

Oberklaid, F., Wake, M., Harris, C., Hesketh, K. and Wright, M. (2002). **Child Health Screening and Surveillance: A Critical Review of the Evidence**. Canberra, ACT: National Health and Medical Research Council.

[www.nhmrc.gov.au/publications/pdf/ch42.pdf](http://www.nhmrc.gov.au/publications/pdf/ch42.pdf)

## **EXECUTIVE SUMMARY**

The early detection of health and other problems in children is a worthy goal. At first glance the benefits appear to be self evident. It is not until one begins to systematically review the evidence for screening and surveillance, topic by topic, that the complexity of this endeavour becomes apparent.

There is surprisingly little evidence for the effectiveness of screening programs in many domains (that is not to say that the corollary is true – for some, there is inadequate evidence for lack of effectiveness either). There are scant data about cost effectiveness. There are major issues of program quality, monitoring of compliance with referrals for assessment, and whether facilities exist in many communities for assessment and follow up. There are concerns that much attention is paid to the test or procedure itself and little to the main elements of a community-wide program. In some cases, there is little evidence that therapy alters outcomes. These issues are outlined in relevant chapters on topics. Finally there are problems with terminology, and the way in which terms such as “ screening” and “ surveillance” are used.

This report should not be construed as diminishing the importance of attempts at early detection. Rather it provides evidence that perhaps we need to rethink how best we do this; a discussion of these issues is to be found in Chapter 9.

A distillation of the evidence, recommendations and research agenda for each topic is summarised in this section; details are to be found in the body of work for each topic. The topics are listed alphabetically for convenience.

### **1.1 TERMS OF REFERENCE**

The NHMRC last undertook to review child health screening and surveillance activities in 1993. Since the last review there has been an emerging acknowledgment of the importance of evidence based medicine. In 1999 the NHMRC instigated an update of the 1993 review employing evidence based techniques. Initially this work was to be carried out as a single task of critically reviewing the evidence and developing clinical guidelines. Prior to commencement of the review, NHMRC decided to divide the task into a two stage process: Stage 1, consisting of a critical review of the evidence for screening and surveillance activities around fourteen child health topics, followed by Stage 2, consisting of the development of evidence based clinical guidelines for practitioners informed by Stage 1. This report constitutes Stage 1 of the process. Following consultation with the NHMRC-appointed Steering Committee and the contractors, the list of fourteen child health topics to be included in the review was revised to twenty-one. However one of the reviews (post-natal depression) was ultimately excluded from the final document, as a more comprehensive review of this subject had recently been released by NHMRC.

The guidelines set out by NHMRC to conduct the review included:

- identification of relevant information through searches of peer-reviewed literature published since the release of the last review (post-1993), consultation with groups undertaking similar work both within Australia and internationally, and consideration of submissions received by NHMRC during two public consultations.
- critically evaluate the scientific literature and other information for quality and strength of evidence using the ratings outlines in the NHMRC publication “A Guide to the Development, Implementation and Evaluation of Clinical Practice Guidelines”. This requirement was modified to allow rating of non-clinical trial evidence which is not comprehensively covered by the NHMRC publication.

It was decided by the NHMRC-appointed Steering Committee that the review should be limited to the consideration of children from birth to eighteen years of age, and thus would not cover prenatal screening activities.

The findings within this report represent those of the authors of this report, based on their critical review of the literature up to August 2000 and informed by comment from the project’s Reference Group and a limited number of experts in each field. This report does not necessarily represent the views of the NHMRC. Because clinical guidelines were not developed, this review has not been subject to the wide external consultation usually associated with guideline development.

## **1.2 METHODOLOGY**

An expert Reference Group was established (Section 5.1). This group provided advice and feedback to the project, from its inception to its completion. In addition, individual members were asked to advise on specific topics. For each topic, experts in that topic area were also consulted to ensure that information in the report was factually correct, and that key works in that area had not been missed.

In consultation with experts in searching the electronic literature, search strategies were developed to locate studies published in 1993-2000 (ie since the previous NHMRC review) according to quality of evidence ratings, as defined in the NHMRC publication “Guidelines for the Development and Implementation of Clinical Practice Guidelines”. Search strategies are included in Appendix B. A limited review of the pre-1993 peer-reviewed literature was undertaken. Other data sources were also searched (Section 5.2).

Final selection of topics to be reviewed, search strategies, specified frameworks to assess the quality of different aspects of the literature (Sections 5.3 and 5.4), terminology (Chapter 3), and the proposed structure for the topic reviews were agreed between the Steering Group and the project team prior to commencement of individual topic reviews.

The draft of the final review was scrutinised by both the Steering and the Reference Groups and modifications made in light of their comments. This also led to the addition of Chapter 9, “The language of early detection: time for review?”.

### 1.3 FINDINGS: OVERVIEW

Relatively few topics could be recommended for formal screening programs.

For many conditions we could not recommend formal screening programs, for a variety of reasons including:

- Multi-dimensional conditions on a continuum of normality-abnormality do not lend themselves to pass/fail criteria (see Chapter 9)
- available screening tests not considered sufficiently acceptable to the target population, based on reported uptake rates of either the screening test or definitive referral
- sensitivity could not readily be balanced against specificity – ie very large numbers of false positives a by-product of capturing all or most of those with the target condition
- the target condition itself was too variable over time to justify screening at a single time point, but evidence to support periodic screening (surveillance) not available.
- management for those detected by screening not shown to significantly alter outcomes
- not yet an agreed therapy.

Some conditions seemed important and management likely to alter outcomes, but there was inadequate research evidence to support this. For these conditions, further research seems vitally important.

For some conditions 'systems issues' seemed especially important. A good example is clinical examination in the newborn period. Early discharge (<72 hours post partum) is now commonplace in Australia, thus there is potential for screening to be missed or incomplete if the infant is discharged from hospital within a few hours of birth. It is important that there is good communication between hospitals and the community, that clinicians in both settings are knowledgeable and competent in the management of neonates, and that a line of responsibility is clear. The findings of examinations undertaken in the maternity unit and any follow up required should be clearly communicated to health professionals in the community. Systems are required to ensure that:

- a point of accountability exists for each child's neonatal examination
- a complete neonatal examination has been undertaken
- the findings of the examination and any subsequent action, referrals, etc have been documented
- correspondence between health professionals in maternity settings and those in the community (particularly GPs and MCHNs) is adequate and appropriate.

Similar systems issues need to be considered for other conditions with respect to early identification.

For some conditions, the balance of harm vs benefit may change over time as new therapies or detection techniques become available, and thus suitability for screening will need to be reviewed. For example, prenatal screening may lower the yield of newborn screening for some conditions, while effective therapy for overweight might make screening or surveillance appropriate in the future.

## **9. THE LANGUAGE OF EARLY DETECTION: TIME FOR REVIEW?**

It has long been accepted that efforts at prevention and early detection form a core component of health services. For some conditions it is possible to prevent any manifestations of what might otherwise be a serious or fatal condition. For others, early detection and early treatment can alter the natural course of the condition to prolong life expectancy and/or improve the quality of life. Early detection has also been considered to be desirable for conditions related to development and behaviour. There is a growing body of evidence documenting the effectiveness of early intervention for children at biological and/or environmental risk of poor or sub optimal developmental outcomes (Shonkoff & Hauser-Cram 1987).

In recent years, three loosely-related terms have commonly been used to describe these activities: screening, surveillance, and health promotion. In this chapter, we review the current status of these terms, and then make some suggestions to better focus these terms.

### **9.1. CURRENT TERMINOLOGY**

#### **9.1.1. Screening**

Screening tests have been used in clinical practice and in population programs for many years, and the number of conditions screened for gradually expanded. A screening test is not intended to be diagnostic, but sorts out those who are likely to have the condition from those who likely do not.

As can be seen in the body of this report, there is clear evidence for the effectiveness of screening for a number of conditions, many of which are readily detectable in the neonatal period. There are well-established criteria against which to benchmark screening tests, suitability of conditions for screening, and overall effectiveness of screening programs – these are outlined in Chapter 4. Implicit in these criteria is that one either “has” or “does not have” a disease or condition, and that if one has, it can and will be treated effectively. This assumption underlies several of the most fundamental properties of screening programs, such as sensitivity, specificity, positive and negative predictive values, and number needed to treat. Screening programs that work well target conditions in which this distinction can be readily made for the great majority of children (eg PKU, congenital hypothyroidism, congenital adrenal hyperplasia). But what of conditions for which early detection and intervention may seem a useful goal, but for which no such distinction can readily be made - ie where there is considerable difficulty into categorising into 'has' or 'does not have' the condition?

Typical of this are the diverse problems that may occur in child development (including language) and behaviour. Screening for these problems is discussed in detail in Chapter 6.3.18. The terms “development” and “behaviour” artificially collapse constructs which are multidimensional along numerous axes. The concept of “having” or “not having” a problem in any one of these domains is mostly artificial and therefore problematic. Nonetheless, often driven by requirements of service systems to determine eligibility, generations of professionals have attempted to make such distinctions, and in turn to develop screening tests that are “valid” in terms of

sensitivity, specificity, and positive predictive value. The extent to which this has failed can be appreciated in the commonly-held view that, for developmental screening tests, sensitivity of 70% and specificity of 70–80% is “acceptable” - values which would not be admissible when considering introduction of a new biochemical screen for a metabolic condition. Yet most developmental screening tests fail to reach even this low level. We are even less well off when it comes to distinguishing diagnosable behavioural problems from the normal and expected developmental problem behaviours. We should not blame the screening tests for this. While one can tinker with sensitivity at the expense of specificity, and vice versa, this begs the question: perhaps child development and behaviour are areas that are fundamentally unsuitable for screening, and an alternative approach is needed.

### **9.1.2. Surveillance**

Given the issues and problems described above, it has been suggested that for child development and behaviour the concept of screening be replaced by surveillance (Hall 1996; Dworkin 1989).

There is no consensus on how surveillance is defined. For many it has a narrow definition, focusing on early detection of problems. For others it has been expanded to embrace many other activities. Hall noted that existing conceptualisations of surveillance could include “the oversight of the physical, social, and emotional health of children; measurement and recording of physical growth; monitoring of developmental progress; offering and arranging intervention when necessary; prevention of disease by immunisation and other means; and health education” (Hall 1996). However, he noted this concept to be problematic in its breadth.

Dworkin defined surveillance as “a flexible, continuous process that is broader in scope than screening, whereby knowledgeable professionals perform skilled observations on children throughout all encounters during child health care (Dworkin 1989)”. While screening consists of a professional administering a test to a child, often as an activity divorced from routine care and excluding parental input, surveillance in this context actively involves parents by eliciting and responding to their concerns, putting these concerns into a historical and family context, and providing advice and information. Standard screening tests could be used as a component of surveillance, but were not the central focus (Glascoe & Dworkin 1993). Surveillance could be employed with individual children, but “also in the process of monitoring the health of a whole community of children” (Hall 1996).

By acknowledging and potentially addressing the biological and environmental factors that can influence child health and development, surveillance may seem to overcome many of the limitations of screening. However, its very breadth poses the danger of making both the term and the process vague and meaningless. Further, through multiple encounters rather than just one, its fundamental goal may still be similar to that of screening - to identify children who do or do not have developmental delay or a behaviour problem. What activities are included and what excluded? Do we include any or all activities that might conceivably improve children’s health? Should we try to develop a more precise scientific definition with specified, clearly definable activities? How well do efforts at early detection sit with prevention activities and health education?

The term surveillance may have another different connotation for consumers. There is anecdotal information, difficult to quantify, that some parents perceive that the term surveillance somehow involves professionals evaluating and judging their competence as parents, or checking up on them. Such perceptions, even if wrongly conceived, nevertheless are of concern; if widely held, they may undermine population programs of surveillance in the epidemiological sense.

### **9.1.3. Activities designed to prevent problems and promote health**

There are a number of activities, often subsumed under the umbrella of health surveillance, that are designed to either to prevent problems from occurring or to promote or enhance health outcomes. These include making sure that children are fully immunised, ensuring a safe environment to prevent injuries, reading to young children to promote early literacy, encouraging breastfeeding and subsequent good nutrition, discouraging parental smoking, anticipatory guidance, and providing parents with accurate information on the basis that well informed parents will be more likely to make appropriate decisions in relation to their children.

Some of these activities have a strong evidence base (eg immunisation, breast feeding); for others there are some limited data suggesting they are likely to be effective in facilitating improved outcomes (eg early literacy, injury prevention programs); some appear to be intuitive yet at present there is no compelling evidence as to their effectiveness in improving outcomes (anticipatory guidance, the provision of information to parents).

As for child health surveillance, there appears to be little consensus as to which preventive and health promotion activities should be systematically incorporated into child health programs and how their outcomes should be judged. To date, with the notable exception of immunisation, such activities have tended to be ad hoc and time-limited, and to occur outside a quality framework.

## **9.2. MOVE AWAY FROM THE CONCEPT OF PASS/FAIL**

For many conditions in childhood where early support is likely to be of benefit – development, language, behaviour, family psychosocial issues – their very nature is such that there will never be a suitable screening test. These are complex, multi-dimensional areas that are not appropriate to categorise into pass/fail on the basis of a test; such a mutually exclusive categorisation flies across the face of the nature of children's development and of family functioning.

Development, language, behaviour, attention, family stress, quality of parenting and other variables exist on a continuum. At one end of the spectrum are those children and families where deficit, delay or dysfunction clearly exists (eg Down Syndrome or other chromosomal disorders), and where it is likely that this will be recognised by parents or professionals in the course of routine contact with a high-quality health and education system. At the other end of the spectrum are well-functioning children in well-functioning families without biological or environmental risk factors, and where good outcomes are likely without any specific intervention. Neither of these groups

are likely to need a specific screening or identification program, though the former will certainly need appropriate assessment and management systems and facilities.

Many children and families exist toward the lower end of this continuum. They are not in the clearly abnormal range for any condition, yet various poor outcomes are frequent. Children in these groups, variously described as “borderline”, “low-functioning”, “high risk” or in the “gray” or “intermediate” zone, may have very real additional needs but can be virtually invisible to screening tests/programs that by definition recognise only normal or abnormal. Yet by virtue of their position on the normal curve they are usually more numerous than those who have a clearly-defined disorder. While risk and protective factors may be fluid, some will have a combination of biological and environmental risk which is relatively stable and potentially leads to poorer outcomes – the concept of “double jeopardy” (Escalona 1982; Parker, Greer & Zuckerman 1988).

Many of these children and families could benefit from some sort of support or intervention. For example, individual maturational differences are the most likely reason that a two year old might have limited language. If there is no intervention, two thirds of these children will have language which is considered “normal” or “within normal limits” by age four, although half of these will continue to carry subtle evidence of their earlier difficulties. Early language support might benefit the child who would otherwise go on to have true language impairment. It might also benefit the children who “recover”, not only in terms of ultimate language functioning but also in decreasing the chances of subsequent learning difficulties. Similar scenarios can be constructed for other domains such as behaviour, where a child may not fit into a diagnostic category but where intervention at a young age may be beneficial in terms of improved short and long-term outcomes. Additional support may also allow additional scrutiny of progress, and may facilitate earlier detection of and intervention for a “true” condition or disability.

### **9.3. CLARIFICATION OF TERMINOLOGY**

We would suggest greater precision in the way these terms are used.

#### **9.3.1. Screening**

“Screening” should be confined to those conditions and tests that meet criteria for screening and for which the benefits of screening have been clearly demonstrated. It should be noted that most of these conditions and tests occur in the neonatal period, and are likely to be implemented exclusively by health professionals in health care settings. These are situations where pass/fail criteria are appropriate, where it is almost always clear what further action needs to be taken on the basis of a failed screening test, and where management instituted as a result of screening is expected to substantially alter outcomes for that child.

Even though there is strong evidence for the effectiveness of screening with these conditions, it is still important to establish quality processes for the tests themselves, their interpretation, the expertise of the testers, and the program as a whole including systems and standards for referral, tracking, management, support and reporting. Many clinical conditions currently screened for do not approach ideal system

standards (eg congenital dislocation of the hip, undescended testes, ophthalmic screening).

### **9.3.2. Surveillance**

We would argue for a more restrictive definition of surveillance, linked to its original intention of early detection of specific problems.

Surveillance should refer to a set of activities designed for the early detection of clearly-defined and specific problems which, unlike those appropriate for one-off screening, would not be expected to be reliably detected at a single point because they may develop or fluctuate over time. Essentially, surveillance would equate to periodic screening over time; “failing” the screen one or more times would indicate a high likelihood of having a disorder. Unlike adult periodic screening programs, such as for breast and cervical cancer, the instruments used might vary at different ages to be appropriate to the child’s age and/or developmental stage.

Such conditions might include short stature, overweight/obesity, hypertension, and vision and hearing problems in the school years, if the evidence were there to support these activities. We would also argue that, like screening, surveillance activities should adhere to the evidence-based principle that each such activity should lead to more benefit than harm. Generally research in this area of child health is almost non-existent. Thus, few of these activities can currently be recommended on the existing evidence; this does not mean that detection of these problems is not important.

### **9.3.3. Activities designed to prevent problems and improve or promote health**

These activities should include a broad range of issues, for which different approaches may be needed. However, these activities are likely to be implemented in a range of settings by many of the professionals who come into contact with children and their families, not just in the health care system. They may be divided into universal and targeted activities.

**Universal activities**, directed at all children, may:

- promote healthy lifestyles, for example food choices, physical activity, safety in the home, and immunisation.
- directly enhance health & well-being, for example promotion of breastfeeding, early language/literacy, anticipatory guidance for behaviour and sleep management, and immunisation.

In contrast, **targeted activities** are directed at children “at risk” of specific adverse outcomes. Risk may be determined by activities such as monitoring BMI, eliciting parent concerns, and determining psychosocial or sociodemographic risk factors. While the process of determining risk may for some issues resemble screening, it differs in that the response may be flexible or graded, and that services may be offered on the basis of risk alone – a diagnosable disorder does not need to be identified (although ideally the “at risk” group would contain all or almost all of those with true disorder). Importantly, the “borderline” or “grey” areas become issues worthy of consideration in their own right. Typically, far more children fall into these

borderline zones than have a diagnosable disorder; thus the potential for health benefit at a population level from secondary prevention activities delivered in the primary care or community setting is large.

Issues for which promising outcomes have already been demonstrated in the community include language enhancement programs for those with very early communication delay, high-quality day care for children at psychosocial and cognitive risk, and simple infant sleep management programs. Others urgently require research, for example secondary prevention of overweight.

This flexible, longitudinal process would seem particularly suited to areas such as child development or to psychosocial and other risk factors that exist on a continuum and which are neither easily nor desirably categorised into pass/fail. It could consist of any activities that lead to identification of risk – eliciting parent concerns, physical examination, informal observations, measurement of growth, administration of tests and procedures, referral for further assessment. These activities may be initiated by professionals but involve a partnership with parents. While they are a core component of primary health care, they may also take place in non-health care settings – child care centres and family day care schemes, preschools and kindergartens, schools – by all professionals who come into contact with young children and their families.

These activities should be continued on an individual and population basis, and further research encouraged to demonstrate efficacy and effectiveness of various strategies. We would argue that prevention and health promotion should be regarded as an integral component of a high quality primary care health system, and not separate from service delivery.

#### **9.4. PRINCIPLES AND PROCESSES FOR SERVICE DELIVERY**

Prevention, early detection and health promotion activities are inexorably linked with the delivery of services to children and their families. The three-tier system suggested – screening, surveillance, and prevention/promotion – should be considered a core component of service delivery. All children and all families would benefit from access to a universal platform of primary services – health, childcare, and education - where these activities are adopted. We suggest that consideration be given to the following principles and procedures to provide an optimal context for the detection of problems at an early stage, as well as providing interventions and support to children and families who are likely to benefit from it.

1. Early detection of problems, early intervention and health promotion activities should not be confined to the health sector. One of the advantages of an accessible and universal service system is that it provides numerous opportunities over a period of time to detect emerging or established problems and to implement activities designed to improve outcomes.
2. All professionals working with children and parents should engage the parents actively in discussions regarding their children. Formal instruments such as the Parents' Evaluation of Developmental Status (Glascoe 1997) can be used for this purpose.

3. All parents are likely to benefit from being provided with credible, age-appropriate and culturally relevant information about their child's health, development and behaviour. Where parents have voiced concerns, these are addressed in a timely and appropriate manner taking into account the context. Context factors include the child (medical and developmental history, biological risk, age), family and environment (environmental risk), the setting (childcare, community nursing, general practitioner, preschool, school) and the provider (level of training expertise).
4. Most children and families will benefit from some level of support – this will range from the provision of advice to parents through to more intensive intervention targeted to children with moderate to severe problems. The level and type of support/enrichment/ intervention again will be dependent on context (context includes child, age, perceived strengths and weaknesses), parents (level of knowledge, confidence, support available, level of concern) and available services.
5. Risk and protective factors are dynamic and their balance will likely change over time. There is variability with respect to these factors for different children, families, and in different age periods. The specific balance of risk and protective factors is complex and rarely fixed. The higher the number of risk factors, the more likely there is to be a poorer outcome if there is no additional support or intervention. A combination of biological and environmental risk factors is especially problematic. (Escalona 1982; Parker, Greer & Zuckerman 1988).
6. For development and behaviour, we recommend a conceptual move away from tests, check lists, and especially categorising children into pass/fail groups towards the concept that most children and families would benefit from ongoing contact with a universal system that is responsive to their needs. Ideally parent concerns and risk factors would be systematically elicited and addressed; a range of graded interventions offered in context; longitudinal follow-up would occur to take into account the changing nature of development and risk and protective factors; seamless referral and follow-up systems would be put into place in community networks; and the whole system would be underpinned by a system of quality assurance to ensure that structures and processes are consistent with contemporary knowledge.

## REFERENCES

Dworkin PH. British and American recommendations for developmental monitoring: The role of surveillance. *Pediatrics* 1989; 84.

Escalona SK. Babies at double hazard: Early development of infants at biologic and social risk. *Pediatrics* 1982; 70: 670-676.

Glascoe FP, Altemeier WK, MacLean WE. The importance of parents' concerns about their child's development. *Am J Dis Child* 1989; 143: 855-858.

Glascoe FP, Dworkin PH. Obstacles to effective developmental surveillance: Errors in clinical reasoning. *J Dev Behav Pediatr* 1993; 14: 344-349 (No.5).

Glascoe FP, Dworkin PH. The role of parents in the detection of developmental and behavioural problems. *Pediatrics* 1995; 95: 829-836.

Glascoe FP, MacLean WE, Stone WL. The importance of parents' concerns about their child's behaviour. *Clin Pediatr* 1991; 30: 8-11.

Glascoe FP. Can clinical judgement detect children with speech-language problems? *Pediatrics* 1991; 87: 317-322.

Glascoe FP. Parents' concerns about children's development: Prescreening technique or screening test? *Pediatrics* 1997; 99: 522-528.

Glascoe FP. The importance of discussing parents' concerns about development. *Ambulatory Child Health* 1996; 2: 349-356.

Hall DMB 1996, *Health For All Children* (3<sup>rd</sup> Edition). Oxford University Press, Oxford.

Parker S, Greer S, Zuckerman B. Double Jeopardy: The impact of poverty on child development. *Pediatr Clin North Am* 1988; 35: 1227-1240.

Shonkoff JP, Hauser-Cram P. Early intervention for disabled infants and their families: A quantitative analysis. *Pediatrics* 1987; 80: 650-658.

Starfield B, Borkworf S. Physicians' recognition of complaints made by parents about their children's health. *Pediatrics* 1969; 43: 166-172.