An evaluation of three methods used to assess gestational age of aboriginal neonates.

Article in Journal of Paediatrics and Child Health · July 1995

2 authors:

Susan Sayers
Charles Darwin University
57 PUBLICATIONS 445 CITATIONS

Jennifer Robyn Powers
University of Newcastle
67 PUBLICATIONS 1,248 CITATIONS
An evaluation of three methods used to assess the gestational age of Aboriginal neonates

S. M. SAYERS and J. R. POWERS

Menzies School of Health Research and Royal Darwin Hospital, Darwin, Northern Territory, Australia

Abstract. Many Aboriginal women do not recall their last menstrual period date, so alternative methods of estimating gestational age are necessary for optimal obstetric and neonatal care. In this retrospective review of 605 Aboriginal infants born at the Royal Darwin Hospital, the gestational age was estimated by the Dubowitz method and compared with available gestational age estimates from first fundal height and first ultrasound measurement. There was good agreement between the Dubowitz and ultrasound estimates of gestational age with best agreement occurring when ultrasound was done in the first trimester and worst agreement in the third trimester. Agreement between fundal height and Dubowitz estimates was poor but the measurement of fundal height was not standardized. When accurate last menstrual period information is absent, these findings suggest that good estimates of gestational age in Aboriginal neonates can be determined from the Dubowitz assessment at birth and from ultrasound measurements taken in the first trimester.

Key words: Aboriginal; foetus; gestational age; neonate.

Despite improvements over the last two decades, mortality and morbidity rates of Aboriginal neonates and infants are still high relative to the Caucasian rates. Previous studies of Aboriginal infants have focused on the high incidence of low birthweight and its contribution to morbidity and mortality. However low birthweight is now recognized to be heterogeneous and includes preterm infants with fetal growth within the standard range, and small-for-date infants with fetal growth outside the standard range for gestational age. Because these two groups of infants differ in aetiology, morbidity and mortality, it is important to be able to distinguish between them clinically and epidemiologically. This requires accurate gestational age data particularly to avoid misclassification of small-for-date infants as preterm. In clinical practice, the estimated gestational age is usually obtained from maternal recall of the last menstrual period, often in association with early fetal ultrasound measurements. Unfortunately women from many traditional societies do not record their last menstrual period date. For example a study of Aboriginal women from East Arnhem Land showed that only 12% could state the date of their last period. Alternative methods of gestational age estimations are therefore necessary.

At the Royal Darwin Hospital (RDH), fundal height and fetal ultrasound measurements are used to estimate fetal maturity for the obstetric management of Aboriginal pregnancies, and the Dubowitz scoring system is used to estimate the gestational age of Aboriginal neonates. Previous studies have established the accuracy of fundal height measurements, fetal ultrasound measurements and the Dubowitz scoring system for non-Aboriginal populations using accurate last menstrual period dates. The Dubowitz scoring system has also been validated using reliable last menstrual data in Nigerian born infants, and in Bantu, Indian, and Malay infants in Capetown. The determination of gestational age by these methods has not yet been evaluated for the Aboriginal population.

The aim of this study was to assess these methods as they are currently used at the RDH for the estimation of gestational age in Aboriginal neonates. Because accurate last menstrual period data were not available, this study compared the gestational age estimations obtained from the fetal ultrasound and fundal height methods with the Dubowitz scoring system. Estimates of gestational age by ultrasound are reported to be more accurate when ultrasound is performed in the first trimester. Therefore subsets of ultrasound measurements taken at different times during gestation were also compared with Dubowitz score estimations.

METHODS

Subjects

The RDH serves a population of approximately 110,000 people. This represents 85% of the population of the Northern Territory, and consists of approximately 20,000 Aborigines and 90,000 non-Aborigines. The percentage of Aboriginal women having babies outside the hospital is low, and in 1987, 89.2% and 1988, 90.7% of Aboriginal mothers delivered in the hospital. In this study the majority of mothers (84.6%) are routinely referred from the Darwin health region (Table 1), some come by choice or are specifically referred from the nearby health regions of East Arnhem and Katherine (13.9%) and a few (1.5%) are high risk pregnancies referred from northern Western Australia.
The neonates were liveborn singletons delivered at the RDH, between January 1987 and March 1990. During this time a total of 1207 neonates were recorded in the Delivery Suite Birth Register as Aboriginal, defined as a neonate born to a woman of Aboriginal or Torres Strait Islander descent, who identifies herself as Aboriginal or Torres Strait Islander and is accepted as such by the community in which she lives. The paediatric investigator was absent for 471 of these deliveries and could not examine a further 131 neonates within 4 days of delivery, as required for the Dubowitz assessment. The neonate could not be examined mainly because the paediatric investigator could not find the mother in RDH or its environs, despite repeated attempts; otherwise the mother and baby were discharged early or had absconded.

The remaining 605 neonates had an examination including a clinical estimation of gestational age using the Dubowitz scoring system. Gestational age estimates based on fetal ultrasound measurements and fundal height measurements were obtained by a retrospective review of the obstetric case notes of all these neonates. To check for any bias in the selection of neonates, mean birthweights and male/female ratios were compared for the different gestational age assessment groups (Table 1). There were no significant differences in the birthweights or sex ratio for each of the groups. Health region data were not collected for non-Aboriginal neonates.

Exclusions in the analysis were multiple pregnancies. By observation and direct questioning of the mothers, 73% of the neonates examined were thought to be Aboriginal and 27% of Aboriginal descent; neonates of Aboriginal descent mostly had Aboriginal mothers and non-Aboriginal fathers.

Procedures

The gestational age was estimated from the first fundal height and the earliest fetal ultrasound measurement recorded in the obstetric notes. The first fundal height in the case notes was recorded by any one of 40 observers (Aboriginal health workers, community clinic sisters, district medical officers, general practitioners, hospital staff and specialists). As fundal height measurement has not yet been standardized, it is not possible to be specific about the different methods used; observers used anatomical landmarks, linear measurements or other methods. The fetal ultrasound measurements were made by seven self-taught district medical officers, of varying expertise and experience, using a single portable ultrasound machine in the rural health community clinics (Alaska Echo camera, 210 model US-74F). For gestational age assessment, the crown-rump length was measured in the first trimester and the biparietal diameter was measured in the second trimester.

The paediatric investigator examined the Aboriginal neonates in the postnatal and intensive care wards of the RDH. Gestational age was estimated according to the neurological and external physical criteria according to the Dubowitz scoring system. As the Dubowitz scoring system was central to the evaluation of the three different methods of gestational age estimation, the paediatric investigator’s accuracy was checked by estimating the gestational age in 30 non-Aboriginal neonates whose mothers accurately recalled their last menstrual period dates. Using the Dubowitz scoring system, the paediatric investigator underestimated the gestational ages of this sample by an average of 4 days compared with last menstrual period dates. For the general population, the investigator’s estimates of gestational age by the Dubowitz scoring system lie within ± 1.6 weeks and 2.7 weeks of gestational age based on last known menstrual period dates, which is close to the ± 2 weeks 95% confidence interval reported for the Dubowitz scoring system.

Analysis

Using a gestational age calculator (Down Bros. Mayer Phelps Ltd) an expected date of confinement was determined from the first fundal height measurements, the fetal ultrasound measurements and the Dubowitz score. The ideal way to validate these three methods would be to compare them with the valid standard based on last menstrual period data. As such a standard was not available and the true values of the gestational ages were unknown, indirect methods of evaluation assessing the degrees of agreement of the gestational age estimations between two methods were undertaken.

The estimations of gestational age may vary as a result of errors, as well as differences between subjects and the method used to estimate gestational age. Two-factor analysis of variance was used to determine the variance in gestational age explained by subjects, methods and errors. The agreement between methods was then given by the intraclass correlation coefficient using the formulae derived by Bartko. When the mean square terms for methods, subjects and errors are MSM, MMS, and MSE, and there are m methods and n subjects, the intraclass correlation coefficient (ICC) is

\[
\text{ICC} = \frac{(MSS - MSE) / (m - 1)MSE + M(MSM - MSE) / n}
\]

ICC values of less than 0.4 indicate poor agreement and values more than 0.75 indicate good agreement.

A simpler way of measuring the degree of agreement between two methods is to plot the difference between the estimates against their mean. This has the advantage of displaying the between method differences visually. From these data, the limits of agreement can be calculated to lie between \( d - 2s \) and \( d + 2s \)

<table>
<thead>
<tr>
<th>Gestational age assessment</th>
<th>Total no.</th>
<th>% Darwin health region</th>
<th>Male/female ratio</th>
<th>Birthweight mean (s.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All neonates</td>
<td>1207</td>
<td>-</td>
<td>1.092</td>
<td>3032 (647)</td>
</tr>
<tr>
<td>Dubowitz score</td>
<td>605</td>
<td>84.6</td>
<td>1.017</td>
<td>3012 (660)</td>
</tr>
<tr>
<td>Fundal height</td>
<td>224</td>
<td>91.5*</td>
<td>1.000</td>
<td>3058 (584)</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>344</td>
<td>87.8</td>
<td>1.072</td>
<td>3038 (639)</td>
</tr>
<tr>
<td>Ultrasound &lt; 31 weeks</td>
<td>310</td>
<td>88.7</td>
<td>1.081</td>
<td>3032 (657)</td>
</tr>
</tbody>
</table>

* Significantly more neonates from the Darwin health region than the Dubowitz group.
where $d$ is the mean difference and $s$ the standard deviation of the differences. The mean of the difference in gestational age estimates and its 95% confidence interval can also be determined.

**RESULTS**

The agreement between the estimates of gestational age obtained by Dubowitz score, fetal ultrasound and fundal height measurements is shown in Table 2. While the mean difference between Dubowitz and fundal height estimations was only $-4$ days for the neonates sampled, the limits of agreement for all neonates could range between $-7.0$ weeks and $5.9$ weeks. The agreement was best between the Dubowitz and ultrasound estimates, with an intraclass correlation coefficient of $0.69$ and limits of agreement of $-3.3$ to $4.0$ weeks. However the degree of agreement between Dubowitz and ultrasound estimates is better when the ultrasound measurement is performed before $14$ weeks gestation and unsatisfactory if taken after $30$ weeks gestation (Table 3). The agreement between the Dubowitz estimations and fundal height estimations taken at the same intervals in gestation remained poor with the intraclass correlation coefficient remaining below $0.48$ for all the gestational intervals analysed.

Comparison of gestational ages, as estimated by the Dubowitz score, for each of the gestational age assessment groups showed no significant differences in the groups for mean gestational age nor the per cent of preterm neonates (Table 4). Although the fundal height group was slightly more mature and the early ultrasound group (ultrasound < 14 weeks) less mature than the other groups, these differences in mean gestational age were not significant.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Agreement of Dubowitz score with ultrasound and fundal height estimates of gestational age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method of estimation</td>
<td>No. neonates</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>344</td>
</tr>
<tr>
<td>Fundal height</td>
<td>224</td>
</tr>
</tbody>
</table>

CI = confidence interval.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Agreement of Dubowitz score with ultrasound estimates of gestational age by age when ultrasound done</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestational age when ultrasound done (weeks)</td>
<td>No. neonates</td>
</tr>
<tr>
<td>&lt;14</td>
<td>23</td>
</tr>
<tr>
<td>14 to 20</td>
<td>129</td>
</tr>
<tr>
<td>21 to 25</td>
<td>97</td>
</tr>
<tr>
<td>26 to 30</td>
<td>61</td>
</tr>
<tr>
<td>&gt;30</td>
<td>34</td>
</tr>
</tbody>
</table>

CI = confidence interval.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Preterm percentage and mean gestational age (estimated by Dubowitz scoring system) of neonates assessed by the different methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method of gestational age assessment (neonates)</td>
<td>Gestational age estimated by Dubowitz score Mean (S.D.)</td>
</tr>
<tr>
<td>Dubowitz score (605)</td>
<td>38.7 (2.0)</td>
</tr>
<tr>
<td>Fundal height (224)</td>
<td>38.9 (1.9)</td>
</tr>
<tr>
<td>Ultrasound (344)</td>
<td>38.7 (2.1)</td>
</tr>
<tr>
<td>Ultrasound &lt;31 weeks (310)</td>
<td>38.7 (2.1)</td>
</tr>
<tr>
<td>Ultrasound &lt;14 weeks (23)</td>
<td>38.3 (2.8)</td>
</tr>
</tbody>
</table>

CI = confidence interval.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Agreement of Dubowitz score with ultrasound estimates of gestational age by birthweight and race</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td>No. neonates</td>
</tr>
<tr>
<td>Birthweight &lt;2500 g</td>
<td>54</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>252</td>
</tr>
<tr>
<td>Aboriginal descent</td>
<td>93</td>
</tr>
</tbody>
</table>

CI = confidence interval.
The agreement between the Dubowitz estimates and the ultrasound estimates was not affected by the Aboriginality (Aboriginal or Aboriginal descent) of the neonates but was unsatisfactory for low birthweight neonates (Table 5).

DISCUSSION

Many Aboriginal women do not comply with medical care and a number of difficulties arise when working in this area. Antenatally, Aboriginal women often present late in the first trimester and postnatally, well Aboriginal women have a reluctance to remain in their beds or even in the hospital. An abscending rate of 5% is reported for those admissions related to complications of pregnancy, childbirth and the puerperium.13 Because of these problems with neonatal numbers and the concern that the retrospective design of the study may have resulted in bias, comparisons of gestational age, preterm rate, birthweight, sex ratio and health region were done between each gestational age subgroup and the entire sample. The only significant difference found was that the proportion of neonates from the Darwin health region in the fundal height group was greater than in the Dubowitz group (Tables 1, 4). Mean birthweights and sex ratios for each subgroup were consistent with those of all Aboriginal neonates born between January 1987 and March 1990 (Table 1).

The accuracy of fundal height measurements, fetal ultrasound measurements and Dubowitz scoring systems for the estimation of gestational age has been established previously for specific populations using accurate last menstrual period dates. Because of the lack of last menstrual period data in Aboriginal mothers, this study compared two methods used in the routine antenatal care of Aboriginal women with a method that had been checked against known last menstrual period data in non-Aboriginal women.

Many studies give the correlation coefficient of the results of two methods as an indicator of agreement. However the correlation coefficient can only indicate the relationship between two variables and not the degree of agreement between them. This is demonstrated in a gestational age study investigated by Bland and Altman.18 In this example the babies with a gestational age of 35 weeks by one method had gestational ages between 34 and 39.5 weeks by another method. This is a relatively poor agreement, yet the correlation coefficient was high (0.85). An appropriate way to analyse these data is to examine the degree of agreement of the estimations of the fetal ultrasound and the fundal height with the estimations of the Dubowitz scoring system.

Fundal height

Considering the large number of clinicians using different methods of measurement and the lack of uniform reference tables, it is not unexpected that the estimations of gestational age based on the fundal height measurement have a poor agreement with the Dubowitz method, and few conclusions can be drawn from this retrospective study. However, when ultrasound measurements are not readily available, the fundal height measurement is probably still the best alternative in developing countries. Recent evidence supports the use of fundal height measurements for the estimation of gestational age and the detection of retarded intra-uterine growth, provided appropriate reference graphs are used and standardized measurements recorded.20,21,22 A prospective study using standardized fundal height methods could be more conclusive in determining the accuracy of fundal height estimates of gestational age in Aboriginal neonates.

Fetal ultrasound

Fetal ultrasound measurements are becoming more readily available in the developing world. An aim of the WHO 'Health for all by the year 2000' programme is the construction of a simple low maintenance portable Echoscope.23 The feasibility of Aboriginal health workers performing ultrasound measurements in Arnhem Land has already been demonstrated.24

The accuracy of the fetal ultrasound measurements in this study is doubtful due to the number of observers of varying expertise, and further investigation using one experienced ultrasonographer is necessary. Nevertheless, the gestational age estimates based on the ultrasound measurements agreed relatively well with those derived from the Dubowitz scoring system. The earlier the measurements were taken in gestation, the better the degree of agreement, and ultrasound estimates based on measurements taken < 14 weeks had a very good intraclass correlation with the Dubowitz estimate of 0.88. This finding is consistent with other studies that have validated fetal ultrasound measurements with last menstrual period dates.2,5

In this study the majority of measurements (226 out of 344) were taken in the 14–26 week gestational age interval. This is a time when gestational age can still be accurately determined, but with the added clinical benefit of viewing fetal morphology, placenta site and amniotic fluid volume. There was a mean difference of 2 days between the Dubowitz and ultrasound estimate during this interval with the 95% confidence interval of this mean being 1 to 4 days, although the limits of agreement of -3.4 to 4.0 weeks are a little more than could be clinically acceptable.

Sabbagh has argued that ultrasound measurements performed after 26 weeks gestation are unsatisfactory predictors of gestational age.3 However, in this study a poor intraclass correlation (0.13) for estimates calculated from ultrasound measurements did not occur till after 30 weeks gestation (Table 3), suggesting that gestational age estimates from ultrasound measurements may be valid as late as 30 weeks gestation.

Dubowitz scoring system

The assessment of gestational age of newborns using a combination of neurological and external characteristics has been described for some years. Different workers using combinations of criteria have developed various methods and scores.26,27,28 The Dubowitz scoring system supported by a manual has been the most thoroughly documented and widely used method in clinical practice with reported 95% confidence limits of ± 2 weeks.

This study evaluates for the first time the Dubowitz scoring system for the estimation of gestational age in Aboriginal neonates. The gestational age estimates based on the Dubowitz scoring system showed good agreement with the estimates based on the fetal ultrasound measurements taken throughout pregnancy, although as expected the degree of agreement was best with ultrasound measurements taken in the first 14 weeks.
There has been some concern about the Dubowitz scoring system overestimating the gestational age in low birthweight infants.\(^{29,30}\) In this study, the Dubowitz estimate agreed fairly well with the fetal ultrasound estimate among neonates weighing < 2500 g, as judged by the intraclass correlation coefficient. However, the degree of agreement between the estimates, with a mean difference of 9 days and the limits of agreement of −2.9 to 5.4 weeks, tends to confirm the positive bias of Dubowitz scores in low birthweight infants.

Certain criteria of the Dubowitz scoring system were difficult to interpret in the Aboriginal neonates. For example, some criteria were influenced by the nutritional state of the neonate. Aboriginal neonates with loose skin folds and thin subcutaneous tissue had mature neurological scores but immature external signs, as poor subcutaneous tissue decreased the bulk of the labia majora and so decreased the maturity score. Similarly, some malnourished neonates had small breast nipples and soft ear cartilage. Breast nodule growth is determined by a complex interplay of hormonal influences\(^{31}\) and perhaps be influenced by placental insufficiency. Very dark Aboriginal neonates made the scoring of skin colour and opacity difficult; a difficulty previously noted with Ugandan born infants.\(^{28}\) Other investigators have found that if the assessment was done within a few hours of birth, skin colour and opacity was easier to assess in Nigerian infants.\(^{11}\) Some cases of immature ear forms seemed to be racial or familial rather than a reflection of maturity, as the mother was noted to have similar ears. Even so, the estimations of gestational age from the Dubowitz scoring system agreed well with those obtained from the fetal ultrasound measurements, giving support to its use in Aboriginal neonates.

CONCLUSION

In a study of 344 Aboriginal infants born at the RDH, good agreement was found between the Dubowitz and ultrasound estimates of gestational age with an intraclass correlation coefficient of 0.69 being obtained. Agreement was greatest in the first 14 weeks (ICC = 0.88 for 23 infants) and least in the last 9 weeks (ICC = 0.13 for 34 infants). Fundal height agreed poorly with the Dubowitz scoring system, probably because of the different techniques used. A prospective study with standardized methods of measurement and recording of fundal height may be more conclusive.

When accurate information about the last menstrual period is absent, as is often the case with Aboriginal mothers, these findings suggest that valid estimates of gestational age in Aboriginal neonates can be determined from the Dubowitz score at birth or from ultrasound measurements in the first trimester.

ACKNOWLEDGEMENTS

We thank Professor J. D. Mathews, Dr T. Nienhuys and Dr A. Plant for their assistance in the form and content of the manuscript; Dr A. C. Walker for his encouraging encouragement; William Craig, Hilary Bowers and Aletta Dawes for their help in collating the data; the Obstetric and Paediatric Departments of the Royal Darwin Hospital for access to the patients and records in their departments; colleagues of the Menzies School of Health Research for advice and assistance in aspects of the work; and most important of all the Aboriginal mothers who agreed to their babies being part of this study. This study was partially funded by the National Health and Medical Research Council.

REFERENCES

Dear Editors

13 January 1994

AN EVALUATION OF THREE METHODS USED TO ASSESS GESTATIONAL AGE OF ABORIGINAL NEONATES

Recently, on reviewing the neonates in this study we regrettably found fetal ultrasound measurements recorded in the case notes that were not included in our published article (J. Paediatr. Child Health 1992; 29: 312–17). We would like to correct this omission. We have repeated the analysis and confirmed that the conclusions of the original article remain the same.

In a retrospective review of 604 Aboriginal infants born at the Royal Darwin Hospital the gestational age was estimated by the Dubowitz method and compared with 552 gestational age estimates from fetal ultrasound measurements. There was good agreement between the Dubowitz and ultrasound estimates of gestational age with best agreement occurring when the ultrasound was done in the first trimester and worst agreement in the third trimester. When accurate last menstrual period information is not available, these findings suggest that good estimates of gestational age in Aboriginal infants can be determined from the Dubowitz assessment at birth and from ultrasound measurements taken in the first trimester.

J. R. POWERS
Menzies School of Health Research
Darwin, NT 0800
Australia