INTERNATIONAL DEVELOPMENTS IN NEWBORN SCREENING QUALITY ASSURANCE

Dianne Webster

on behalf of the International Society for Neonatal Screening (ISNS) Quality Assurance Committee
National Testing Centre, Auckland, New Zealand

Abstract. The five dimensions of quality are Accessibility, Acceptability, Effectiveness, Efficiency and Safety. A newborn screening program consists of a number of interlinked parts, not usually managed by the same individual, and the quality of the program is dependent on the quality of program planning, specimen collection, laboratory testing and follow up, diagnosis and treatment. Outcome assessment and feedback into program planning complete the loop. Use of the quality parameters can quantify ongoing improvements to our screening programs.

QUALITY ASSURANCE COMMITTEE ACTIVITIES

The Quality Assurance Committee of the ISNS has worked on a number of projects and has published a minimum dataset for comparison of neonatal screening program results (Webster, 1998). A survey of available quality assurance programs for neonatal screening is ready for update (Webster *et al*, 1993). Quality Assurance meetings have been held in association with the international meetings in 1991 (Webster and Hannon, 1992), 1993 (Dhondt *et al*, 1994), 1996 (Dhondt *et al*, 1996), 1999 (Webster *et al*, 1999) and a meeting is planned for 2002. In response to significant TSH kit differences observed by various quality assurance programs, the committee has produced a TSH dried blood spot standard material. This is currently under manufacturer evaluation.

HEALTH QUALITY DIMENSIONS APPLIED TO NEWBORN SCREENING

The five dimensions of quality in health are accessibility, acceptability, effectiveness, efficiency and safety (Campbell *et al*, 2000). A newborn screening program consists of a number of interlinked parts, not usually managed by the same individual, and the quality of the program is dependent on the quality of program planning, specimen collection, laboratory testing and follow-up, diagnosis and treatment. Outcome assessment and feedback into program planning complete the loop. It is essential that health outcome goals be part of the screening program planning since unless the desired outcomes are achieved, the faultless functioning of part of the system costs health dollars without contributing to improved health outcomes.

1. Accessibility

Accessibility involves who and what proportion of a population are screened. Few newborn screening programs can tell exactly which babies have been screened due to difficulties correlating birth information with newborn screening information (Loeber *et al*, 1999). This information is important as it enables systematic gaps in coverage (for example a particular region or cultural group) to be identified and remedied. Loeber surveyed 31 screening programs from 26 countries and found coverage >99% for 13 programs, 95-98% for 7 programs and <95% for 3 programs. No data was available for 6 programs and only 13 programs checked newborn screening information against a birth register (Loeber *et al*, 1999).

2. Acceptability

Acceptability refers to the extent to which newborn screening meets the needs of the population screened. It may be considered necessary to involve the population in these decisions, which will be influenced not only by genetic differences between populations but also by availability and cost of treatments competing with other local healthcare priorities. Screening programs also need to be acceptable to healthcare professionals (Wilson and Junger, 1968). It has been suggested that screening programs have advisory committees and that these involve community representation (Therrell *et al*, 1992). The Advisory Committee can contact disorder support groups and can help ensure that the screening program notification processes meet community expectations.

3. Effectiveness

Effectiveness of the screening program can be assessed only after clearly stating the health goals of the program (Human Genetics Society of Australasia). For instance, the goal of a program screening infants for congenital hypothyroidism may be to have all children with the condition show height and development parameters equivalent to that attained by their siblings. This type of outcome measure can be expensive and difficult and surrogate measures (such as time taken to normalise thyroxine, which has been shown to correlate with outcome in congenital hypothyroidism) may be used.

4. Efficiency

Program efficiency is measured by the classic screening parameters such as sensitivity, specificity and odds of being affected given a positive screen result (OAPR). To make these results comparable between programs, clear definitions of disease must be developed and comparable cut-offs used. Good performance indicators enable comparison of critical variables over time and allow evaluation of interventions designed to improve the workings of the screening program.

Performance indicators for sample collection might include

- % babies born receiving a test
- % samples suitable for testing
- % samples collected at the recommended age
- % samples with acceptable transit time to laboratory
- % samples with acceptable demographic information provided

Indicators for Follow-up might include

- % requested follow-up achieved
- % requested follow-up achieved within an appropriate (specified) timeframe
- % requested second samples arriving in an appropriate (specified) timeframe

Outcome performance indicators are poorly developed internationally. They should be specific for each disorder. Fig 1 shows the times contributing to normalisation of thyroxine levels after newborn screening as observed by the New Zealand program. The age at which normalisation occurs is the sum of the age at which the sample was taken, the time taken for shipment to the laboratory, the time taken for laboratory testing and notification of the LMC (lead maternity caregiver, the person responsible for ensuring screening happens), the time taken for treatment to commence, and the time taken for the treatment to normalise the thyroxine level. Fig 1 shows that the overall age at normalisation of thyroxine levels in New Zealand has decreased from an average of 31 days in 1995 to 16 days in 1998. Major contributors to this improvement were increased awareness among midwives of the appropriate time for sample collection and replacement of an in-house radioimmunoassay which took 3 days to complete (and hence 5 days for a confirmed result) with a commercial fluorescence assay (about 30 hrs to a confirmed result).

5. Safety

Safety includes results and the people involved. Safety of laboratory staff is covered by the framework of local legislation and safety regulations under which all laboratories work. The safety of screened infants is assured by having reliable processes inside and outside the laboratory to collect the samples, ship them and produce the screening test results. False negative and false positive tests are minimised by selection of the best available screening test with appropriate cutoffs.

A seminal study (Holtzman *et al*, 1986) showed that of cases missed by screening programs, about half the errors occurred within the laboratory and half outside. Of the 28 due to laboratory errors, 23 were due to 'clerical' errors, 1 to a misread result, 2 to switched samples and 2 to improper cutoffs. The use of 'blind' quality assurance samples can give insights into aspects of laboratory processes which may give rise to the 'clerical' type of error. Blind quality assurance samples are recommended for screening programs (Therrell *et al*, 1992).

Accreditation can reassure laboratories that they have good quality systems in place. Assessment for accreditation should be against National – International standards such as ISO9002. Accreditation against a standard incorporating documentation of laboratory processes and policies, equipment maintenance schedules and records and a quality improvement system ensures that the laboratory systems are good. Where possible, assessment against a special screening standard, eg the HGSA (Human Genetics Society of Australasia) or CORN (Therrell *et al*, 1992), guidelines (covering program organisation, communication, quality assurance, funding, diagnosis, program evaluation, liability etc) and occasional peer review ensure that appropriate screening systems, technology and cutoffs are used. Analysis of program incidence reports and adverse events (eg missed cases) enables continuous program quality improvement.

Laboratory performance indicators are necessary for the same reasons as program performance indicators

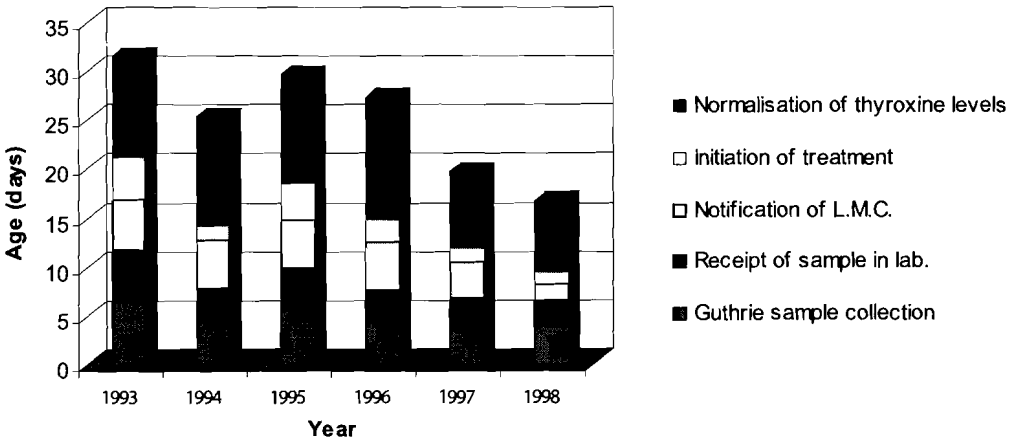

Fig 1. Times contributing to normalisation of thyroxine levels in patients with congenital hypothyroidism diagnosed through the New Zealand newborn screening program.

and may include:

- % samples tested within a defined timeframe
- % samples requiring repeat testing
- % controls out of range
- % QA samples with acceptable results

Initiatives such as the production of international standard material (ISNS TSH and European and CDC aminoacids) work with in-lab and between-lab quality assurance activities to ensure quality of testing. Use of quality parameters can enable programs to quantify ongoing improvements to screening programs and program comparisons.

REFERENCES

Campbell SM, Roland MO, Buetow SA. Defining Quality of Care. *Soc Sci Med* 2000; 51: 1611-25.

Dhondt J-L, Hannon H, Torresani T, Webster D. Summary of the meeting 'Quality Assurance in Newborn Screening', Hardelot, September 11-13 1993. In: Farriaux J-P, Dhont J-L, eds. New horizons in neonatal screening. *Exerpta Medica* 1994:379-82.

Dhondt J-L, Hannon H, Harada S, Torresani T, Webster D. Report of the Quality Assurance Workshop; Boston October 20 1996. In: Proceedings, Third Meeting of the International Society for Neonatal Screening, HL Levy, RJ Hermos, GF Grady, eds, Third International Society for Neonatal Screening, Boston, 1996: 87-90

Holtzman C, Slazyk WE, Cordero JF, Hannon WH. Descriptive epidemiology of missed cases of phenylketonuria and congenital hypothyroidism. *Pediatrics* 1986; 78:553-8.

Human Genetics Society of Australasia, Newborn Screening Policy. http://www.hgsa.com.au

Loeber G, Webster D, Aznarez A. Quality evaluation of newborn screening programs. *Acta Paediatr* 1999; 432(suppl): 3-6.

Therrell BL, Panny SR, Davidson A, *et al*. U.S. newborn screening system guidelines: Statement of the Council of Regional Networks for Genetic Services (CORN),1992; 1:135-47.

Webster D. A minimum dataset for newborn screening. *J Med Scr* 1998;5:109.

Webster D, Dhondt J-L, Hannon WH, Loeber G, Torresani T. Quality assurance and standardisation; summary of the satellite meeting, Turku, Finland 11-12 June 1999. *Acta Paediatr* 1999; 432(suppl): 7-12.

Webster D, Dhondt J-L, Hannon WH, Torresani T. Survey of quality assurance programs in newborn screening. *Screening* 1993; 2: 219-28.

Webster D, Hannon WH. Arthur Veale Memorial Meeting, Summary Report, Highlights and Issues. In: Wilcken B, Webster D, eds. Neonatal screening in the nineties. Australia: 8th International Neonatal Screening Symposium, 1992: 253-8.

Wilson JJMG, Junger G. In: Principles and Practice of Screening for Disease. Geneva: World Health Organisation, 1968.