1. Introduction

The incorporation of newborn screening into the states’ departments of public health represented the formal initiation of population-based predictive medicine with the goal of prevention of morbidity and mortality from genetic disease. Starting with a methodology based on microbiology, the bacterial inhibition assay [1], the field has incorporated far more advanced technologies such as molecular genetic analyses [2–7] and tandem mass spectrometry (MS/MS) [8–12]. Prediction and prevention are fundamental to public health and genomic medicine. Effective integration of screening across the lifespan for genetic predisposition and disease will require utilization of advanced technologies and information systems, and incorporation of appropriate safeguards to protect autonomy, privacy and confidentiality.

As newborn screening testing menus expand and population-based genetic screening extends through adulthood, pilot programs must be developed and evaluated to determine their ability to achieve the goals and objectives for which they are designed [13]. Such pilot programs will require careful design, not only to assess the technological bases of these systems, but also to evaluate their impact on the individuals tested. Determination of impact must include not only traditional effectiveness measures such as changes in morbidity and mortality from disease, but also the ethical, legal and social implications of testing such as genetic discrimination [14–18].

Design and implementation of genetic screening programs will require broad-based expertise. In addition to the traditional membership on screening program advisory committees, such as medical specialists and subspecialists, laboratorians, and public health and government officials, these groups must represent all stakeholders [13]. Affected individuals and their family members have the best-informed experience with the process and consequences of screening, and they should be incorporated as full and equal members.

1.1. Role of screening in public health

The goal of screening has been to identify disease before irreparable damage is done, to determine individuals whose offspring are at risk for disease, and to evaluate new screening strategies and their potential impacts [19–21]. Implementation of these principles within a public health paradigm means that they will translate to presymptomatic identification of affected individuals, recognition of carriers of genetic diseases, and performance of research to improve screening programs. Therefore, screening for genetic disease will fulfill the public health need of providing an ever-evolving “safety-net” to predict and prevent disease.

While screening may involve clinical evaluation of individuals, here we will be concerned primarily with laboratory-based screening programs. By their very nature, such programs will be driven by advances in laboratory technology. Until very recently population-based screening programs, such as neonatal screening, have relied on biochemical measurements of analytes that have included: metabolites (e.g., phenylalanine for phenylketonuria, or galactose and galactose 1-phosphate for galactosemia); hormones (e.g., thyroid stimulating hormone and thyroid hormone for hypothyroidism, or 17-hydroxyprogesterone for 21-hydroxylase deficient congenital adrenal hyperplasia); and proteins (e.g., hemoglobin for hemoglobinopathies, galactose 1-phosphate uridyltransferase for galactosemia, or immunoreactive trypsinogen (IRT) for cystic fibrosis (CF)) [21,22]. Beginning in the 1980s, the tools of molecular genetics
were shown to be applicable to newborn screening for analysis of DNA (e.g., for hemoglobinopathies [3–5,7], cystic fibrosis [23–26], and medium chain acyl-CoA dehydrogenase (MCAD) deficiency [6,27,28]); and RNA (e.g., for hemoglobinopathies [29,30]).

1.2. Decision: who to screen?

A fundamental decision in the construction of a screening program is determination of the group of individuals who should be screened in order to meet the goals for that program. Newborn screening represents a program designed for an entire population cohort, i.e., all neonates. Diseases selected for such a program will have a broad distribution throughout the population [19–21]. Alternatively, for diseases that have much higher frequencies in certain subpopulations, a specific ethno-cultural group or a cohort originating in a defined geographic region may be identified for screening. An example of such a program is carrier screening for Tay–Sachs disease among Ashkenazi Jews [31]. As we will discuss below, the decision of which group is to be screened will begin to determine the venue for screening in order to maximize participation within that group. For example, if one desires to screen all neonates, then the best place to accomplish this within the health care systems of North America is in newborn nurseries. Screening of populations after the neonatal period will present significant challenges to achieve total participation.

1.3. Principle: minimize false negatives

The goal of screening is to identify 100% or as close as possible to 100% of the individuals at risk for a disorder [19,20]. If a screening test determines that an individual is not at risk for developing a disease, but that individual develops the disease, then this is called a false negative result or a missed case. Conversely, if a screening test indicates that an individual will develop a disorder, but the diagnostic test shows that they do not have the disease, then this is called a false positive result or a false alarm. While a false positive result is concerning to the individual family until the diagnostic test reveals that the individual does not have the disease, the false negative result is much more devastating. If an individual eventually develops the symptoms of a disease that could have been prevented with early diagnosis and treatment, the screening program has failed that individual.

1.4. Population screening as a system

One goal of screening is to identify individuals with the intent of intervention [19,20]. For screening to achieve this goal, a system must be developed that involves not only the screening test, but also a comprehensive systematic integration of the preanalytic, analytic, and postanalytic phases [13]. All phases of the screening system must be well integrated so that voids are not created in which samples and information will be lost [13]. The preanalytic phase includes education, informed decision-making, sample collection, and sample and identifying information transmission. The analytic phase involves sample preparation and testing in the laboratory. The postanalytic phase includes follow-up on all tests and appropriate management of a positive test result, which will entail communication of the test results to the individual’s medical home, and to the individual, or, in the case of a neonate, to the parent(s). If the screening test is positive, then diagnostic testing is indicated to determine if the individual has the disease. If the screening test result is a false positive, then the individual or the family should receive counseling and reassurance that the individual is not affected. If the individual is diagnosed with the disorder, then the individual and/or family should be counseled about the implications for other family members and for future reproductive decisions. In addition, the affected individual should be referred to a specialist who would prescribe treatment and provide follow-up care. Therefore, the duration of the postanalytic phase, and consequently of the entire screening system, could be life-long.

1.5. Risks of screening

The risks of screening include, but are definitely not limited to, the possibilities of stigmatization, genetic discrimination in life and health insurance and employment, and determination of misattribution of parentage [14–18,32]. An individual with a positive screening result, or a group with a high incidence of positive screening results, may be stigmatized by the screening test or even the screening process. In addition to the impact on self-concept, the screening test result could also impact the individual economically when applying for life or health insurance. The positive screening result and diagnosis could lead to higher premiums or denial of insurance due to a preexisting condition. This could also impact employment, since in the case of small independent companies that are self-insured, the employer is also the health insurer. Sometimes the determination that an individual has a disease may result in the finding that one or both of the parents are not the biological parents. A committee of the American Society of Human Genetics concluded that it was not appropriate to reveal misattribution of paternity unless the testing was specifically for paternity [33]. However, such a policy can lead to misinformation regarding carrier status or even the risk of disease. It is optimal to include discussion of risks through the pro-
cess of education and informed decision-making in the preanalytic phase.

2. Newborn screening

Newborn screening for PKU using neonatal blood specimens was initiated in the early 1960s using a bacterial inhibition assay [1]. From the very beginning of newborn screening, a key component was the use of dried blood on blotter paper. This specimen collection format was used because it was shown that the analyte phenylalanine was stable in the dried blood spots, and there was a desire to have specimens sent to centralized testing laboratories to facilitate quality control.

State laws that were passed in the 1960s mandating neonatal PKU screening resulted from grass-roots efforts on the part of the National Association for Retarded Children, whose membership was motivated to eliminate the mental retardation from untreated PKU [34]. Advocacy by grass-roots organizations resulting in legislative action for newborn screening has continued to the present (e.g., http://www.tylerforlife.com).

Advances in the screening for congenital hypothyroidism, the hemoglobinopathies, and other disorders resulted in the addition of new diseases to the state newborn screening testing batteries, beginning in the 1970s [21,35–41]. Pilot programs in the 1990s established the feasibility of applying tandem mass spectrometry (MS/MS) to newborn screening, and states began to include this technology in their screening programs in the late 1990s [8–11].

2.1. Concept of two-tiered screening

Since the goal of any screening program is to minimize false negatives, any strategy that increases sensitivity without reducing specificity will be desirable. A well-conceived and empirically driven two-tiered testing approach will achieve this goal.

Newborn screening for congenital hypothyroidism was implemented when it was demonstrated that such a two-tiered strategy was feasible on a large scale. Since the vast majority of patients with congenital hypothyroidism have the primary form of this disorder with low thyroid hormone (T4) and elevated thyroid stimulating hormone (TSH), it was recognized that it would be optimal to include both of these analytes [36]. The competing strategies involve which test, T4 or TSH, will be the first or second tier [21]. The strategy that is used by any program or region is determined by the technology available, the relative expense of measuring the respective analytes, and the beliefs of those who establish the program. If T4 is the first-tier test and only those specimens with low T4 are tested for TSH, then those individuals with congenital hypothyroidism but residual thyroid gland function and borderline low T4 will be missed [21], even if they have elevated TSH, since they will never have TSH tested. Both strategies will miss those with secondary (pituitary abnormality with TSH deficiency) and tertiary (hypothalamic abnormality with deficiency of thyroid releasing hormone, TRH) hypothyroidism, since specimens from these individuals, although they will have low T4, will not have elevated TSH; if T4 is the primary test, but a positive test report requires elevated TSH, they will not be reported, and if TSH is the primary test they will not be tested with T4 since TSH will be below the threshold [21,36–38]. Although the two-tiered screening strategy for congenital hypothyroidism demonstrates improved sensitivity over either test individually, it also illustrates the need for the health professional to maintain clinical vigilance and exercise clinical judgment despite the presence of an excellent newborn screening program. If a patient has the signs and symptoms of one of the disorders included in a screening program, then clinical suspicion should override the normal newborn screening test result, and the definitive diagnostic test(s) should be performed [21]. Screening tests are not infallible, for biological [21,36–38,42] and/or technical or clerical [43] reasons.

The demonstration that DNA is stable in the neonatal dried blood spots [2] added another potential analyte for two-tiered strategies. Typically, addition of molecular genetics to screening approaches involves analytical measures for the biochemical phenotype as the first tier and analysis of genotype as the second tier. Examples include newborn screening for the hemoglobinopathies, in which evaluation of the protein phenotype is the first tier [5]; and CF, in which the measurement of IRT is the first tier [23]. In both of these examples, the second tier involves analysis of selected mutations. For hemoglobinopathy screening this approach permits more rapid initiation of antibiotic prophylaxis [5] and reducing the risk for death from overwhelming sepsis [40,44]. In newborn screening for CF, DNA analysis on the initial blood spot allows the cut-off for IRT to be reduced without overwhelming false positives, i.e., improved sensitivity without loss of specificity [25]. The costs of DNA analysis for the second tier in hemoglobinopathy and CF screening have been estimated at $10–25 and $3–5 per test, respectively [5,25].

2.2. Newborn screening task force report

Immense disparities exist in diseases included and technologies utilized in newborn screening among the District of Columbia and the 50 state programs [13]. All 51 programs screen for PKU and congenital hypothyroidism, but some screen for only three disorders and others screen for more than 30. The technologies used by the various programs are also remarkably discrepant.
For example, some laboratories screen for PKU using the original bacterial inhibition assay introduced in the 1960s, in which the phenylalanine concentration is measured by the size of the growth zone around the disk punched from the blood spot. Other laboratories, using more recently developed technologies, simultaneously measure the phenylalanine and tyrosine concentrations [45,46], an approach long-recognized to improve sensitivity and specificity [42,47–49].

On May 10–11, 1999, a Newborn Screening Task Force was convened, which was co-sponsored by the American Academy of Pediatrics (AAP) and the Genetics Disease Branch of the Maternal and Child Health Bureau (MCHB), Health Resources and Services Administration (HRSA) [13]. Among the conclusions of the Task Force were that newborn screening must be regarded as an integrated system and that there was need for a national agenda in newborn screening. The Task Force recommended that this national agenda should consider both the diseases and technologies that all US infants should have the right to experience.

The March of Dimes Birth Defects Foundation challenged the Task Force report [50]. They argued that the Task Force should have developed a specific list of recommended tests, and should not have been satisfied with stating the need for a national agenda. They also maintained that there was too great an emphasis on the requirement for cost-effectiveness of screening in the report, and stated that specific tests should be implemented if they would improve health.

2.3. New directions in newborn screening

Many blood-based neonatal screening programs are adding new diseases and new technologies. Among the diseases being added are CF and CAH [13,51–54]. The addition of MS/MS will increase the number of diseases by up to 20 or more [10,12]. Some experts are concerned that if resources are not increased for the screening system, then the addition of new tests will erode the funding for the established programs such as for follow-up and management of PKU and congenital hypothyroidism [13]. In addition only a few of the disorders that are identifiable by MS/MS, such as MCAD deficiency [55] and glutaric academia type 1, appear to be treatable, though for children with these treatable disorders intervention may be life saving [13,56]. However, if one of the currently untreatable diseases should be diagnosed eventually when signs and symptoms develop, then MS/MS screening may prevent expensive “diagnostic odysseys” [57], and earlier diagnosis and coordinated interventional investigations may improve outcomes [58–60].

Universal neonatal hearing screening uses a functional assessment of a physiologic parameter, and therefore represents a departure from blood-based, “heel-stick” screening [13]. The goals of these screening programs are quite similar: to identify the neonate pre-symptomatically, to intervene before irreparable damage is done, and to screen 100% of the newborn population. To achieve these goals the appropriate screening venue is the newborn nursery. The frequency of deafness is estimated to be as high as 1/500 [61]. Approximately 50% of congenital deafness is due to genetic causes and 80% of inherited deafness is nonsyndromic and autosomal recessive [61]. Mutations in the gap junction protein, connexin 26, represent about 50% of autosomal recessive, nonsyndromic inherited deafness [61]. Connexin 26 mutations, therefore, may account for 40% of childhood deafness or up to 1/2500 newborns. A specific mutation, 35delG, accounts for 75–80% of connexin 26 mutant alleles in a mixed North American population [61]. The frequencies of individual mutations, however, have been shown to differ among different ethnocultural groups [62]. If genotypic analysis is to be used as a second tier after initial physiological screening, then the molecular epidemiology of mutations in connexin 26 and other genes responsible for inherited deafness must be understood. Mutation evaluation may accelerate definitive diagnosis and intervention, and therefore has been recommended for inclusion in neonatal hearing screening programs [63].

3. Screening selected populations

An effective screening program requires complete testing of the population identified for screening. This may involve testing of an entire population such as in the evaluation of all neonates in newborn “heel-stick” and hearing screening. Other programs may target specific subpopulations for screening. We will examine two such programs. One of these, hemoglobinopathy screening, will illustrate the importance of an evidence-based strategy to show how misconceptions may lead to inappropriate targeting of a subpopulation when the entire population should be engaged in the program. The second, Tay–Sachs and β-thalassemia carrier screening in Montreal, demonstrates the creation of a screening venue outside of the newborn period and the educational opportunities that can be achieved through a screening program.

3.1. Hemoglobinopathy screening

The hemoglobinopathies include sickle-cell disease (SS, SC, and S/β-thalassemia), Hemoglobin E, and many others [64]. Sickle-cell disease occurs in one out of every 400 African Americans [21]. This autosomal recessive disorder is also observed among individuals of Arab, East Indian, Middle Eastern, and Mediterranean
descent. Individuals with sickle-cell disease are at risk for potentially lethal pneumococcal sepsis early in life. A multicenter, randomized, placebo controlled trial showed an 84% reduction of the incidence of sepsis among infants with sickle-cell disease treated with penicillin prophylaxis [44]. The study was prematurely terminated when it was recognized that there were three deaths in the placebo group compared with no deaths in the penicillin-treatment group. The investigators concluded that newborn screening for sickle-cell disease would permit affected children to begin penicillin prophylaxis before four months of age. An NIH Consensus Development Conference on Newborn Screening for Sickle-Cell Disease and Other Hemoglobinopathies recommended that "every child should be screened for the hemoglobinopathies to prevent the potentially fatal complications of sickle-cell disease during infancy" [40]. Despite this recommendation by a national body more than 14 years ago that there be universal screening for sickle-cell disease, some states still do not screen their neonates for the hemoglobinopathies, arguing that their population demographics do not warrant such a program. Other states have "targeted" newborn screening for the hemoglobinopathies. Based on the number of individuals not being screened for hemoglobinopathies it was estimated in 1992 that approximately 90 infants with sickle-cell disease were being missed each year, and that number has not changed substantially in the interim [65]. Newborn screening for the hemoglobinopathies demonstrates that we must not only determine a national agenda, because that was accomplished for sickle-cell disease in 1987 [40], we must also have the political will to implement this agenda.

3.2. Tay–Sachs and β-thalassemia carrier screening in Montreal

Screening for Tay–Sachs disease is typically focused on the Ashkenazi Jewish community, since the disorder occurs at increased frequency among members of this ethnocultural group [31]. β-Thalassemia likewise occurs in specific ethnocultural groups, including individuals from the Mediterranean region [21,64]. For more than 25 years Montreal senior high school students have had genetics education and the opportunity to volunteer for Tay–Sachs or β-thalassemia carrier screening [66]. The student and a parent must consent to participate in this screening effort. Confidential results of carrier screening are reported directly to the student, and individuals identified as carriers are provided with genetic counseling. Subsequent evaluation has shown that carriers identified by this program remember their status. When they are considering child-bearing, many have their partner's carrier status determined if it is not known, and they and their partner are seen for prenatal counseling. The incidences of these two diseases decreased by 90–95% since the inception of the program. Targeted screening is supported by ethnic community organizations, parent–teacher organizations, and the school board.

Given the differences between the fragmented, for profit US health care system and the Canadian universal health care system, there are significantly different consequences to knowing your carrier status in the US and Canada [67]. Carrier status would not lead to discrimination in the Canadian health care system. In addition, the US health care system does not universally recognize 14 years of age as the age of health care independence as Canada does. The Montreal model for focused screening in high school shows that adolescents make appropriate use of genetic information, but legal and cultural differences between Canada and the US will make it difficult to model this program in the USA.

4. Screening programs “under construction”

Larger-scale, population-based screening outside of the neonatal period is on the horizon for a number of genetic diseases. We will discuss several examples.

4.1. Hemochromatosis screening

Hereditary hemochromatosis (HH) involves excess intestinal iron absorption, leading to increased serum ferritin saturation, increasing iron stores, and tissue damage [68]. The College of American Pathologists (CAP) Task Force concluded that survival of HH patients was extended by phlebotomy and improved by absence of cirrhosis and diabetes mellitus [69]. The CAP Task Force documented improvement with phlebotomy, liver enzymes normalizing in 50–90% of HH patients and diabetes control improving in one-third of HH patients. The CAP Task Force recommended HH screening for everyone over 20 years of age.

A Center for Disease Control–National Human Genome Research Institute (CDC-NHGRI) Expert Panel [70] concluded that genetic testing should not be used for screening for HH. The CDC-NHGRI Panel was concerned that there were not enough data on the prevalence and penetrance of the mutations, or on the best care of those with mutations, and that there was the possibility of genetic discrimination against those with HH.

Pilot screening programs for HH often involve a two-tiered approach, with initial screening for phenotype followed by genotype. The penetrance of HH is less than 100% among those with mutations, and at this time it is impossible to predict on an individual basis who will develop the HH phenotype [70–75]. Therefore, unnecessary overtreatment of and discrimination against those who will not develop disease are feared if genotypic
screening is performed [17,71]. Genotypic analysis could become the primary screen if the costs of screening follow-up for all positive individuals and the possibility of discrimination were reduced [74]. Any screening program would have to include a large-scale effort of public education regarding the availability, risks, and benefits of screening [32,71]. Care must be given to understanding the distribution of mutations in ethnocultural groups, because without these data we will not have effective genotypic primary screening or confirmatory testing [72,76].

4.2. Cystic fibrosis carrier screening

Cystic fibrosis has a carrier frequency of 1/25–1/30 among those of Northern European origin, with a lower frequency in other ethnocultural groups [21]. In 1997, an NIH Consensus Development Conference [77] recommended that mutation screening for CF mutations be offered for adult family members of patients with CF, partners of CF patients, couples planning a pregnancy, and couples seeking prenatal care. The biggest obstacle to establishing carrier screening for CF is the mutational heterogeneity with more than 900 mutations reported [78]. A joint committee with representatives from the American College of Medical Genetics (ACMG), American College of Obstetricians and Gynecologists (ACOG), and National Institutes of Health (NIH) developed guidelines for population-based screening for CF [79]. This group also provided guidelines for health care professionals and described the need for educational materials for individuals considering CF carrier testing. The group recommended a panel of 25 mutations with a frequency of \( \geq 0.1\% \) in the general population. Because of the complexity of CF mutation testing, they encouraged laboratories to participate in the CAP/ACMG quality assurance and proficiency testing programs.

4.3. Type 2 diabetes mellitus screening

Genetic screening for common disorders will certainly be part of predictive genomic medicine in the future. Type 2 diabetes mellitus will represent such an opportunity [80], though at this time the genetic component is known for only a small group with this disorder [81]. Maturity onset diabetes of the young (MODY) is an autosomal dominant form of non-insulin-dependent, type 2 diabetes mellitus [81]. The criterion for the diagnosis of MODY includes three generations showing autosomal dominant inheritance and two patients with onset \( \leq 25 \) years of age. Lehto et al. [82] screened Scandinavian families with MODY and found that 13% had a mutation in one of the four MODY genes for which they tested. Screening for mutations in MODY genes before adolescence would provide the opportunity to intervene with diet and exercise to prevent the obesity associated with MODY. The molecular epidemiology of diabetes must be elucidated more fully and the screening venue will have to be identified.

5. Summary and conclusions

5.1. Population screening: utilization of genomic technology for the public’s health

Informed population screening avoids health disparities and allows the application of improved technology for the benefit of all. This requires sufficient data about mutation frequency and penetrance for various ethnocultural groups in order to provide effective genetic counseling regarding the clinical utility of the information from the test.

5.2. Systems development and evaluation: ethical, legal, and social issues

The public needs to be educated to the fact that any collection of blood or other tissues is a DNA database. This includes residual newborn screening blotters and forensic specimens. Anyone contemplating genetic testing should be counseled that they may receive unanticipated information that could include information regarding misattribution of parentage or an unexpected disease association. An overriding principle in the development of public health screening systems is the requirement of benefit to the individual being tested when performing presymptomatic testing.

5.3. Oversight of screening programs: role for all stakeholders

The AAP/MCHB Newborn Screening Task Force recommended that all stakeholders be involved in setting policy for newborn screening [13]. These individuals include affected individuals, parents, health professionals, public health officials, laboratory personnel, and legislators. Among the issues that should be reviewed are the diseases included in newborn screening and the methods to be used. Policy should be established to ensure effective education of the parents before the newborn screening sample is taken and informed consent for research use of the sample. Because residual newborn screening specimens are DNA databases, newborn screening programs need to establish or develop relationships with Institutional Review Boards to consider applications from researchers for access to the samples and to determine what information would be linked to each sample while providing privacy and confidentiality. Newborn screening laboratories participate in national quality assurance programs through the
CDC. They also need to combine data from pilot testing of new methods or for new diseases so that all programs can make an informed decision regarding the efficacy of newborn screening methodologies and/or inclusion of new disorders. The AAP, MCHB, and March of Dimes will continue to provide leadership in newborn screening with the goal of developing effective, cost-effective programs that are similar across state lines, so that no child will be denied the opportunity for life-saving medical intervention on the basis of geography.

References


[40] NIH Consensus Development Conference Statement on Newborn Screening for Sickle Cell Disease and Other Hemoglobinopathies, vol. 6, No. 9, April 6–8, 1987.


[74] P.C. Adams, A.E. Kertesz, C.E. McLaren, R. Barr, A. Bamford, S. Chakrabarti, Population screening for hemochromatosis: a comparison of unbound iron-binding capacity, transferrin satu-


