An Epidemiological Assessment of Neonatal Screening for Dislocation of the Hip

IAN LECK, MB, PhD, DSc, FRCP, FFCM
Professor of Community Medicine, University of Manchester

Screening for congenital dislocation of the hip was recently described in The Lancet as 'a mess'[1]. This article attempts to assess the deficiencies of screening by measuring it against the main criteria that any population-wide screening programme should meet. These criteria are that the disorder to be detected should be important and of known natural history and be best treated early, before the usual clinical time of presentation; the screening test should be accurate and safe, the treatment should be effective and its indications agreed; and both test and treatment should be practicable and worth the cost. In this article the term 'dislocation of the hip' (DH) will be used of cases in which the femoral head is either totally outside the acetabulum (luxation) or only in partial contact with it (subluxation); and the screening procedures considered will be those undertaken to detect neonates who, if untreated, would present clinically with DH on one or both sides when they began to walk.

Importance of DH

Importance depends on frequency, severity and duration. As to frequency, records made before the introduction of screening suggest that, without it, DH would become clinically apparent in early childhood in 0.8–1.6 per 1,000 children born in Scandinavia, North America and the UK (the regions where most of the available data on screening originated). Much higher figures have been reported for Amerindian, Japanese and Lapp communities whose methods of cradling or clothing infants tend to keep their hips extended and adducted (Fig. 1, left side). Severity is more difficult to quantify, but a DH that is severe enough to produce clinical signs in the infant or toddler will tend, if untreated, to lead to the formation of a false acetabulum, to degenerative changes, and to the knee joint problems and other consequences of limping caused by the affected leg being shorter than the other. The duration of these problems in the untreated is of course life-long.

The dislocated hip may be even more important than these facts would suggest, since the cases that present in early childhood may be only the tip of an iceberg. The number of children with hips in which the signs of subluxation, laxity or other dysplasia can be elicited at birth greatly exceeds one per 1,000, and although many of these hips may become normal later, it is widely assumed that an 'unknown number . . . persist as dysplasia or subluxation into adult life and ultimately become painful due to the development of secondary osteoarthritis'[12].

Fig. 1. Prevalence per 1,000 (logarithmic scale) of clinically diagnosed dislocation of the hip(s) in populations not studied neonatally, and of neonatal instability of hip(s) in screened populations. Screened hips have been assumed to exhibit instability if the published data were consistent with their being clinically dislocated or dislocatable. (Numbers show where sources appear in the list of references.)

- O = UK
- = Scandinavia (excluding predominantly Lapp series)
- = North America (excluding predominantly Amerindian series)
- = Japan
- = Other

Vol. 20 No. 1 January 1986
Natural History of DH

Epidemiological data suggest that DH has the natural history shown in Fig. 2[13]. Most of the risk factors fall into two groups, one associated with decreases in the resistance of the hip to dislocation and the other with external constraints. The risk factors affecting resistance include shallowness of the acetabulum and a generalised laxity of connective tissue (both of which are reported to be relatively common in children with DH and their relatives) and the female genotype (which may make joints liable to become lax under the influence of ovarian hormones). The risk factors that may owe their effects to external constraints include at least three in which these constraints are likely to be prenatal (oligohydramnios, primigravidity—when the uterus is ‘tighter’ than subsequently—and breech presentation), one which may operate before and/or after birth (birth around the coldest time of year, which is when both women in late pregnancy and newborn infants are most likely to be well wrapped up), and one with postnatal effects (membership of an ethnic group in which infants’ legs are kept extended and adducted; the prevalence of DH has fallen sharply in Japanese communities in which this practice has been abandoned[11]).

Most of the above risk factors that are likely to act prenatally appear to apply both to neonatal instability of the hips as detected by screening tests and to established DH in unscreened toddlers[14]. For this reason and because it seems inherently more likely that a hip which is dislocated or dislocatable at birth will remain abnormal than that a stable hip will become so, it is generally assumed that the unstable hip of the neonate and the dislocated hip of the toddler are early and late stages of the same process—although the fact that neonatal instability is almost always treated when found means that we have little direct evidence as to what happens if it is not.

Two points, however, are well established from studies of children who have been screened for dislocated and dislocatable hips at birth. The first is that the prevalence of these abnormalities in screened neonates tends to be much higher than the prevalence of recognised DH in toddlers who were not screened at birth (Fig. 1). Second, most of the reports of screened-populations (Fig. 1) make reference to children in whom hip abnormalities were not detected and treated as a result of screening but who were later diagnosed as having DH. In Fig. 3 the proportions of children in this category are plotted against the proportions in the same series in whom screening led to treatment. Like an earlier analysis reported by Parkin[51], Fig. 3 reveals no association between the frequency with which neonatal instability is treated and the prevalence of cases presenting late—suggesting that screening programmes with a high yield of cases of instability have in general had no more success than those with a low yield in picking up infants at risk of lasting dislocation. However, the prevalence of late cases in nearly three-quarters of the screening studies was below the lower limit of the range for unscreened populations, even though the figures for late cases in some of these studies included children whose hips only appeared unstable a few months after birth and might never have progressed to established DH in early childhood.

This evidence that treating neonatal instability prevents some cases of DH in early childhood provides further support for the view that these two conditions are early and late stages of the same process. However, the higher prevalence of neonatal instability in the screened than of childhood DH in the unscreened indicates that most cases of the former do not progress to the latter if left untreated. Presumably the cases most likely to progress are the most unstable and those in which there is most interference with hip flexion and abduction during infancy.

There is probably more than one explanation for the finding that hips that appear stable in the neonate

---

Fig. 2. Suggested natural history of ‘congenital’ hip dislocation.
sometimes manifest DH later. Sometimes abnormalities are present at birth which standard screening tests do not detect—notably irreducible dislocation—or which are missed because the observer is inexperienced; but there are also well-documented cases on record in which hips that an experienced examiner had found to be clinically stable at birth showed increasingly obvious dysplasia on radiological follow-up and eventually became dislocated.[52]

In picturing the natural history of DH, we should therefore envisage some cases as developing de novo after birth, perhaps as a result of postnatal constraints interacting with a genetic predisposition in the way portrayed by the interrupted line in Fig. 2. One component of the genetic predisposition in these cases may be acetabular shallowness, to which the parents of children with established DH seem to be more predisposed than parents whose children’s hips were unstable at birth.[53]

It is not possible to estimate how many cases of DH involve hips that were stable at birth, although the data shown in Fig. 3 (which are confined to populations of European origin) could be interpreted as suggesting that careful neonatal screening can detect instability in a substantial majority of the 0.8-1.6/1000 infants in such populations who would develop DH if unscreened.

In the higher-risk communities (e.g. in Japan) in which infants’ legs are kept extended and adducted, we should expect screening to be less sensitive if such postnatal constraints can indeed lead to DH in initially stable hips. This expectation is borne out by a report that, in a series of Japanese infants of whom 2.7 per cent had DH according to a one-year follow-up, less than one-fifth of this 2.7 per cent had shown signs of instability one week after birth[27].

Best Time for Treatment of DH

The question whether DH is best treated earlier than the time at which unscreened cases present ignores the neonates whose hips are unstable at birth but would stabilise later even if not treated, and focuses our attention on those who, without treatment, would present with clinical DH as toddlers. If splinting in abduction, the recommended early treatment, is effective, they must clearly stand to benefit by receiving this treatment, partly because otherwise their hips will become more abnormal, and also because neonatal treatment is less radical than the alternative later treatment, probably carries less risk of ischaemic necrosis of the femoral head (the main hazard of later treatment)[51], and comes at a time when it is likely to have less impact on the physical and psychosocial development of the child and on family life.

Accuracy of Neonatal Screening for DH

Although radiographic and clinical methods have both been advocated for neonatal screening, radiography cannot really be regarded as suitable for general use, given its hazards and the fact that the head of the femur and much of the pelvis are not yet ossified. The most widely used clinical procedure has two parts. The first is the Ortolani test of abducting the flexed thigh, which is designed to enable the already dislocated hip to be detected by causing the femoral head to slip into the acetabulum, when a 'clunk' should be felt or heard. The second part, introduced by Barlow to identify unstable hips which the Ortolani test misses because they are not already dislocated, is to see whether dislocation followed by reduction can be detected if one presses first on the lesser trochanter and then on the greater while holding the thigh in 90° of flexion and 45° of abduction.

It is debatable how accurately these signs pick out unstable hips. Some workers[49,54] have argued that the other clicks and grating sensations which these manoeuvres sometimes cause should also be accepted as evidence of instability, although this view is not generally held[1,55]. However, it seems beyond debate that these tests for dislocation and dislocatability are far from accurate in identifying future cases of established DH, given that the latter are far less common than are positive tests (Fig. 1) and sometimes arise in infants in whom the tests were negative (Fig. 3). The figures which these results yield for the standard indices of accuracy—sensitivity, specificity, and predictive value of a positive test—are of course very variable, because of the wide variations in the frequency both of positive tests and of 'false negatives' (negative tests in infants who are later found to have DH). Table 1 shows estimates of accuracy for what may be regarded as the best available recent English data, based on prolonged follow-up of a large series of Bristol infants screened by a teaching unit with particular exper-
Table 1. Indices of accuracy yielded by a series of 23,002 screened Bristol children[48], given the assumption that DH occurs in 1.25/1,000 unscreened children (midpoint of range for non-Lapp populations of European origin in Fig. 1).

| Screening test result | Positive | Negative | Total |
|-----------------------|----------|----------|-------|
| Prognosis if not treated: |
| Dislocation            | 19       | 10       | 29    |
| No dislocation         | 426      | 22,547   | 22,973|
| Total                  | 445      | 22,557   | 23,002|
| Sensitivity            | 19/29 = 66% |
| Specificity            | 22,547/22,973 = 98% |
| Odds against developing established dislocation, given a positive test result, = 19:426 = 1:22.

In this case[48,55]. The most disappointing of these figures is the estimate that even in such a favourable setting the test yields as many as 22 false positive results for every one true positive. The reason for this figure being so large when sensitivity and specificity are as high as 66 per cent and 98 per cent respectively is the relatively low prevalence of DH; with a test of this accuracy the number of true positive results will only exceed the number of false positives for conditions affecting over 2.9 per cent of subjects.

Safety of Neonatal Screening for DH

It has been suggested that one of the manipulations involved in screening—the manoeuvre described by Barlow for seeing whether the hip can be dislocated—may itself render some hip joints less stable[42]. Although this hypothesis has received very little attention, it should perhaps be borne in mind as a possible reason why there have been screened series in which the reported prevalence of late-diagnosed cases of DH (Fig. 2), and even the proportions of all infants undergoing surgical operations for DH[9,30,49], exceeded the overall prevalence of DH in unscreened populations. A rigorous test of the hypothesis that screening (perhaps when inexpertly performed) can be harmful would be to screen but not treat a group of neonates so that they and an unscreened group could be compared for the subsequent prevalence of DH; but this would raise ethical problems.

Indications for Treating Neonatal Hip Joint Instability

The clinician who accepts both the possible benefits of early treatment and the limitations of early screening for DH faces a major problem. Should every neonate in whom the results of screening appear positive be immediately subjected to some form of splinting in abduction, even though most of these neonates would not develop DH even if left untreated; or should the numbers treated unnecessarily be reduced by restricting splinting to those whose hips are still abnormal at some later date, even though the latter children may by this time be harder to treat?

Not surprisingly, there is no agreed answer to this question. In 1981, Parkin[51] pointed out that different authorities had recommended policies of (a) immediate treatment of all cases with evidence of instability; (b) a few weeks' observation of all cases followed by treatment of those remaining abnormal; and (c) a combined approach whereby those in whom screening provides evidence of spontaneous dislocation are treated immediately while those whose hips only appear dislocatable are observed for a few weeks and then treated if the abnormality persists. More recent papers indicate that the first two of these policies still have their advocates among experienced British workers[9,48].

Disagreement about the indications for treatment is also implicit in the above-quoted debate about the clinical signs that should be treated as evidence of instability. Even among the minority who accept as evidence not only signs that the hip is or can be dislocated but also other clicks and grating sensations, there are some who splint hips on the basis of such evidence alone[54] while others do no more initially in cases with the latter signs only than follow them up, splinting at a later stage those that then prove to be dislocated, dislocatable, or radiologically abnormal[49].

Effectiveness of Treatment of Neonatal Hip Joint Instability

The treatment of neonatal instability of the hip joint has been advocated as a measure for preventing osteoarthrosis[12] as well as established dislocation of the hip. It is unlikely that we shall be able to evaluate its effectiveness against osteoarthrosis, much the commoner condition, until data based on follow-up of screened and unscreened populations for several decades become available. Even the success of neonatal treatment in preventing established dislocation of the hip cannot be fully evaluated, since there appear to have been no randomised clinical trials. The most that published data enable one to do (and this applies in a small number of studies only) is to examine the overall frequency of significant residual abnormalities in screened populations, including cases both in children given early treatment for instability and in other children.

Table 2 shows the proportions of children in these studies whose residual abnormalities were apparently considered significant enough to be submitted to surgical operations. As will be clear from the final column of the table, it is questionable whether each figure relates to precisely the same surgical procedures: the procedures included in the Aberdeen statistics for operative treatment were not specified, and the papers from Edinburgh and Uppsala made no mention of adductor tenotomies. The most informative sets of results appear to be the two from England. Neither of these suggests that neonatal screening and treatment are highly effective: in the Mansfield series the proportion requiring surgery was...
close to what might have been expected in the absence of screening, and even the proportion for Bristol (the centre of expertise from which the figures in Table 1 also came) only seems to have been halved by the screening programme. However, these findings may reflect more discredit on the screening procedure than on the treatment: in both the English series and in three of the others a substantial majority of the operations were carried out in children who had not been selected for splinting when originally screened.

The different types of abduction splints have not been compared here because epidemiological data on their relative merits are lacking. However, the keeping of infants with unstable hips in double nappies, which has been advocated as an alternative to the use of special splints, seems to emerge as particularly ineffective from a recent report; when splinting was replaced by double nappies in Southampton, the percentage of cases responding to treatment among those diagnosed by screening fell from 92 to 55[56].

### Table 2. Frequency of surgery for DH in screened populations.

| Place                      | Population size | Proportion undergoing surgery (per 1000 total population) | Nature of surgery (No.) |
|----------------------------|-----------------|----------------------------------------------------------|-------------------------|
|                            |                 | Splinted as neonate | Not splinted as neonate | Total (No. in brackets) |
| Uppsala, Sweden[23]        | 11,868          | 0.25               | 0.42                  | 0.67(8)                 | Open reduction (2), closed reduction (6) |
| Aberdeen, Scotland[28]     | 71,169*         | 0.11               | 1.10                  | 1.21(86)                | ‘In-patient operative treatment’ |
| Edinburgh, Scotland[29]    | 31,961          | 0.22               | 0.13                  | 0.35(11)                | Surgical reduction (9), excision of limbus only (1), derotation osteotomy (1) |
| Aberdeen, Scotland[9]      | 53,033          | 0.19               | 1.64                  | 1.83(97)                | ‘Operated on’ |
| Bristol, England[48]       | 23,002          | 0.22               | 0.43                  | 0.65(15)                | ‘Open surgery’ (7), adductor tenotomy (8) |
| Mansfield, England[49]     | 7,864           | 0.25               | 1.14                  | 1.40(11)                | Adductor tenotomy (+ open reduction in 1) |

*Estimated number remaining through follow-up period.

### Practicability of Providing Screening and Treatment for Neonatal Hip Joint Instability

For a screening programme to be practicable, it must have both the ability to reach those for whom it is intended and the resources to screen and, where necessary, treat them. These conditions are readily attainable by screening for DH in countries such as Britain where almost all births occur in hospitals, so that the neonates to be screened form a captive population, and staff are available who can be trained to carry out the screening.

### Cost of Providing Screening and Treatment for Neonatal Hip Joint Instability

Estimates of the current financial costs of screening and splinting for neonatal instability and of surgical treatment for established DH in the UK are not readily available, but figures have been published for British Columbia in 1981-82[57] which can be used to make an approximate

### Table 3. Health care costs of screening and not screening a population of 100,000 with a distribution like that estimated for the Bristol population in Tables 1 and 2, given unit costs based on British Columbia experience in 1981-2[57].

| Type of care                        | Number of recipients | Cost (Canadian $) |
|-------------------------------------|----------------------|-------------------|
|                                     |                      | for each recipient| for total population|
| Requirements if screening is omitted|                      |                   |
| Surgical treatment                  | 125                  | 9,133             | 1,141,625            |
| Requirements if screening is practised|                     |                   |
| Screening                           | 100,000              | 1.90              | 190,000              |
| Splinting                           | 1,935                | 167               | 323,145              |
| Surgical treatment                  | 65                   | 9,133             | 593,645              |
| Ratio of costs with screening to costs without screening = | $rac{1,106,790}{1,141,625}$ = 0.97:1 |

60 Journal of the Royal College of Physicians of London Vol. 20 No. 1 January 1986
assessment of the relative health service costs of screening and not screening. When these figures are applied to the Bristol-based statistics in Tables 1 and 2, screening emerges as marginally more economical (Table 3).

More than half the estimated health care costs associated with screening are attributable to surgical treatment of children in whom early splinting either failed or was not carried out. However, the British Columbia workers assumed that the cost of surgical treatment per case for these groups was the same as for unscreened children with DH. If, as is, possible, the former tended to require less treatment than the latter, or if the accuracy of screening or the effectiveness of splinting in early infancy were to improve, the financial advantages of screening could increase considerably. The converse is, of course, equally true.

To try to quantify and compare the costs for infants and their families which policies of screening and not screening incur would involve many other aspects of life besides money. It could be that the two policies are as finely balanced in respect of these costs as in respect of the costs of health care; the main difference between the two policies in both cases may be that, with screening as compared to without, the costs are shared between many more individuals.

Conclusions

There seem to be four main problem areas in the field of neonatal screening and treatment for hip joint instability as currently practised.  

1. False Positives. Most of the screened cases in which unstable hips are reported are ‘false positives’, at least in the sense that they would not present later with established DH if untreated. Although it is widely believed that these false positives have a predisposition to osteoarthrosis, which may be reduced by treating the neonatal instability, this hypothesis has not been tested epidemiologically. It is therefore at least possible that several infants—more than ten in many areas—suffer unnecessary splinting for every one who benefits from this procedure.

2. False Negatives. The prevalence of ‘false negatives’ (cases in which a hip that appeared stable when screened is later reported to be dislocated) in screened populations varies widely, but averages about half as much as the prevalence of DH in unscreened populations. In other words, the number of false negative results yielded by screening tests seems to be of the same order as the number of true positives. It is not known how often in these false negative cases instability is present but missed at birth and how often it develops later.

3. Treatment Policies. There are at least three questions about treatment on which experts differ. First, should the indications for treatment include abnormal physical signs short of those accepted as manifestations of dislocation and dislocatability? Second, should treatment be started immediately in all in whom the relevant signs (however defined) are observed, or postponed for a few weeks and confined to those in whom signs persist? Third, what type of splint should be used?

4. Outcome of Early Treatment. Although there are some screened populations in which established DH is substantially less common than in the unscreened, there are others in which it is so common as to raise the possibility that, at least in some hands, one or more of the procedures involved in screening may actually increase prevalence. The example of Bristol shows that, even at a centre of expertise, screening may do no more than halve the frequency of established DH and have negligible effects on health care costs.

It should be noted that there are unanswered questions in all these four areas. Some of the more important of these questions—whether unstable hips other than those which would progress to established DH benefit from splinting in early infancy, and whether screening procedures can themselves predispose to DH—may be extremely hard to answer, the former because very long follow-up would be needed and the latter because of the ethical problem mentioned when the safety of screening was discussed. However, it should be possible to set up randomised clinical trials which would answer ethically and relatively quickly the three questions listed under ‘treatment policies’. These trials should be designed to enable the costs and benefits of current neonatal screening and treatment policies to be compared as comprehensively as possible. The costs and benefits of the best of these policies should then be compared with those of the policies of screening later in infancy and of not screening but treating DH when it presents clinically. Whether screening is justified will remain an open question until this is done.

This article is based on a paper read at the Screening Conference held at the Royal College of Physicians in December 1984.

References

1. Robертон, N. R. C. (1984) Lancet, 1, 909.
2. Getz, B. (1955) Acta Orthopaedica Scandinavica, Supplement 22 (misprinted as 18).
3. Harris, L. E., Lipscomb, P. R. and Hodgson, J. R. (1960) Journal of the American Medical Association, 173, 229.
4. Palmén, K. (1961) Acta Paediatrica Scandinavica, 50, Supplement 129.
5. Rabin, D. L., Barnett, C. R., Arnold, W. D., Freiberger, R. H. and Brooks, G. (1965) American Journal of Public Health, 55, No. 2, Supplement.
6. Leck, I., Record, R. G., McKeown, T. and Edwards, J. H. (1968) Teratology, 1, 263.
7. Walker, J. M. (1977) Canadian Medical Association Journal, 116, 501.
8. Lehmann, E. C. H. and Street, D. G. (1981) Canadian Medical Association Journal, 124, 1003.
9. Mackenzie, I. G. and Wilson, J. G. (1981) Journal of Bone and Joint Surgery, 63-B, 38.
10. Paterson, D. (1982) Congenital Dislocation of the Hip, p.145. (ed. M. O. Tachdijian.) New York: Churchill Livingstone.
11. Yamamura, T. and Ishida, K. (1984) Clinical Orthopaedics and Related Research, 184, 34.
12. Coleman, S. S. (1982) Congenital Dislocation of the Hip, p.181. (ed. M. O. Tachdijian.) New York: Churchill Livingstone.
13. Leck, I. (1983) Obstetrical Epidemiology, p.263. (ed. S. L. Barron and A. M. Thomson.) London: Academic Press.
14. Leck, I. (1977) Handbook of Teratology, Vol. 3 (Comparative, Maternal and Epidemiologic Aspects), p.243. (ed. J. G. Wilson and F. C. Fraser.) New York: Plenum.
15. Coleman, S. S. (1956) Journal of the American Medical Association, 162, 548.
The Value of Preventive Medicine, edited by D. Evered and Julie Whelan. Ciba Foundation Symposium 110. Pitman Medical, London, 1985. 258 pages. Price £27.95.

This is a rather unusual topic for a Ciba Foundation Symposium dealing not, as in most of the series, with the details of a fairly circumscribed subject but with the many broad considerations, social and economic as well as medical, that need to be considered in prevention. Contributors and participants came from the UK, USA, Belgium, Switzerland, Bulgaria, Sweden, Norway, Australia, Japan and Canada. The objectives of preventive medicine are considered by Sir Richard Doll. It may not be very long before the great majority of deaths will occur at 80 to 85 years of age but a general increase in longevity beyond 90 or 100 years seems unlikely. Improvements that are possible depend on a substantial reduction in mortality from three groups of diseases that are now the main causes of death under the age of 85—neoplasms, ischaemic heart disease and other vascular disease. Whether a reduction in age-specific mortality will also result in less disability is unclear. David Weatherall considers the impact of new methods of gene analysis on screening for genetic disease. Beneficial results that would follow from the control of tobacco-related disease are dealt with by Richard Peto. The controversial issue of the preventability or otherwise of coronary heart disease is discussed at length. Between them, several contributors review the major primary prevention trials in some detail. Much of the discussion in the chapters concerned deals with the difficulties of designing and interpreting such trials and, being a good deal less formal than the way in which these points are usually handled, it is of special value. Is the randomised controlled trial the pinnacle of scientific evidence? Do preventive trials test aetiological hypotheses or evaluate policies—or both? These are some of the questions raised by Geoffrey Rose. Screening for cancer, screening and intervention in alcohol-related disease and the value of physical fitness are further medical or physiological topics included.

Other contributors outline the economic aspects that have to be borne in mind. Particularly useful in what can very easily become a confusing subject is the chapter by Jeffrey Koplan on the benefits, risk and costs of immunisation programmes. This reviews studies on a number of particular vaccines such as those against poliomyelitis.