statistics (2007) *Mortality statistics: Childhood, infant and perinatal, England and Wales, 2005*, Series DH3 No. 38, ONS: London.
23. The National Community Child Health Database for Wales, Information Products Unit, Health Solutions Wales.
24. SMR02, Information Services Division, NHS National Services Scotland.
25. World Health Organisation, Article 23 of the Constitution of the World Health Organisation Off. Rec. Wld Hlth Org 1950 (28)17, 1967 (160), 11 and Annex 18, and 1976 (233)18.
26. Confidential Enquiry into Maternal and Child Health (2007) *Perinatal Mortality 2005: England, Wales and Northern Ireland*, CEMACH: London.
27. Information Services Division, SMR02 Data Dictionary: birthweight, ISD, NHS National Services Scotland, accessed 14 Jan 2008, available at: www.datadictionaryadmin.scot.nhs.uk/isddd/9686.html
28. Joseph K S, Kramer M S, Allen A C *et al* (2001) 'Implausible birth weight for gestational age', *American Journal of Epidemiology* **153(2)**, 110–113.
29. Hilder L, Moser K, Dattani N *et al* (2007) 'Pilot linkage of NHS Numbers for Babies data with birth registrations', *Health Statistics Quarterly* **33**, 25–33.

30. Alexander G R, de Caunes F, Hulsey T C *et al* (1992) 'Validity of postnatal assessments of gestational age: a comparison of the method of Ballard et al and early ultrasonography', *American Journal of Obstetrics and Gynaecology* **166(3)**, 891–895.
31. Chervenak F A, Skupski D W, Romero R *et al* (1998) 'How accurate is fetal biometry in the assessment of fetal age?', *American Journal of Obstetrics and Gynaecology* **178(4)**, 678–687.
32. Mongelli M, Gardosi J (1996) 'Gestation-adjusted projection of estimated fetal weight', *Acta Obstetricia et Gynecologica Scandinavica* **75(1)**, 28–31.
33. Savitz D A, Terry J W Jr., Dole N *et al* (2002) 'Comparison of pregnancy dating by last menstrual period, ultrasound scanning and their combination', *American Journal of Obstetrics and Gynaecology* **187(6)**, 1660–1666.
34. Yang H, Kramer M S, Platt R W *et al* (2002) 'How does early ultrasound scan estimation of gestational age lead to higher rates of preterm birth?', *American Journal of Obstetrics and Gynaecology* **186(3)**, 433–437.