

Carlo Corbetta and Luisella Alberti

5.1 Introduction

The aim of a newborn screening program is to identify affected babies in the newborn population. This implies that screening tests and laboratory methodologies are characterized by performance parameters (accuracy, analytical range, analytical specificity, blank reading, detection limit, interferences, precision, reagent stability, etc.) able to meet the medical usefulness of a newborn screening program and specific criteria such as sensitivity (positivity in disease), specificity (absence of disease), predictive value of positive test (percent of patients with positive test results who are diseased), predictive value of negative test (percent of patients with negative test results who are nondiseased) [1].

For what strictly concerns newborn screening for congenital hypothyroidism (CH), many aspects should be taken into consideration. A crucial aspect is to define the spectrum of pathologies to be screened: primary CH, secondary (central) CH, severe forms of CH, mild forms of CH. Moreover, the complex regulation of thyroid hormone biosynthesis, as well as the rapid modification of reference intervals of thyroid and pituitary hormones in the first weeks of life, may affect many laboratory aspects. In addition, there are features concerning the biological matrix universally used for newborn screening, the *dried blood spot* (DBS). DBS is collected on a special filter paper with a high degree of uniformity and designed to absorb a specific volume of blood. The quantitative nature of DBS allows the quantitative assessment of biomarkers of many congenital diseases and the definition of cutoff values to differentiate asymptomatic newborns that may have a disease from those who

C. Corbetta (✉) • L. Alberti

Dipartimento della Donna, della Mamma e del Neonato, Laboratorio di Riferimento Regionale per lo Screening Neonatale, Ospedale dei Bambini “V. Buzzi”, AO Istituti Clinici di perfezionamento, Via L.Castelvetro n° 32, Milan I-20154, Italy
e-mail: carlo.corbetta@icp.mi.it; luisella.alberti@icp.mi.it

may not [2]. These features meet the needs of the analytical process and guarantee efficiency and effectiveness of the newborn screening program.

It is universally accepted that only high-technology laboratories which use standardized operating procedures and have appropriately trained staff with large experience in automated immunoassay procedures, quality assurance policy, and information technology [3] can successfully process the high workload requested by a newborn screening program for CH. In fact, it is recommended that the number of babies screened yearly by a screening center is at least 35,000-50,000 [4].

The criteria on the basis of which a laboratory test may be applicable to a newborn screening program have been clearly defined [5, 6]. A screening test must be applicable to DBS samples and feasible for high-throughput platforms. Furthermore, it should have high sensitivity and specificity in order to identify the highest number of affected babies in the newborn population. It should also be inexpensive. Finally, in the USA, the Centers for Disease Control and Prevention (CDC) classified the biochemical genetic testing and newborn screening as essential laboratory services for the screening, detection, diagnosis, and monitoring of inherited metabolic diseases, endocrine and hemoglobin disorders, and other rare diseases [7]. Biochemical genetic tests and newborn screening tests are considered high-complexity tests. Laboratories that perform these tests must meet more severe regulations for the total testing process in comparison with general laboratories. Moreover, the qualification of laboratory personnel, including training and experience, is a critical factor for ensuring the quality of laboratory test results.

5.2 Biomarkers and Screening Strategies Used in Newborn Screening for Congenital Hypothyroidism

In general terms, the laboratory tests [8] most commonly used to evaluate thyroid hormone dysfunction are classified into

- Hormone concentration tests:
 - Total thyroxine (T_4)
 - Total triiodothyronine (T_3)
 - Free thyroxine (fT_4)
 - Free triiodothyronine (fT_3)
 - Thyrotropin (Thyroid-stimulating hormone) (TSH)
 - Reverse Triiodothyronine (rT_3)
- Serum-binding protein tests:
 - Thyroxine-binding globulin (TBG)
 - Thyroxine-binding prealbumin (Tranthyretin) (TBPA)
- Autoimmune Thyroid Disease tests:
 - Antithyroglobulin antibodies (TgAb)
 - Antithyroid peroxidase antibodies (TPO Ab)
 - TSH receptor antibodies (TRAb)

- Other Hormones and Thyroid-Related Protein tests:
 - Thyrotropin-releasing hormone (TRH)
 - Thyroglobulin (Tg)
 - Calcitonin (CT)

Some of these analytes (TSH, T_4 , fT_4 , fT_3 , TBG) can be detected in DBS samples (http://www.cdc.gov/labstandards/pdf/nsqap_analyte_list.pdf).

According to the clinical target of a newborn screening program for CH, i.e., detection of babies with primary CH and/or central CH, the screening protocol should use appropriate biomarkers in order to obtain a highly effective screening program [11, 12]. Generally, an analyte expressing the effect of a genetic and/or functional defect is an optimal biomarker. For example, phenylalanine is the best analyte in newborn screening for phenylketonuria (PKU). This rare inherited disorder is caused by a genetic defect causing the lack of Phe-hydroxylase, an enzyme needed to process phenylalanine. Without the enzyme, phenylalanine increases in blood. This vision explains why T_4 was identified as an optimal biomarker when newborn screening for primary CH was introduced. However, it was realized that T_4 is not an ideal indicator of thyroid status, in part because of the effects of variations in serum-binding protein levels and also because the relationship between T_4 and T_3 (the primary active thyroid hormone) is not always predictable [8]. Differently, TSH serum (or blood) concentration reflects an integrative action of all thyroid hormones in one of its target tissues, the pituitary cells that secrete TSH. Moreover, TSH serum or blood levels are inversely proportional to T_4 concentration, and a twofold change in fT_4 causes an approximate 100-fold change in serum (or blood) TSH concentration [13]. This log-linear relationship explains why little changes in T_4 level are reflected in great variations in TSH levels. This feature, which allows the identification of subclinical thyroid diseases, together with the increasing accuracy of TSH measurement which has been achieved over the years, are the main factors on the basis of which TSH is considered the best biomarker to identify primary CH in the first days of life. Even though the measurement of TSH alone is not appropriate to identify babies with central CH, which requires the combined use of other biomarkers (mostly T_4 or fT_4), the TSH-centered strategy for screening evaluation of thyroid function in newborns is both cost effective and medically efficient.

Although TSH as primary screening test is the most used screening strategy for CH worldwide, other strategies are possible:

1. T_4 -backup TSH
2. Tandem T_4 and TSH (in all samples)
3. T_4 , TSH, TBG (combined method)
4. Tandem fT_4 and TSH

In the T_4 -backup TSH strategy, T_4 is the primary screening test and TSH is measured only in samples whose T_4 concentrations are below the 10th centile. In the 1990s, this strategy was the most used worldwide. Over the years, the

improvement of the analytical and functional sensitivity of TSH assays, as well as the increasing use of the retest at 2–4 weeks of life (*serial testing*) in preterm or sick newborns [14], has increased the use of TSH as primary screening test for CH. At present, T4 is used as primary screening test only in some states of the USA and in Israel, whereas the strategy with simultaneous (*tandem*) T₄ and TSH testing is applied only in a limited number of screening programs [15]. The widespread use of TSH as primary screening test for CH is also confirmed by annual reports of the most common international program of quality assurance for newborn screening, the *Newborn Screening Quality Assurance Program* (NSQAP) of the Centers for Disease Control, Atlanta [16]. In 2013, 578 newborn screening laboratories in 73 different countries took part in the Program. Among the 492 laboratories participating in the Proficiency Testing Program, 310 participated for TSH and only 84 for T4. Similarly, among the 445 laboratories participating in the Quality Control Program, 294 participated for TSH and only 72 for T4.

It is important to understand which strategy provides the best performance metrics in relation to the aims of a newborn screening program. In a recent study, performance metrics of four screening strategies used from 1994 to 2010 were compared: T4-backup TSH, tandem T4 and TSH, TSH (no serial testing), TSH plus serial testing [17]. In terms of effectiveness, the TSH plus serial testing strategy resulted to have the best performances, although the tandem T4 and TSH strategy allows the identification of cases with central CH. It has been reported that the measurement of fT₄ can be used instead of T4 in tandem strategy. The use of fT₄ avoids the effects of variations in serum-binding protein levels on the T4 measure. In Japan, some NBS laboratories are using the fT₄ measure as primary marker with the simultaneous measure of TSH in all DBS samples: the fT₄ measurement enables the detection of CH of central origin, with an estimated incidence in Japan's regions of 1:30.833 live births [18, 19]. In the Netherlands, the Dutch NBS program is using a different strategy: primary T4 test with sequential TSH measured in the lowest 20 % and TBG measured in samples with the lowest 5 % of T4 values. Also, this strategy allows the detection of both primary and secondary CH, the latter with an incidence of 1:15.000 live borns [20].

5.3 Analytical Methods and Technology Used for Screening of Primary CH

In the last 60 years, the measurement of TSH and thyroid hormones has been performed by means of immunochemical methods based on the antigen-antibody reaction, i.e., the antigen is detected by an antibody which is used as a reagent. The availability of labeled antigens and antibodies has allowed the development of highly sensitive and specific immunochemical tests for the assessment of hormones in biological samples.

Immunochemical methods can be classified in “competitive” and “noncompetitive” methods. In a competitive immunoassay, all reactants are mixed together, simultaneously or sequentially with different sensitivity. In a noncompetitive

immunoassay, a capture antibody is adsorbed or covalently bound to a surface of a solid phase. The antigen to measure reacts with the solid-phase capture antibody. After a washing action to remove other proteins, a labeled antibody is added and reacts with the bound antigen through a second distinct epitope: after a new washing action, the bound label is determined, and its concentration or activity is directly proportional to the concentration of the antigen. The invention of noncompetitive immunoassay technology has generally been credited to Miles and Hales, who in 1968 labeled anti-insulin antibodies with ^{125}I and used them in a noncompetitive two-step immunoradiometric assay (IRMA) of insulin [21]. The analytical detection limits of competitive and noncompetitive immunoassays are determined principally by the affinity of the antibody and the detection limit of the label used. Since the sensitivity of competitive assays is defined by the association constant of antibodies, while the sensitivity of noncompetitive assays is defined by the total error, nonspecific binding, and the affinity of antibodies, it has been possible to create noncompetitive assays that are several orders of magnitude more sensitive than competitive assays. The principal examples of labeled immunoassay are

1. Radioimmunoassay (RIA): developed in the 1960s, RIA methods use radioactive isotopes of iodine (^{125}I , ^{131}I) and tritium (^3H) as labels.
2. Enzyme Immunoassay (EIA): EIA methods use the catalytic properties of enzymes to detect and quantify immunochemical complex. EIA assays are classified in
 - (a) Enzyme-Linked Immunosorbent Assay (ELISA)
 - (b) Enzyme Multiplied Immunoassay Technique (EMIT)
 - (c) Cloned Enzyme Donor Immunoassay (CEDIA)
3. Chemiluminescence Immunoassay (CLIA): chemiluminescence is the name given to light emission produced during a chemical reaction. Isoluminol and acridium esters are two examples of chemiluminescent labels used in CLIA.
4. Fluoroimmunoassay (FIA): FIA methods use a fluorophore as label. In this assay, the problem of the background fluorescence has been overcome by the use of rare earth (lanthanide) chelates and background rejection (time-resolved) procedures.

Recent data of the CDC NSQAP show that the time-resolved fluoroimmunoassay (TR-FIA) is widely used in newborn screening programs for congenital endocrinopathies (CH and congenital adrenal hyperplasia). In this technology, the design of the immunoassay involves the use of a lanthanide chelate fluorophore (generally, europium) and its detection by means of time-resolved fluorometry. Typically, the fluorescence from a europium chelate lasts many times longer than from a conventional fluorophore. This means that the measurement of the signal can take place well after nonspecific interfering fluorescence has faded away. Again, the wavelength of the fluorescence light is significantly different from that of the light used to excite the europium chelate. This difference in wavelength is known as “the Stokes shift”: a large Stokes shift allows a more sensitive measurement of the fluorescence. The time-resolved fluorometry relies on two important properties of the fluorophore which contribute to the sensitivity of the assay: the long decay time and the wide “Stokes shift.”

Preanalytical and analytical phases of the newborn screening process need technologies with a high sensitivity and suitable for sampling a small volume of blood. Generally, DBS is a blood punch of 3.2 mm of diameter, in which the mean serum volume is about 1–1.5 μl according to the neonatal hematocrit. Only technologies characterized by high sensitivity and reproducibility can guarantee an accurate, precise, and sensitive measure of a biomarker. From this point of view, FIA methods are highly reliable.

In the last two decades, there has been a technological evolution of newborn screening laboratories toward methods and immunoassay analyzer platforms able to guarantee high sensitivity, high analytical speed, and a high workload. This technological evolution has determined a roughly 100-fold improvement in the lower detection limits of analytical methods to measure TSH, the availability of highly automated immunoassay platforms, and the development of multiplex assays by using tandem mass spectrometry. Specifically, the improvement in the lower detection limits of analytical methods has led to a widespread use of blood TSH as primary screening marker for CH. It is possible to classify TSH methods according to the concept of “functional sensitivity” which is the functional detection limit of serum (and blood) TSH assays determined on the basis of low-end interassay precision characteristics. The Nomenclature Committee of the American Thyroid Association recommended that precision [as coefficient of variation (CV)] at the lower reporting limit should be in serum 10–15 % and no worse than 20 % [22]. At present, the functional sensitivity is defined as the lowest concentration of TSH at which an interassay CV of 20 % can be achieved. The first-, second-, and third-generation methods have functional sensitivity of 1.0, 0.1, and 0.01 mIU/L, respectively. Some more recent methods (fourth-generation) declare a detection limit of 0.001 mIU/L.

Concerning the availability of automated immunoassay platforms, automation of preanalytical phase has been greatly improved by the introduction of multiplex platforms punching in 96-well microplates and able to analyze a high number of samples with a complete traceability. Currently such preanalytical systems run parallel with highly automated immunoassay platforms able to simultaneously test for multiple diseases from a single sample (*multiplex platforms*). The widespread use of multiplex platforms started in the 1990s with the introduction of tandem mass (MS/MS) technology in the newborn screening for inborn errors of metabolism. This technology allows screening of about 50 metabolic diseases by measuring more analytes simultaneously (amino acids and acylcarnitines). Therefore, over the years, MS/MS technology has contributed to the dramatic increase of the number of the diseases candidate for screening programs. On this regard, the possibility to detect T_4 from a filter paper blood spot using essentially the same method as MS/MS analysis of amino acids and acylcarnitine has been recently demonstrated [23]. This method may provide a cost-effective means of analyzing both T_4 and TSH by consolidating a T_4 analysis into the MS/MS panel. Therefore, when central CH is the clinical target of a newborn screening program for CH, MS/MS may help and facilitate the necessary dual analysis of TSH and T_4 by eliminating the need for a separate assay (T_4). Laboratories utilizing only TSH as screening marker could add T_4 for minimal costs and resources if they are using MS/MS for acylcarnitines and amino acids.

5.4 Quality Assurance Policy in the Neonatal Screening Laboratory

Newborn screening represents one of the most important results of preventive medicine in childhood, as it provides useful information for the prevention of diseases characterized by high morbidity and mortality. This implies that reliability of the result of the screening test is guaranteed by newborn screening laboratories because the quality of future life of an affected baby depends on this result. To this end, criteria of good laboratory practice have been well defined. These include quality assurance policies [24] which can be classified into

- *External Quality Assessment* which allows laboratories to evaluate accuracy of their measurements on the basis of target values and to compare their performances with those of other laboratories (proficiency assay)
- *Internal Quality Assessment* which allows the daily monitoring of the test performance in terms of precision and accuracy and the early identification of any trend of systematic errors before the analytical process is out of control

Newborn screening is a complex process which starts with a blood sample (DBS) and ends with a negative or a positive screening test. The latter case results in the identification of a baby with a high risk for the screened pathology in whom the confirmation of the diagnosis with further tests and clinical referral are needed [25]. Therefore, each phase of this multidisciplinary process needs specific quality assurance policies.

Newborn screening laboratories should have policies and procedures to address the time-sensitive issues of testing and the handling of varying conditions of the infants, including specimen collection for preterm or low birth weight infants, sick newborns, or those in need of special care. Written procedures addressing specimen-related issues (timing of specimen collection and submission, sample's quality, number and sources of samples) should be consistently applied. Again, newborn laboratories should continuously monitor the performance of their screening tests and determine the need for reevaluating performance specifications as new disease information or test performance data become available [26]. Guidelines and reference procedures are currently available [7, 14, 27, 28] and allow laboratories to define their quality assessment policy.

For what strictly concerns newborn screening for CH, the wide variability of reference intervals of TSH and thyroid hormones at birth, which may be affected by neonatal factors such as gestational age, low birth weight, sex, ethnicity, age at screening, and type of delivery, should be also taken into consideration. The definition of correct reference intervals for TSH in each category of newborns (preterm, at term, acutely ill newborns, etc.) is needed to define the cutoff values that will be used to select an affected baby. The reference interval is generally the nonparametric central 95%. In a screening program, the cutoff value represents the "decision value" at which a result is considered positive. Therefore, each laboratory, to correctly determine the test's cutoff, should calculate the test's sensitivity and specificity and weigh increased detection of mild cases vs harm from recall of normal infants.

Bibliography

1. Koch DD, Peters T (1998) Selection and evaluation of methods. In Tietz Textbook of Clinical Chemistry, Thyrd Edition. C.A. Burtis and E.R. Ashwood, eds. Philadelphia, PA: WB Saunders, pp 320–335
2. De Jesus VR, Chace DH (2012) Commentary on the history and quantitative nature of filter paper used in blood collection devices. *Bioanalysis* 4(6):645–647
3. Demers LM, Spencer CA (2002) NACB Laboratory Medicine Practice Guideline. Laboratory support for the diagnosis and monitoring of thyroid disease, Section 3-I,87-97 National Academy of Clinical Biochemistry, Washington DC. Available at: <https://www.aacc.org/~media/practice-guidelines/thyroid-disease/thyroidarchived2010.pdf?la=en>
4. Cornel M, Rigger T, Weinreich S et al (2012) EU Tender “Evaluation of population newborn screening practices for rare disorders in Member States of the European Newborn Screening in Europe” Expert Opinion document. Available at: http://ec.europa.eu/health/rare.diseases/screening/index_eu.htm
5. Wilson JMG, Jungner G (1968) Principles and practice of screening for disease. WHO, Geneva. Available at site <http://www.who.int/bulletin/volumes/86/4/07-050112BP.pdf>
6. Watson MS, Lloyd-Puryear MA, Mann MY et al (2006) Newborn screening: toward a uniform screening panel and system. *Genet Med* 8(5 Supplement):12S–252S
7. Chen B, Mei J, Kalman L et al (2012) Good laboratory practices for biochemical genetic testing and newborn screening for inherited metabolic disorders. *MMWR* 61(2):1–43
8. Whitley RJ (1999) Thyroid function. In Tietz textbook of clinical chemistry, Third Edition, pp 1496–1529
9. Rastogi MV, LaFranchi SH (2010) Congenital hypothyroidism. *Orph J Rare Dis* 10;5:17, <http://www.ojrd.com/content/5/1/17>
10. Gruters A, Krude H (2011) Detection and treatment of congenital hypothyroidism. *Nat Rev Endocrinol* 18;8(2):104–13
11. American Academy of Pediatrics, American Thyroid Association, Lawson Wilkins Pediatric Endocrine Society (2006) Update of newborn screening and therapy for congenital hypothyroidism. *Pediatrics* 117:2290–2303
12. Leger J, Olivieri A, Donaldson M et al (2014) European Society for Paediatric Endocrinology consensus guidelines on screening, diagnosis, and management of congenital hypothyroidism. *J Clin Endocrinol Metab* 99:363–384
13. Spencer CA, LoPresti JS, Patel A et al (1998) Applications of a new chemiluminometric thyrotropin assay to subnormal measurement. *J Clin Endocrinol Metab* 66:588–592
14. Clinical and Laboratory Standards Institute (2009) Newborn screening for preterm, low birth weight, and sick newborns: approved guideline. Clinical and Laboratory Standards Institute Publication, Wayne, n°ILA32-P
15. Ford G, LaFranchi SH (2014) Screening for congenital hypothyroidism: a worldwide view of strategies. *Best Pract Res Clin Endocrinol Metab* 28:175–187
16. CDC National Center for Environmental Health, Division of Laboratory Sciences, Newborn Screening Quality Assurance Program, Annual Summary Report, 2013, vol 31. Available at site http://www.cdc.gov/labstandards/pdf/nsqap/nsqap_summaryreport_2013.pdf
17. Korzeniewski SJ, Grigorescu V, Kleyn M et al (2012) Performance metrics after changes in screening protocol for congenital hypothyroidism. *Pediatrics* 130:e1252–e1260
18. Adachi M, Sameda A, Asakura Y et al (2012) Mass screening of newborns for congenital hypothyroidism of central origin by free thyroxine measurement of blood samples on filter paper. *Eur J Endocrinol* 166:829–838
19. Sameda A, Adachi M, Muroya K et al (2014) Overall usefulness of newborn screening for congenital hypothyroidism by using free thyroxine measurement. *Endocr J* 61:1025–1030
20. Lanting CI, van Tijn DA, Loeber JG et al (2005) Clinical effectiveness and cost-effectiveness of the use of the thyroxine/thyroxine-binding globulin ratio to detect congenital hypothyroidism of thyroidal and central origin in a neonatal screening program. *Pediatrics* 116:168–173

21. Miles LE, Hales CN (1968) Labelled antibodies and immunological assay systems. *Nature* 219(5150):186–189
22. Hay ID, Bayer MF, Kaplan MM (1991) American Thyroid Association assessment of current free thyroid hormone and thyrotropin measurements and guidelines for future clinical assays. *Clin Chem* 37(11):2002–2008
23. Chace DH, Singleton S, DiPerna J et al (2009) Rapid metabolic and newborn screening of thyroxine (T₄) from dried blood spots by MS/MS. *Clin Chim Acta* 403:178–183
24. Slazyk WE, Hannon WH (1993) Quality assurance in newborn screening. In: Therrel BL Jr (ed) *Laboratory methods for neonatal screening*. American Public Health Association, Washington DC, pp 23–46
25. *Serving the family from birth to the medical home-newborn screening: a blueprint for the future – a call for a national agenda on State newborn screening programs* (2000) *Pediatrics* 106 Suppl 2:389–422
26. McHugh DM, Cameron CA, Abdenur JE et al (2011) Clinical validation of cutoff target ranges in newborn screening of metabolic disorders by tandem mass spectrometry: a worldwide collaborative project. *Genet Med* 13:230–254
27. Clinical and Laboratory Standards Institute (2007) *Blood collection on filter paper for newborn screening programs: approved guideline*. Clinical and Laboratory Standards Institute Publication, Wayne, n°I.A4-A5
28. Clinical and Laboratory Standards Institute (2006) *Newborn screening follow-up: approved guideline*. Clinical and Laboratory Standards Institute Publication, Wayne, n°IL.A27-A