Effectiveness of screening for congenital dislocation of the hip

E G KNOX, E H ARMSTRONG, AND R J LANCASHIRE
From the Department of Social Medicine, University of Birmingham, Medical School, Edgbaston, Birmingham B15 2TJ

SUMMARY The incidence of congenital dislocation of the hip (CDH) in Birmingham was measured in the period 1942–52 and again in 1950–54; it is measured now in 1974–83. Despite the introduction of neonatal screening for CDH in November 1966, between the times of the earlier and the present measurements there was no indication that CDH requiring prolonged treatment had declined. The performance of the procedure itself was poor. It detected only a third of genuine cases, and the false positives outnumbered the true positives by 10:1. The incidence of ‘definite’ cases varied markedly from year to year in a manner suggesting diagnostic idiosyncrasies rather than genuine fluctuations. There was no evidence in particular maternity units that high levels of early detection led to low levels of late diagnosis. Indeed, maternity units with the worst results—the highest rates of late diagnosis and of prolonged treatment—had the highest rates of early detection and early treatment. Current theories of natural history, aetiology, and treatment should be reappraised.

In the period 1942–52, before the introduction of neonatal screening, the incidence of treated congenital dislocation of the hip (CDH) in Birmingham was 0.65 per 1000 births.1 In the partly overlapping period 1950–54, it was 0.91 per 1000 births.2 The second investigation used a wider range of data sources. Very few infants were diagnosed in the neonatal period: about 0.03 per thousand births. A major effort was subsequently made in Birmingham, from November 1966, to screen all newborns.3 It was hoped that early detection would greatly reduce the incidence of delayed diagnosis, limit the need for surgery and prolonged treatment, and reduce the prevalence of disability.

Department of Health and Social Security (DHSS) guidelines on clinical techniques were published in 19694 and in 1986,5 but little thought was given in 1966 to the quality control of the programme as a whole, and no provision was made to measure its effectiveness in terms of a reduced incidence of late diagnosed CDH. Notification of congenital malformations to the Office of Population Census and Surveys (OPCS) on a nationwide scale began in 1964, and CDH had become a notifiable condition. Numbers of neonatal detections could thus be recorded, but limitation of the notification scheme to malformations recognised in the first 7 days of life meant that it could not monitor national or local reductions in the frequency of late presentations, the essential criterion of success.

From 1976 onwards reports began to appear which cast serious doubts upon the effectiveness of CDH programmes.6–13 They showed little evidence of substantial reductions of late presenting CDH and, in some cases, an increase. In some investigations it was suggested that the screening procedure itself, or perhaps early treatment of unstable hips, might have caused a dislocation to become established.

Our purpose in this report is to examine the problem in Birmingham using the facilities of the Birmingham Birth Register and the Birmingham Malformations Register.

Materials and methods

The Birmingham Birth Register (BBR) and Birmingham Malformations Register (BMR) maintain summarized records of all Birmingham births and all congenital malformations. The initial ascertainment of malformations is based upon birth notifications, maternity reports, and notifications of malformations to OPCS. These initial reports are later supplemented by health visitor records and by Hospital Activity Analysis (HAA) searches for admissions of children up to 5 years of age where there was mention of a malformation. Ascertainment of malformations evident before 5 years of age is essentially complete except for children who may have left the city and a small number of late “inward
Detailed classification of the malformations relies upon the reported diagnoses of clinicians in charge, supplemented in fatal cases by necropsy reports and in subsequently treated cases by HAA diagnostic codings. For "doubtful" diagnoses (including talipes, heart murmurs, clicking hips, and "generic malformations otherwise unspecified"), and for those which lacked a subsequent consistent confirmation, we examined clinical records directly. For CDH we examined the original neonatal and HAA-listed hospital records for all disorders of the hip up to the age of 5 years among children born to Birmingham residents between 1974 and 1983. We also searched diagnostic indexes of outpatient clinics of children's hospitals, paediatric and orthopaedic units, and neonatal follow-up clinics of maternity units and, where hip disorders were indexed, the clinical notes were examined. Substantial periods of treatment or surveillance would certainly have been recorded by the health visitor even if a child had not been admitted to hospital, but shorter episodes might have escaped our notice. With this additional material, the only cases we were likely to 'miss' or 'misclassify' were (a) those who moved out of the city before diagnosis, and (b) infants referred to outpatient clinics or neonatal follow-up clinics where no diagnostic index was kept and where the child was discharged from care without ever being admitted to hospital.

The poor organisation of outpatient record systems, and a general absence of diagnostic indexes, partly frustrated these additional searches. Only two maternity units and none of the orthopaedic outpatient clinics had follow-up clinics with a consistently maintained diagnostic index. Enquiries at plaster rooms and at x-ray departments proved to be fruitless. In the event, the two "indexed" neonatal clinics recorded transient instabilities of the hip in numbers sufficient to suggest the untraced attendance of similar children at other units. In-patient record systems were well kept. All HAA-indexed records were eventually traced and inspected.

These extended searches failed to reveal any undoubtedly genuine dislocations of which we had been unaware. The additional ascertainment came entirely from the indexed neonatal clinics and consisted entirely of infants receiving no treatment or short and tentative treatments or whose referrals for an additional opinion led to neither a substantive diagnosis nor additional treatment.

Results

Incidence

From 1974 through 1983 there were 144,246 total births to Birmingham residents (excluding residents of Sutton Coldfield) in the area covered by BMR and BBR. During this time BMR identified 557 children in whom CDH was at some time diagnosed. This amounts to 3.86 per thousand births. Of these 557, 314 had a "notification code"; that is, they were notified to OPCS through the national Congenital Malformations Notification Scheme. On this basis, the incidence would appear as 2.17 per 1000 total births.

Other infants were reported to OPCS and/or to BMR with a "clicking" hip. OPCS classifies them under a rubric separate from CDH, and they were not accepted as "valid" malformations within BMR. They are therefore excluded from both measures of incidence. It was possible subsequently to link late presenting cases of CDH with neonatal reports to see how many of them were in fact preceded by that condition; none was.

In order to compare the current incidence of CDH with the estimates of 1942–52 and 1950–54 we must identify comparable cases. Among the 557 children reported to have CDH, 83 had surgery or prolonged splinting of 16 weeks or more, and another 13 had shorter periods of treatment beginning at 18 weeks of age or later. These 96 cases probably represent the closest approach we can achieve to the definition used in the earlier studies, and this gives an incidence of 0.67 per 1000 total births, almost exactly the same as in 1942–52 (0.65) but somewhat less than the more reliable estimate of 1950–54 (0.91).

There have been major demographic changes in Birmingham since the earlier measures of incidence were made. In the period 1974–83, 69% of babies were to Caucasian mothers compared with the almost total predominance of Caucasian births in 1950. Ninety-three of the 96 cases mentioned above were born to Caucasians, giving an incidence of 0.93 per thousand. There is no evidence of a substantial change.

Sixty-four of the 96 cases in 1974–83 (67%) were first treated after the age of 18 weeks and seemed quite clearly to have slipped through the screen; either they were "missed" or they developed late.

Sensitivity and Specificity

In order to estimate these parameters we must try to distinguish (a) infants diagnosed and treated early, for whom a later disability was otherwise inevitable (the true positives), (b) those whose dislocations were missed or not yet present and who were diagnosed and treated later (the false negatives), and (c) children diagnosed unjustifiably in that they needed no treatment at all (the false positives).

There is no objective criterion for distinguishing "true" from "false" positives, and the best available guide is probably a record of a specialist orthopaedic
assessments and a note of the action taken. It is for this
last reason that we include an operational as well as a
diagnostic element in our definition of a true positive
(above). A “medical behavioural” definition to some
extent begs the questions in hand, but we can add to
the precision of the analysis by classifying the children
according to the ages at which they were first treated,
and the duration and extent of the treatment given.
This emphasis on the timing and duration of treatment
was to some extent enforced by the inadequacies of
the medical records that we inspected. The treatments
were more often and more consistently dated than
were the diagnoses. A two-way distribution in these
terms is presented in table 1.

Table 1 identifies the 83 children who had surgery or
prolonged splinting (mentioned earlier) and the 13
(4 + 3 + 5 + 1) who had less prolonged treatment of late
onset. These are the 96 cases used for comparison with
the earlier periods. The 32 who began a prolonged spell
of treatment starting within 18 weeks were regarded as
true positives on neonatal screening, and the 64 who began late (15 + 36 + 13) were regarded as
false negatives. The 18-week dividing line was chosen
partly because it was the antimode of the distribution;
partly because it seemed likely that if no record of
starting or modifying a substantive treatment had been
made at this age then none had been given; and
partly because increasing infant mobility beyond this
age would suggest that later diagnoses had been
reached on the grounds of visible impairments rather
than through screening.

Table 1 also identifies 114 children who were not
treated at all and suffered no subsequent harm, and
another 114 who had short early periods of treatment,
many for two or three weeks only. All 228 children can
readily be regarded as “false positives” on initial
treatment. Another 37, first splinted in the neonatal
period, had no record of referral or of duration.
Another 121 could not readily be located within the
body of the table; 23 were referred to orthopaedic
clinics, but the notes, when traced, did not mention
CDH at all, while in another 23 referred cases no notes
were found. Eight children had died in the neonatal
period and 18 had additional malformations to which
the dislocation was perhaps secondary. The remaining
children in this group had a note of intended referral in
the neonatal notes but without evidence that an
appointment had been made or that the child had been
seen. We suspect that many were seen the same day,
informally, without a separate appointment being
made or recorded. Except for the eight deaths and the
18 multiple malformations, this group of 121 must be
regarded in the main as false positives.

Another 75 children were splinted for periods of 9 to
16 weeks beginning before 18 weeks of age. It is
impossible to say exactly what proportion should be
regarded as true positives and scored as screening
“successes”, and what proportions should be regarded
as false positives. However, any substantial number of
true positives would raise current estimates of total
incidence well above those for 1950–54. In addition,
table 1 shows that the “sensitivity” of neonatal
screening—as judged by the proportion treated before
18 weeks—seems to increase as the assessed severity
of the condition (duration of treatment) decreases. We
would expect the more severe instabilities to be most
readily detectable within the neonatal period. We infer
that the majority of these positives were false,
although not necessarily all of them.

Precise estimates of the sensitivity and false-positive
rates of neonatal screening cannot be made, but
estimates using different definitions are given in

Table 1  Timing, type, and duration of treatment of CDH

<table>
<thead>
<tr>
<th>Duration of splinting (weeks)</th>
<th>&gt;17*</th>
<th>9–16</th>
<th>&lt;8</th>
<th>None</th>
<th>Duration not recorded</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No treatment</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>114(42)</td>
<td>—</td>
<td>114 (42)</td>
</tr>
<tr>
<td>Early: before 18 weeks</td>
<td>32(21)</td>
<td>75(51)</td>
<td>114(87)</td>
<td>—</td>
<td>37 (19)</td>
<td>258(178)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>15 (1)</td>
<td>4 (1)</td>
<td>3 (1)</td>
<td>—</td>
<td>—</td>
<td>22 (3)</td>
</tr>
<tr>
<td>Late: after 28 weeks</td>
<td>36 (2)</td>
<td>5 (0)</td>
<td>1 (1)</td>
<td>—</td>
<td>—</td>
<td>42 (3)</td>
</tr>
<tr>
<td>No documentation/</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>121 (88)</td>
<td>121 (88)</td>
</tr>
<tr>
<td>wrong diagnosis/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dead/misclassified/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>multiple/secondary . . .</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>83(24)</td>
<td>84(52)</td>
<td>118(89)</td>
<td>114(42)</td>
<td>158(107)</td>
<td>557(314)</td>
</tr>
</tbody>
</table>

* Includes 57 children who had surgery in addition to splinting, and 3 who had surgery alone with no separately recorded splinting.
† Figures in parenthesis indicate those children who were known to OPCS through the notification scheme. Most of them were probably recorded as CDH alone, but some (especially those with intermediate or late treatments) were notified as “multiple malformations” or as other malformations to which CDH was secondary.
Table 2  Incidence of "CDH" based on alternative definitions

<table>
<thead>
<tr>
<th>Definition of CDH</th>
<th>n</th>
<th>Incidence (10^-3)</th>
<th>False positives %*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) All cases traced</td>
<td>557</td>
<td>3.86</td>
<td>83</td>
</tr>
<tr>
<td>(b) Excluding secondary CDH, neonatal deaths, and cases without recorded treatment or referral</td>
<td>322</td>
<td>2.23</td>
<td>70</td>
</tr>
<tr>
<td>(c) Further exclusion of short or unrecorded splinting of early onset</td>
<td>171</td>
<td>1.19</td>
<td>44</td>
</tr>
<tr>
<td>(d) Surgery or prolonged treatment; or any splinting beginning after 18 weeks</td>
<td>96</td>
<td>0.67</td>
<td>0</td>
</tr>
<tr>
<td>(e) Late treated CDH (&gt;18 weeks of age)</td>
<td>64</td>
<td>0.44</td>
<td>0</td>
</tr>
<tr>
<td>(f) Previously measured incidence (1942-52) of all treated cases</td>
<td>148</td>
<td>0.65</td>
<td>0</td>
</tr>
</tbody>
</table>

* We suppose group (d) to represent the "genuine" cases of CDH, then:
  
  Sensitivity [(d-e)/d] = 33%
  Specificity [(a-d)/a](144246-d) = 99.68%

Percentage of false positives for (a) = [(a-d)/a] = 83... and similarly for subsequent rows of the final column.

Table 2. Over 80% of all the diagnosis reported to us were probably false positives. The sensitivity of the procedure was about 33%. If some of the early "positives" were "true" rather than "false", then it might be a little higher but probably not beyond 50%. Approximately 28% of "genuine" cases were known to OPCS, whether under the rubric CDH or some other rubric (another malformation), and about 90% of OPCS-notified cases were false positives.

TIME AND PLACE

Table 3 gives the distribution of the different subgroups in successive years. The "definite" cases in this table correspond with group (d) of table 2; they showed wide variations with 22 and 17 such cases in the worst years and 3 and 2 in the best. Styles of treatment also changed.

Table 3  Year of birth by diagnostic subgroup

<table>
<thead>
<tr>
<th>Year</th>
<th>&quot;Definite&quot;</th>
<th>Early (&lt;18 weeks)</th>
<th>Moderate duration (9-16 weeks)</th>
<th>Early (&lt;18 weeks) Short (&lt;9 weeks) or unrecorded duration</th>
<th>Early (&lt;18 weeks) No treatment</th>
<th>No. of births</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974</td>
<td>6 (1)*</td>
<td>3</td>
<td>13</td>
<td>6</td>
<td>14 488</td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>6 (3)</td>
<td>1</td>
<td>21</td>
<td>5</td>
<td>13 691</td>
<td></td>
</tr>
<tr>
<td>1976</td>
<td>10 (3)</td>
<td>1</td>
<td>26</td>
<td>16</td>
<td>13 448</td>
<td></td>
</tr>
<tr>
<td>1977</td>
<td>17 (5)</td>
<td>8</td>
<td>14</td>
<td>8</td>
<td>13 181</td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>12 (3)</td>
<td>14</td>
<td>12</td>
<td>8</td>
<td>13 844</td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td>22 (9)</td>
<td>7</td>
<td>17</td>
<td>8</td>
<td>14 616</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>7 (1)</td>
<td>10</td>
<td>14</td>
<td>16</td>
<td>15 433</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>11 (4)</td>
<td>14</td>
<td>11</td>
<td>8</td>
<td>15 020</td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>3 (2)</td>
<td>6</td>
<td>18</td>
<td>16</td>
<td>15 265</td>
<td></td>
</tr>
<tr>
<td>1983?</td>
<td>2 (1)</td>
<td>11</td>
<td>5</td>
<td>23</td>
<td>15 260</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>96(32)*</td>
<td>75</td>
<td>1518</td>
<td>114</td>
<td>144 246</td>
<td></td>
</tr>
</tbody>
</table>

* Figures in parentheses are of children treated in first 8 weeks of life
† Ascertainment of late presentations among 1983 births may be incomplete in 1986
‡ 114 from column 3 of table 1 plus 37 from column 5

These diagnostic idiosyncracies are illustrated again in table 4 where the incidence varies widely between the different maternity units. The diagnostic classes are regrouped here to reflect as nearly as possible the successes and failures of neonatal screening. The first column represents the substantive failures of the screening service, dislocations that were both late treated and severe (36+15, from table 1). The 196 infants in the third column are false positives. The central column is intermediate; the majority (see earlier) are probably false positives. But it also contains genuine cases who received early (although extended) treatment, and may contain others where early treatment of intermediate duration averted prolonged splinting. Part of the variation between maternity units in the last two columns reflects the existence of indexing systems in hospitals A and F,
through which many of the false positives were traced. They result in a relative "inflation" of some of the figures.

It might be supposed that those units which exhibit the lowest rates of missed late presenting cases might be those which detected and treated the greatest number of hip instabilities in the early weeks. The data were therefore aggregated as nearly as possible into two equal groups, the first consisting of hospitals with the highest rates of severe CDH, the second group with the lowest rates. These aggregated results are given at the foot of table 4. The three "worst" (A,C,F) hospitals displayed a combined incidence of 0-52 "missed" cases per 1000 births while the four best had an incidence of 0-19 per 1000.

Paradoxically, the hospitals with the 3 worst results displayed the highest rates of early detection and early treatment. High awareness, high early detection rates, high early treatment rates, and prolonged early treatments were followed by a high incidence rather than a low incidence of late recognised dislocations requiring prolonged splinting or surgery.

This might have been explained by the reputed low incidence of CDH among immigrants and a concentration of births to immigrant families in particular hospitals. The data of table 4 were therefore re-assembled for Caucasian births alone (British, Irish, and other European), and the results are given in table 5. The three hospitals with the highest rates of late diagnosed CDH among all births (in table 4) had the three highest among Caucasians. The positive relation between high levels of early and late diagnosis was confirmed. The ethnic hypothesis is not the true explanation, although it is confirmed that the incidence is greater among Caucasians than among Asian and Afro-Caribbean immigrants.

**ONE DISEASE OR TWO?**

An alternative (partial) explanation of the concordant rates for early and late CDH could be the coexistence of two different diseases with different aetiologies. Hypothetically, the first occurs early and is transient while the second is either early or late and requires prolonged treatment. Better screening detects both of them in larger numbers.

Table 6 tests the two-disease hypothesis by comparing the 83 cases of CDH which required either surgery or prolonged splinting (column 1 of table 1) with all other cases. Since almost all cases of CDH occurred in Caucasians, the table provides similar figures for the whole Caucasian population of Birmingham. We found that the two "grades" of CDH differed jointly in several respects from the population at large, but they did not differ substantially from each other. They appear to be elements of a continuum.

The well recognised excess of breech presentations was evident in both sets; there was also an excess of caesarean deliveries. An excess of maternal x rays during pregnancy, again evident in both sets, once more reflects a raised frequency of pregnancy problems. An excess of female children was evident in both groups. The seasonal variation detected in the 1942–52 study—an excess from October to March—was not confirmed. The greatest frequency was between July and December.

The question of heterogeneity was pursued further in a study of laterality, and this interacted with mode of delivery. The ratio of unilateral to bilateral CDH
was 2.22 among infants with a vertex delivery but only 0.84 among those with breech vaginal deliveries or caesarean sections. The pattern was common to the two subclasses of CDH, and there was no distinction between them in this respect.

**Discussion**

(1) The incidence of CDH requiring prolonged treatment was no less in the Caucasian population in the period 1974–83 than it was before widespread neonatal screening was introduced.

(2) Late diagnosis was the rule rather than the exception, and the sensitivity of the procedure in the neonatal period was about 33%. Among “detections” under the age of 18 weeks the false positives outnumbered the true positives by about 10:1. The distinction between false positives and true positives...
Effectiveness of screening for congenital dislocation of the hip 289

was not always clear, but it would not be possible to suppose that many more of the positives were "true" without supposing that the incidence had risen substantially.

(3) Cases needing major treatment following a late diagnosis were more frequent in those maternity units which had high rates of early diagnosis. It has been suggested elsewhere that the test itself may be a cause of CDH; or perhaps the enthusiastic pursuit of early detection results in a number of unjustifiable diagnoses and unnecessary prolonged treatments. Early treatment might itself supposedly consolidate rather than avert the persistence of an early instability.

There was no evidence from the data that these paradoxical associations could be explained through concentrations of births to immigrant women in particular units, and there is no evidence that we are dealing with more than one condition. CDH appears to be part of a continuum extending from mild to major instabilities.

(4) Part of the aetiology must lie in the antenatal period and in the birth process, as evidenced by the increased incidence among both breech and caesarean deliveries and in the children of mothers who were x rayed.

There must also be a postnatal element, and the total absence of severe or late instabilities in infants born to women from the Indian subcontinent is striking. It is certainly not due to back-carriage and maintenance of the "frog" position in the infant, the explanation sometimes offered for the reputed low incidence in African populations. It is also noteworthy that three genuine cases occurred among children of Birmingham Afro-Carribean mothers, a group with a longer period of western acculturation than the Asian population.

(5) The service failures recorded here are no doubt due partly to lack of proper organisation, lack of quality control, lack of monitoring, and lack of feedback to those responsible for clinical procedures. There is no clear assignment of responsibilities for these functions or indeed for the performance of the service. However, the failures seemed to be especially associated with hospitals where the clinical service was well organised, and their causes might therefore be of a more fundamental nature.

(6) These results suggest that the natural history, the age at onset, and the nature of the postpartum determinants of dislocated hip have been misunderstood. Theories and treatments based upon the hypothesis that the persistence of an instability is determined by posture might not be correct. The persistence of an initial instability might be less a function of the position in which the hip is held (extension/flexion) and more a function of immobility of any kind as opposed to unrestrained mobility. Work from Japan, where the incidence is high, could be interpreted in this way. A detailed study of infant handling techniques by Asian mothers could pay dividends.

This work was carried out as part of a Health Services Research Programme supported by DHSS and by the West Midlands Regional Health Authority.

References


3 City of Birmingham Report of the Medical Officer of Health for the year 1967, 115.

4 DHSS. Screening for the detection of congenital dislocation of the hip in infants. Prepared by the Standing Medical Advisory Committee for the Central Health Services Council and the Secretary of State for Social Services, 1969.

5 DHSS. Screening for the detection of congenital dislocation of the hip. Prepared by the Standing Medical Advisory Committee and the Standing Nursing and Midwifery Advisory Committee for the Secretaries of State for Social Services and for Wales, 1986.


Accepted for publication July 1987
Effectiveness of screening for congenital dislocation of the hip.

E G Knox, E H Armstrong and R J Lancashire

*J Epidemiol Community Health* 1987 41: 283-289
doi: 10.1136/jech.41.4.283

Updated information and services can be found at:
http://jech.bmj.com/content/41/4/283

**Email alerting service**

These include:

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/