Technological Journey From Colorimetric to Tandem Mass Spectrometric Measurements in the Diagnostic Investigation for Phenylketonuria

Donald H. Chace, PhD[1] and William H. Hannon, PhD[2]

Abstract

Phenylalanine analysis for phenylketonuria (PKU) detection in newborn screening (NBS) was chosen as the model system to describe how advancements in laboratory technology improved laboratory performance. These advancements have made NBS programs better and have improved the health outcomes of the affected newborn through improvements in accurate early detection over the past 50 years. The most current state-of-the-art technology, tandem mass spectrometry (MS/MS), has proven that it is now the choice in almost all modern NBS facilities because it is a versatile instrument that continues to grow in its application not just for amino acid and acylcarnitine detection but for other metabolites and disorders such as lysosomal storage diseases and second-tier detection of some screen-positive results. The use of MS/MS will continue to expand, even with the anticipated introduction and expansion of molecular screening methods into NBS programs. Regarding technological advancements, the future of NBS will include even newer technologies and approaches that will enhance the detection and treatment of newborns affected by PKU and other inborn errors of metabolism.

Technology

The most utilized and modern analytical chemistry techniques in medicine today are well rooted in the knowledge and experience of all the technology that preceded it. This is clearly the case in newborn screening (NBS) for inherited metabolic diseases (IEM),[1] where today one of those new technologies is tandem mass spectrometry (MS/MS). It is currently the best method to screen for phenylketonuria (PKU), the very first disorder that became NBS in the 1960s, and for many other metabolic disorders. The improvements in the screening of PKU using MS/MS not only relied on its technology of mass separation and detection but also on technology that included sample collection on filter paper, automated sample preparation in 96-well plates, flow injection analysis, and ultimately computer-assisted data interpretation and reporting. Many methods contributed to these improvements as shown in a timeline[2] of the technological advancements for PKU detection as illustrated in Figure 1.

The Fabric of NBS

Hundreds of millions of babies have been screened worldwide for PKU since its public health introduction in the United States and internationally more than 50 years ago.[1] The technology behind the initial population-based screening was a biological assay using a unique sample consisting of a drop of blood on filter (absorbent) paper. The filter paper for blood collection has been improved while we embraced the state-of-the-art mass spectrometric techniques for screening, confirmation, and monitoring. Three technological threads characterize the advancements leading up to today's application of multiplex metabolite analysis. These are the dried blood spot (DBS) sample, highly specific and selective analytical methods, and multiple metabolite analysis/multiple disorder models leading to

[1] Pediatrix Analytical, Mednax Centers for Research, Education and Quality, Sunrise, FL, USA
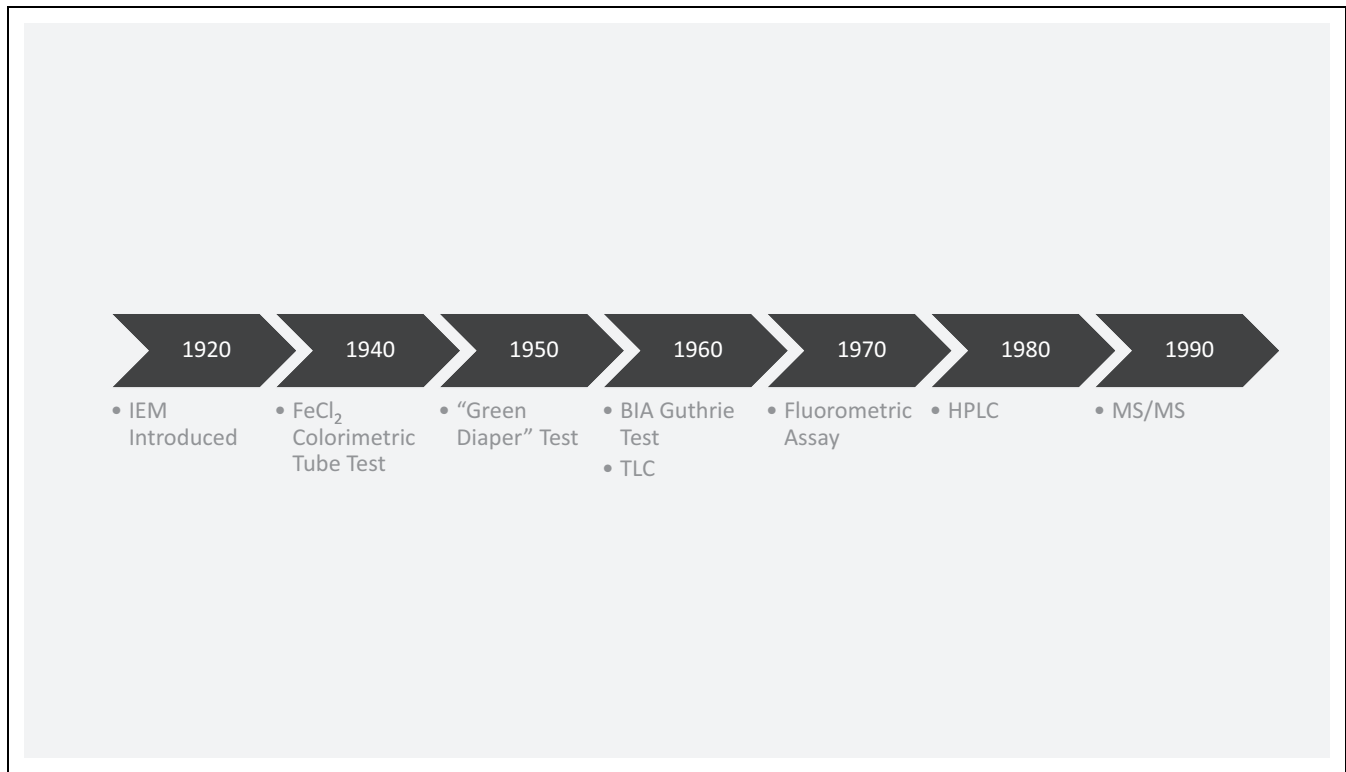[2] CDC Foundation and Newborn Screening Branch, Centers for Disease Control and Prevention (CDC), Atlanta, GA, USA

Corresponding Author:
Donald H. Chace, Mednax National Medical Group, 1301 Concord Terrace Sunrise, FL 33323, USA.
Email: donald_chace@mednax.com

**Figure 1.** Timeline for PKU detection methods from 1920 to present. PKU indicates phenylketonuria.

higher clinical sensitivity and specificity while also increasing the number of disorders from 1 to as many as 34 core conditions.[3] Even with dozens of disorders now being tested, PKU remains the quintessential NBS test and is a key part of these 3 threads, making it the essential fabric of population-based NBS.

The first thread, the DBS. The backbone of NBS by which all analytical technology must be adapted from sample preparation, analysis, and interpretation is the DBS, a matrix for which blood is applied, cotton-based filter paper. An entire book has been written describing the use of DBS in medicine.[4] From a technological perspective, cotton-based filter paper has unique properties, which are its extensive hydrogen bonding that keeps its form intact when blood is applied and then dried. Second, this absorption is sufficiently reproducible with standardized blood collection that some paper-manufactured sources are considered a US Food and Drug Administration–approved device. Third, the dried blood filter paper matrix stabilizes both small molecules (metabolites) and macromolecules (proteins such as enzymes) so that they are not denatured and can be fully reconstituted similar to a lyophilized sample. Figure 2 shows the technical aspects of the paper from a laboratory perspective. Blood spots are 3 dimensional and have volume (a small cylinder, $V = \pi r^2 h$, where r is the radius and h is the thickness of the paper). Typically, a blood spot collection area is bounded by dashed lines and is a circle of approximately 12 to 13 mm. A punch (B) can be obtained anywhere in that zone. Blood is applied to the top of the card (D) where it penetrates

the paper and emerges on the bottom (E). The diameter on the application side is often slightly larger than the penetration side and as such—to ensure obtaining homogeneous blood throughout the punched disk—the punch is taken within the dashed line circle saturated with blood. The third graphic in Figure 2v shows the evolution of punch sizes from 1/4 inch (H) to more common punch sizes used today, 3/16 inch (4.8 mm, G) and 1/8 inch (3.2 mm, F).[5]

The Birth of NBS

In 1934, Asbjörn Fölling, a Norwegian physician, reported the discovery of PKU, the cause of severe mental retardation in 2 young children in Oslo, Norway.[6,7] Phenylketonuria is termed an "error of metabolism" because children who have it cannot metabolize phenylalanine (Phe), an amino acid making up a low percentage of all food proteins.[6] The inborn error is due to the defective enzyme, Phe hydroxylase (PAH), as illustrated in Figure 3. This illustration shows the primary sources of Phe in blood (diet and protein turnover) and alternate metabolic pathways for Phe in addition to catalysis by PAH. Tyrosine (Tyr) is the major product for normal Phe metabolism. Its blood pool is sustained by production from Phe, dietary intake, and protein turnover. Tyrosine is a precursor for several important biomolecules including L-dopa (a precursor to epinephrine and norepinephrine) as well as thyroid hormones T3 and T4 (thyroxine). Phenylalanine and Tyr when present in excess are eliminated either through the urine as phenylketones or via the
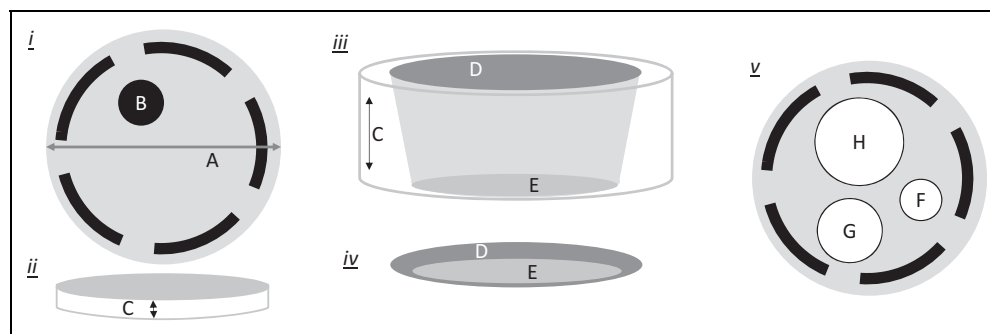
Figure 2. Features of DBS filter paper. i, The diameter of the blood spot (A). The dashed circle simply a target for application. A blood punch from the blood spot (B). ii, Filter paper has width (C). iii, There is a top (D) where blood is applied and bottom side (E) where a punch is made. The blood does not flow vertically through the width (C) but rather narrows from top to bottom forming a conical shape. iv, Typically, there is more blood apparent when viewed from the blood application side (D) versus punching side (E). A punch from side (E) guarantees blood throughout the thickness of the paper. v, Typical punch sizes include 1/4-, 3/16-, and 1/8-inch diameters relative to the 3/4-inch dashed zone. DBS indicates dried blood spot.
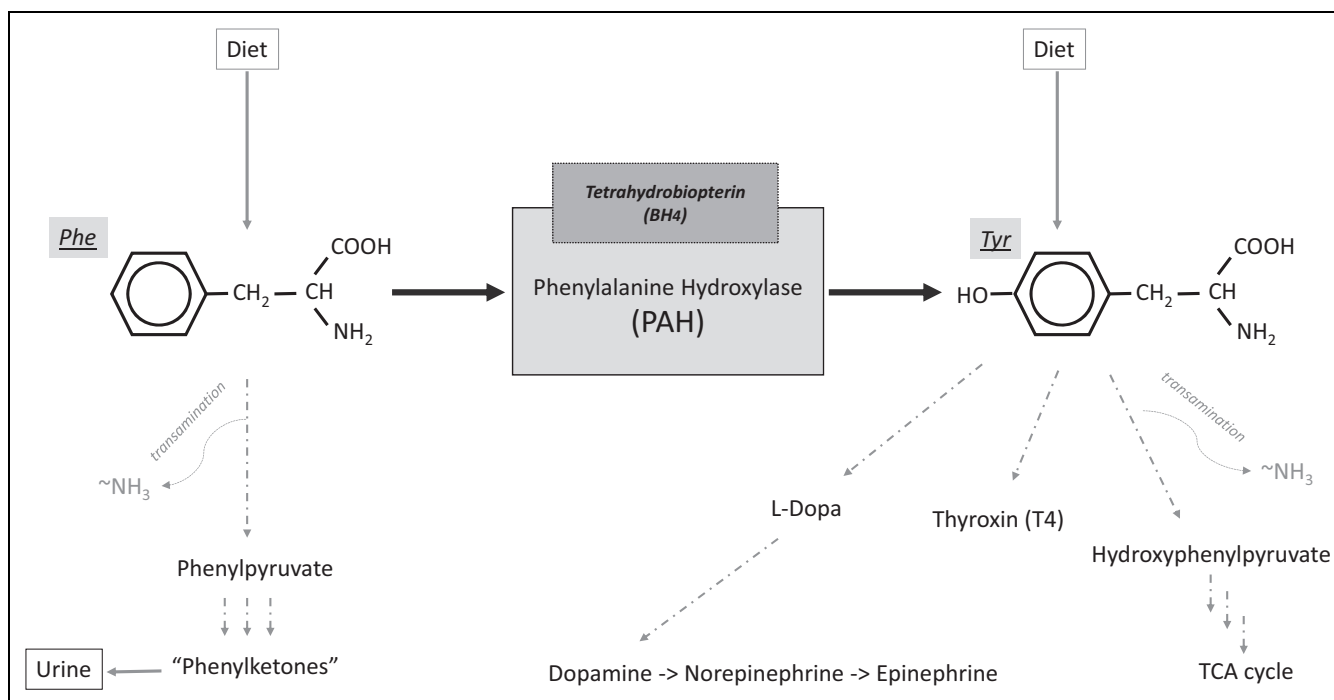
Figure 3. Phenylalanine metabolism.

Krebs cycle in the case of Tyr. From Figure 3, it is clear that 2 approaches are possible for the detection of PKU, an elevated Phe in blood or its phenylketone metabolites in urine.

Pre-Guthrie, Colorimetric Assays for PKU

The colorimetric test-tube assay with ferric chloride originated by Fölling is the oldest, the best known, and was the most widely used of any of the diagnostic urine tests for PKU until the early 1960s. In the late 1950s, an NBS program was initiated in California by utilizing the ferric chloride test on recently urine-wet diapers.[8] The diaper test is a modification of the ferric chloride test-tube assay. A drop of a ferric chloride solution is placed on a baby's urine-wet diaper, and a blue-green to gray-green color appears immediately.[7] Another modification of the ferric chloride test is the Phenistix dipstick, a commercially prepared paper strip that is impregnated with a buffered ferric chloride salt.[8] The dipstick test was an excellent procedure for routine screening of apparently well newborns because the dipstick could be pressed against the urine-wet diaper or dipped directly into a urine solution. Another urine test introduced for PKU detection was the dinitrophenylhydrazine (DNPH) tube test, when properly prepared and used, the DNPH reagent was at its time the most sensitive and reliable of the various urine tests for PKU.[9]

In 1958, the method for collection of urine from newborns absorbed into squares of Whatman filter paper was described as part of a program for the early detection of PKU.[10,11] The filter paper test was performed on ordinary white filter paper, which was wetted with urine, dried, and sent in an envelope by mail to an experienced testing laboratory. The test parameters for the filter paper urine were essentially the same as those for the diaper test. The paper method involved placing an absorbent filter paper of 2-inch square into a newborn's diaper, allowing it to become wet (or dipping directly in urine) and then allowing it to dry. A small piece cut from the dried filter paper urine sample was the aliquot for analysis.[10,11] This technique of using dried urine paper demonstrated its utility in routine testing for phenylpyruvic acid to detect PKU and for a number of other excreted substances caused by metabolic defects. The urine of a PKU case will contain phenylpyruvic acid that will turn dark green within a few seconds upon addition of a drop of 10% ferric chloride solution and the negative, normal case, is clear yellow.[1,7] The filter paper testing procedure offered the advantage of testing urine that may be several days old. Since phenylpyruvic acid usually does not appear in the newborn's urine until weeks after birth, the wet diaper and filter paper tests using a drop of ferric chloride solution were not reliable until the newborn was a few weeks old. Because of this delay, many newborns were not appropriately tested for PKU. This assay variable created the policy that no babies or child should be started on the long-term program of a low-Phe diet without a liquid blood sample analysis for confirmation of the disorder.

In the early 1960s, the main value and indication for the liquid blood serum Phe measurements by enzymatic or fluorometric methods were for final confirmation of the diagnosis of PKU in situations of presumptive positive results by urine tests.[12-15] However, these methods were too slow and difficult for routine screening for PKU. All the experience and evidence up to this historical testing stage continuously pointed to the need for a blood screen or test that could be performed while the newborn was still in the hospital nursery.

The Guthrie Test: Population Screening Begins

The Guthrie test, PKU test, or bacterial inhibition assay (BIA) for PKU is the name applied to the first usable NBS technology for population-based testing in the United States.[1,16,17] The 1963 technology was based on a dose response of specialized bacteria to the presence of metabolites (Phe) diffusing from the dried blood spot into the growth media (Figure 4i). The instruments required to detect changes are as simple as a ruler to measure a growth zone diameter. Bacterial inhibition assay is a simple, low-cost technology and requires no complex instrumentation or equipment. Much of the technology and experience in its application was already familiar to public health laboratories and the microbiologists who were staffed in them. Bacterial inhibition assay was the most appropriate technique because of its negligible cost, its application to blood samples dried on filter paper with batch analysis, and the utilization of the existing microbiologist who could be responsible for the

NBS testing. The BIA overcame the detection concerns with urine caused by delayed production of phenylpyruvic acid and allowed for earlier testing with excellent predictive results for PKU.

Specifically, the BIA technology is a single-analyte detection device for Phe. It requires specialized bacteria that will not grow in an agar culture medium in the absence of Phe and contains an inhibitor of Phe that is overcome by increased Phe levels. The blood spot punches are placed on the surface of the agar gel culture medium in an array similar to that of a microtiter plate and incubated. Blood spots with known concentrations of Phe (controls) and the unknown patients (cases) provided the necessary Phe for the bacterial growth. The observed growth is measured by the diameter of the growth zone developed around the Phe diffusion (elution) from the DBS punch. The measured diameter is directly proportional to the Phe concentration. This BIA technique was also adaptable to single tests of other metabolites for the detection of different amino acid disorders.

The primary method of choice in most laboratories for PKU detection was the BIA method for over 2 decades. The problem with a biological assay is that anything that can impact the bacterial growth could cause an inaccurate test result. Antibiotics present in blood were an obvious problem. Further, the precise nature of measuring a growth zone (diameter) was relatively poor, but its selectivity was excellent. Given the poor precision, most samples were collected no earlier than 2 days after birth, with most on day 3 or later to enable further enhancement of the PKU disorder, which would yield higher Phe concentrations and thus reduce the potential for false-negative test results. The BIA over time established itself as an excellent screening tool and enhanced the status of newborn by screening, saving thousands of babies from a life with an intellectual development disorder.

Embracing Modern Analytical Technology

New methods that are more modern technologically and exceed the current analytical standard will often follow a prescribed path of acceptance, which is often used as a confirmatory test. The next stages include piloting and implementing the test. As experience and technology improves, there is the desire to upgrade the existing assays to something better, more efficient, precise, accurate, and ultimately cheaper. Because the BIA test was only qualitative, other techniques, when developed, were considered as confirmatory tests. A fluorometric assay for PKU detection described by McCaman and Robins[14] in 1962 relied on fluorescence emitted from Phe after incubation with fluorophores (Figure 4iii). Once validated, these methods typically are automated so they have higher throughput and as such reduce the cost per sample. Kirkman et al[18] reported on the use of a fluorometric autoanalyzer with a capability of analyzing more than 80 000 samples per year. Although BIA remained entrenched in many NBS laboratories, and more technologically advanced methods such as high-performance liquid chromatography (LC) were available (Figure 4iv), the
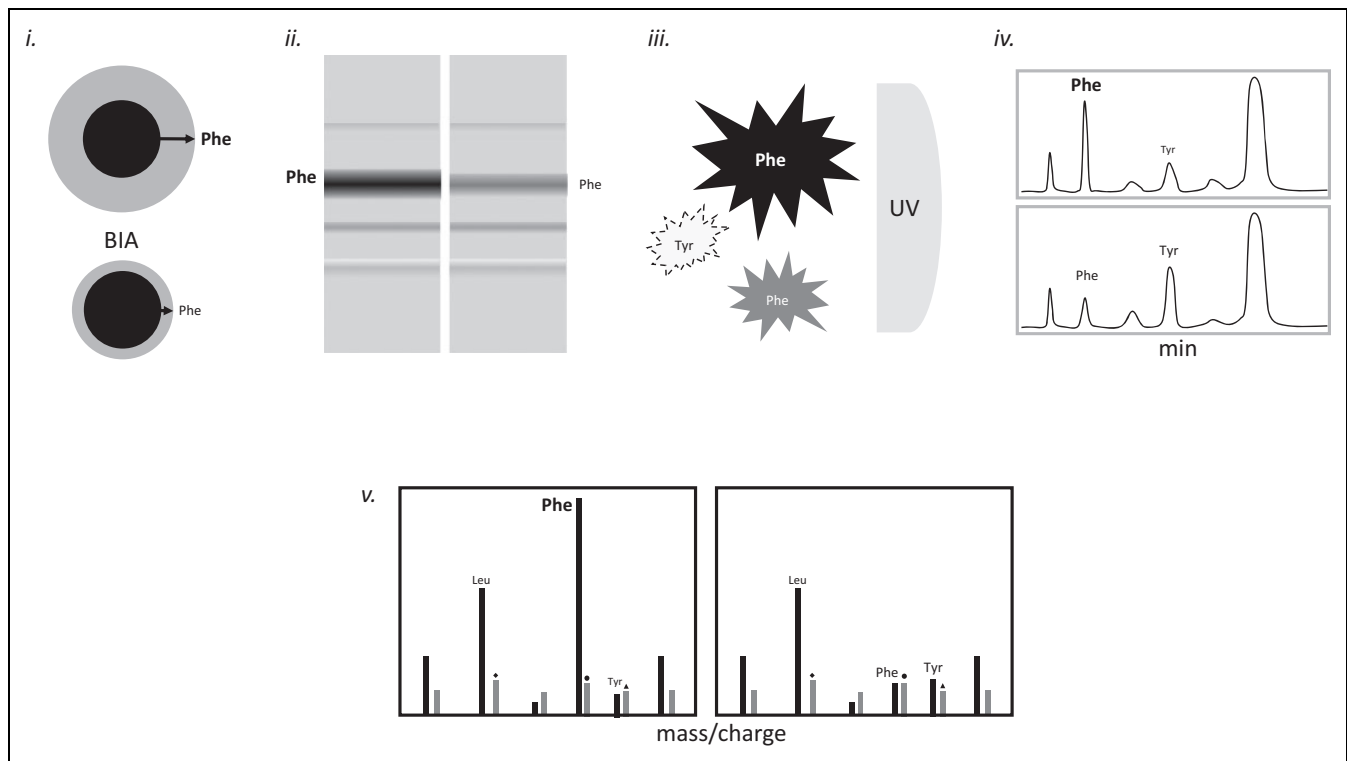
Figure 4. Technology evolution. i, Guthrie BIA test showing growth zone (light gray) is larger with higher Phe concentrations (top illustration) than normal Phe (lower illustration). ii, Thin layer chromatography for Phe showing bands that identify Phe. Higher concentrations (left panel) indicate higher Phe concentrations. iii, Fluorometric Phe analysis showing higher Phe concentrations fluoresce more than normal Phe concentrations. Tyrosine fluoresces at the wavelengths of Phe to a partial extent causing a variable bias based on the concentration of Phe. iv, HPLC chromatographic separation of amino acids show Phe at higher concentrations (top panel) as compared to normal Phe (lower panel). Other amino acids are separated and measured independently. v, Mass of Phe and Tyr (as well as other amino acids, ie, Leu) using MS/MS with isotope-labeled internal standards for each amino acids (*). Higher Phe shows higher peak intensity (left panel) relative to internal standard (*). BIA indicates bacterial inhibition assay; HPLC, high performance liquid chromatography; Leu, leucine; MS/MS, tandem mass spectrometry; Phe, phenylalanine; Tyr, tyrosine.

simplicity and high throughput of automated fluorescence technology led to its widespread adoption for the screening of Phe. The precision of this test was high, but its accuracy was still less than desirable because of the nonselective nature of fluorescence in similar compounds. Other amino acids extracted from the DBS could also fluoresce at wavelengths that would contribute to the signal for Phe and contribute to a higher baseline for the measurement of Phe, thus contributing a bias in the analysis.

Chromatographic methods for measuring Phe and other metabolites of interest in NBS have their roots in paper (thin layer) chromatography[19-21] (Figure 4ii). Gas chromatography (GC) methods[22,23] for Phe and other amino acid were available for a few decades, yet only recently become utilized in some laboratories but are still too labor intensive and time consuming for widespread adoption and practicality. Liquid chromatography methods in NBS (Figure 4iii) didn't have the time to evolve for application because the same system was being developed to flow into a mass spectrometer (MS) for a much more powerful detection system as well as similar issues impacting GC/MS, time, and practicality. Ultimately, the use of LC systems but without chromatography coupled to a MS/

MS became the most powerful, efficient screening systems for Phe and other amino acids (Figure 4v).

Three decades of screening witnessed the evolution from a biological assay to a classical clinical chemistry assay for Phe (Figure 4). During this time interval, up to a half-dozen disorders were added to the NBS panels in many states. The next decade witnessed a "sea-of-change" in screening[24] where not only did the technology for analyzing Phe and other amino acids become more accurate, sensitive, and precise, but an entirely new set of metabolites were also added to the screening panel, ultimately producing an analysis of more than 32 disorders in a single run using one 3.2-mm DBS punch, 1 sample preparation, and 1 analysis by MS/MS.[25]

Capillary GC coupled with MS detection (GC/MS) is still the gold standard in metabolic research and clinical chemistry analysis.[26] The problem with GC/MS has always been the complex sample workup, long run times, and metabolites excreted in urine, which is problematic for some metabolic disorders and NBS sample collection systems. In addition to the logistical problems associated with GC/MS in urine, the analysis of more polar, hydrophilic, larger, complex, or unstable metabolites was difficult by GC/MS if not impossible.
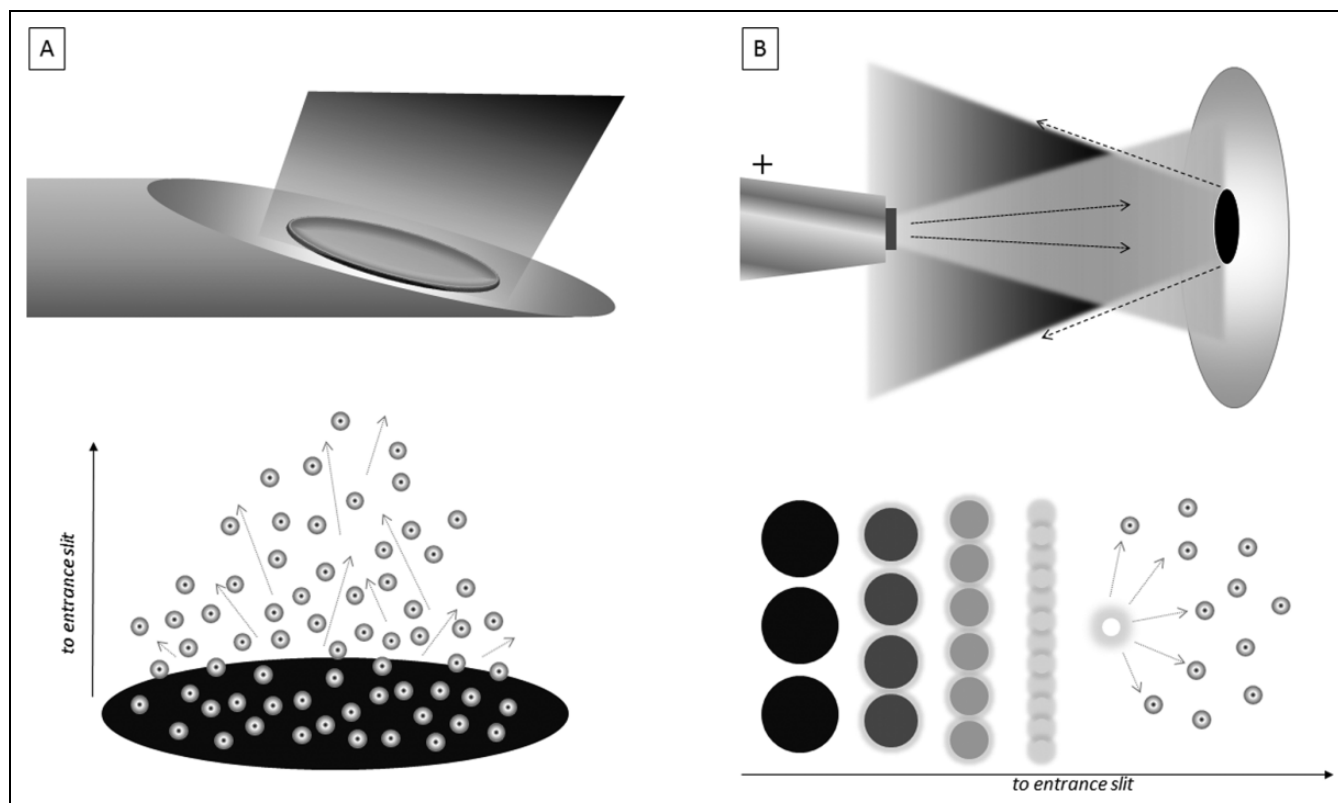
Figure 5. Liquid chromatography ionization methods. A, Liquid secondary ionization illustrated (fast atom or fast ion bombardment). First, MS/MS surface ionization technique utilized. Metabolites dissolved in a glycerol matrix were placed at the end of a stainless steel probe. Ionization occurred when the energy from the ion or atom beam struck the surface, imparting energy and transfer of a proton from the matrix to the analyte (ie, Phe). B, Electrospray ionization illustrated. Electrospray used principles of charge aerosols emerging from a flowing stream containing compounds of interest (ie, Phe). Curtain gases emerging from the mass spectrometer source and the probe tip charged at thousands of volts caused droplets to become charge while the solvent evaporated. When the charge became high enough relative to the droplet size that stabilized it, it exploded and each protonated molecule enter the gas phase and entered the MS for analysis. Note: The MS/MS analysis for both methods was the same. MS indicates mass spectrometer; MS/MS, tandem mass spectrometry; Phe, phenylalanine.

Carnitine and acylcarnitines were found to be important biomolecules in fatty acid metabolism in the 1980s, but the analysis was problematic[27] because these compounds could not be analyzed intact by GC/MS because of their permanent positive charge. Fortunately, improved LC/MS technology (liquid secondary ionization mass spectrometry techniques), more commonly known as fast atom bombardment (FAB) or fast ion bombardment (FIB), was developed. These ionization methods revolutionized a whole host of new analytical approaches from proteins to peptides, carbohydrates, lipids, and small molecules such as amino acids and acylcarnitines.[28-30] Gentle ionization techniques such as protonation cause very little if any fragmentation unlike electron impact ionization is commonly used with GC/MS. In order to generate fragments, MS/MS systems that have a collision chamber for that purpose were used. With this collision chamber, intact protonated molecular ions could be fragmented in a very controlled and specific fashion to fragment (product) ions, which ultimately enabled highly selective analysis using MS/MS (see below). Figure 5A, left panel, illustrates the ionization process of FAB or FIB.

Electrospray ionization succeeded FAB/FIB because both improved the ionization process while also having the ability

to automate the analysis using a flowing liquid stream.[31,32] The process of electrospray ionization is illustrated in Figure 5B (right panel). With electrospray, protonation occurs by charging droplets from an aerosol while also drying the excess solvent gases that form curtains and sheaths, which prevents excess solvent (neutral molecules) from entering the mass separation chambers. Using MS/MS for amino acids and acylcarnitines, there is no chromatography, but rather a simple flowing stream of a plug injection. Separation of compounds is performed by MS/MS rather than a separation column and as such requires no time separation by chromatography. The process of injection without chromatography and separation by MS/MS is known as flow injection MS/MS or FIA MS/MS.[33]

The mass analyzer of a typical MS system separates molecules based on their mass to charge (*m/z*) ratios. Tandem mass spectrometry known as tandem-in-space instruments have 2 mass analyzers separated by a collision chamber. The collision chamber is used to fragment molecules entering it from the first mass spectrometer (MS1). The second mass spectrometer (MS2) receives these ionized fragments where a second *m/z* analysis is performed. Using software native to these instruments, 4 types of scan functions (MS analysis types) can be

performed that provide 4 types of analysis. For example, an MS/MS system can be used as a typical MS-only instrument by simply turning off the collision cell. The system's primary use is in research, and the analysis is a scan known as product ion scan where MS1 focuses on a single protonated molecule of interest, fragmentation occurs, and then MS2 charts the spectra of all the fragments. Two, less commonly used scan functions known as precursor ion and neutral loss scans are the 2 most important analyses in the MS/MS analysis of acylcarnitines and amino acids. A precursor ion function scans all protonated molecular ions in MS1, fragments each of these in the collision cell, but focuses only on a single fragment in the MS2. This second fragment is usually the common fragment of the specific group of compounds that you wish to analyze. For example, all acylcarnitines share a common product ion at m/z 85 (whether derivatized as a butyl ester or as a free nonderivatized acid). The resultant spectra appear as the original molecule analyzed in MS1 but had to have an ion generated at m/z 85, essentially an m/z 85 scan (which is specific to acylcarnitines). No chromatography is needed to sort out the different acylcarnitines, as this is achieved by differences in mass. Amino acids generate a common uncharged fragment molecule, which has a mass of 46 or 102 if it is formate (nonderivatized) or butyl formate (butyl ester derivative). Using a neutral loss scan of either 42 or 102, MS1 and MS2 analyses are offset by the mass differences (which is affected by the loss of the neutral molecule). The result is similar to acylcarnitines in their selectivity but very specific to α-amino acids only. Both analysis can be combined during a sample acquisition to produce what is commonly referred to as the MS/MS NBS analysis.[34]

Substantial clinical research and understanding of metabolic diseases came from developments in the late 1980s and the utilization of FAB, MS, and MS/MS for plasma analysis. Like PKU, many of these disorders benefit from early detection and treatment for the prevention of devastating disability or death. In the late 1980s and early 1990s, before PKU analysis by MS/MS was developed, clinical research found evidence that strongly suggested that NBS of fatty acid disorders (β-oxidation disorders) such as medium chain acyl CoA dehydrogenase (MCAD) deficiency would prevent disability or death. Prevention was mostly a recognition that the disease was present and as a result an avoidance of fasting was the treatment. To understand whether the disorder was frequent in the population and the cost of screening development could be justified, a molecular analysis as a survey for the most common mutation of MCAD was performed. It was determined that the combined frequency of all MCAD deficiency mutations was similar to that for PKU.[35] At the time, although a clinical analysis of plasma using an acylcarnitine profile for MCAD deficiency existed, there was no method to analyze this disorder in blood, let alone the preferred sample type in NBS, the DBS. Molecular studies and postmortem screening data advocated that a screening method should be developed for MCAD deficiency.[36,37]

It was becoming evident that MCAD deficiency met most of the requirements for NBS.[38] The challenges were that in addition to developing a new method based on DBS analysis (a blood profile vs plasma), it would be a prohibitively expensive method for a single disorder using the MS/MS system (which was still large complex expensive instruments at that time). Second, validations of new methods usually require comparison with an older established method. No other method with DBS existed for MCAD deficiency. The solution was to analyze a compound currently performed in routine NBS that was similar in chemical class and a likely candidate for MS/MS analysis. The obvious candidate metabolite was Phe, a metabolite that is commonly analyzed by MS technology.

Accidental Discovery

Developing the analysis of Phe from DBS was relatively straightforward for MS/MS application. Organic extraction (methanol), which can recover metabolites like Phe from the blood matrix without reconstituting proteins or salts, was the first important development. Choosing a different derivatization method was also necessary (butyl esterification vs methyl esterification) to remove a potential interference for Phe. Finally, finding a scan function unique to Phe (vs that for acylcarnitines) led to the discovery that what was common to Phe may be also common to many other compounds. How was this discovered? The observation and use of a standard mix of amino acids had many unknown peaks in the spectra that differed by 14 and 16 Da. Identification of these compounds disclosed amino acids such as Tyr, leucine (Leu), and methionine. All of these compounds were shown to share a common product molecule that was lost on fragmentation leading to the development of an amino acid profile rather than a Phe-only analysis. Incorporating the acylcarnitine profile was easy to achieve by changing derivatization from methyl to butyl esterification.

A Better Standard

Mass spectrometer technology is often called the gold standard of clinical and pharmaceutical analysis. The reason for this is a concept in bodily fluid analysis termed isotope dilution. This is a technique where an isotopically labeled internal standard, which is chemically identical to the target analyte but differs simply by mass (based on its enrichment), is added to the biofluid to be analyzed, that is, plasma, blood, or urine. In NBS, the biomatrix is dried and therefore presents a challenge to classical isotope dilution mass spectrometry (IDMS) techniques. Stable isotope standards are prepared in the extraction solvent where the DBS is immersed. As the target analytes are extracted into the solvent, they become mixed with their stable isotope analogues in a manner similar to isotope dilution. The remainder of the sample preparation and analysis is identical to the IDMS technique. Because the extraction efficiency is a key variable in the analysis and requires a correction factor, but all other aspects are identical to IDMS, it is termed pseudo-IDMS.

The evolving process and design of internal standards for the analysis of NBS metabolites are little known or discussed but were important aspects of the ultimate success of MS/MS.
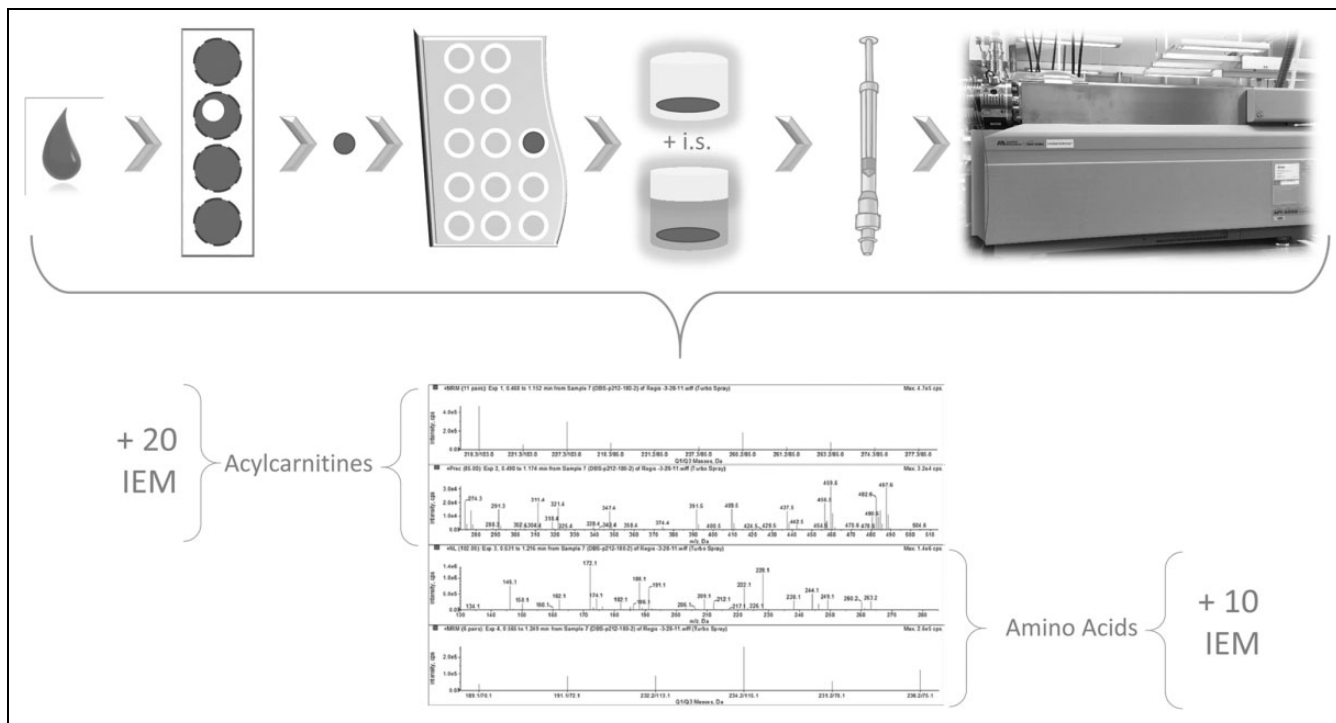
Figure 6. One drop to many disorders. Illustration of the steps from collecting a drop of blood through disorders detected.

Each and every internal standard is carefully designed for enrichment, that is, choice of isotope (deuterium, carbon-13, nitrogen-15), number of isotopes (minimum of 2), location of the isotopes to match the fragmentation pattern and scan functions, and stability of each isotope. Today, at least 20 isotopically labeled standards are present in each MS/MS analysis of a DBS for amino acid and acylcarnitine measurements.

A Better Method for the Detection of PKU

New technology is important because it offers new approaches to solving problems. There is often a strong desire among laboratory professionals to keep their laboratory updated. But new doesn't always mean better in terms of technology. What counts is that any change or improvement that may include new technology makes the results of the procedure better. This process to provide the evidence through clinical trials and or research is slow, and it still takes time for the acceptance. Tandem mass spectrometry for amino acid and acylcarnitine analysis was used in only 2 or 3 centers of excellence worldwide in the early 1990s. It was adopted by a commercial screening laboratory in the mid 1990s and only in the next century was fully embraced by public health laboratories worldwide. It took 10 to 20 years to optimize, refine, and improve the analysis by MS/MS and to provide summary data and experience to show that indeed this was the newest and best practice for screening many inherited metabolic disorders such as PKU. The story of the validation of PKU screening by MS/MS is the same for each and every successive metabolite added to the analysis panel. That journey of discovery and validation

is an excellent model for future metabolic screening developments. Figure 6 shows the components of MS/MS screening for PKU and other metabolic disorders from the dried blood sample through the resulting mass spectra.

In 1993, the first validation article for the MS/MS detection of PKU was published.[39] As described in this article, the original MS/MS analysis measured not only Phe but also Tyr and other amino acids. Each key metabolite had its own unique stable isotopic internal standard and therefore produced a unique concentration value based on its matched standard. The concentration of each metabolite could be obtained, providing an opportunity to use informatics to calculate concentration ratios (molar ratios) that actually have a basis in metabolism for enhanced analytical value. Because Phe is the substrate and Tyr is the product of the targeted potentially defective enzyme, PAH, the ratio of these 2 compounds may be a better indicator of enzyme function or dysfunction. In addition, any errors in the blood spot sampling or characteristics of the blood spot volume variation due to hematocrit or spot size would be minimized in a ratio value because theoretically it would similarly affect each compound equally. A sentinel paper in 1993 demonstrated the potential for a reduction of false results in screening with the use of ratios.

An important study examined the improvements made using MS/MS for screening PKU. Dried blood spot collected from early discharge newborns (sample collection is less than 24 hours in this group) that were previously analyzed by fluorometry and categorized as initially negative (93), confirmed positive (19), presumptive positive but confirmed negative (91) by fluorometric analysis was reanalyzed by MS/MS. Of

Table 1. Analytical Features of Methods Used to Detect PKU.

Metabolite	Method	Characteristics	Response	Risk
Phe only	BIA	Poor precision; low sensitivity	Collect later	False-negative risk for other disorders
	Fluorometry	High bias (interferent)	Raise cutoffs	False-negative risk for PKU
	MS/MS	High selectivity; sensitivity	Lower cutoffs; collect earlier	Reduce false negative and false-positive risk
Phe and Tyr	Fluorometry	Two methods; both high bias	Ratio inaccurate (dual bias)	Expensive without improvement
	MS/MS	One method; higher selectivity	Ratio accurate; collect earlier	Reduce false results with earlier detection
Phe, Tyr, and Leu	MS/MS	One method; verification step	Reduce false results (TPN)	Reduce false results; characterize interferent

Abbreviation: BIA, bacterial inhibition assay; Leu, leucine; MS/MS, tandem mass spectrometry; Phe, phenylalanine; PKU, phenylketonuria; TPN, total parenteral nutrition; Tyr, tyrosine.

the presumptive positive group, 88 of the 91 samples were found to have a completely normal Phe concentration.[40] All the specimens that were previously confirmed positive were positive by MS/MS at a cutoff concentration of 180 μM. Two samples that had a mildly elevated Phe had a normal Phe/Tyr ratio on reanalysis. One sample that had an elevated Phe and borderline Phe/Tyr ratio was presumably a preterm baby since multiple amino acids were elevated. The different results by both methods show that MS/MS has little or no bias in the measurement of Phe because of its high selectivity. Adding other amino acids to the analysis such as Tyr and Leu enable calculations of Phe/Tyr and Phe/Leu ratios, which further increase the specificity of the analysis. Recent studies in preterm infants have shown that many elevated Phe results are due to total parenteral nutrition (TPN) contamination of a blood sample.[41,42] Total parenteral nutrition typically contains high concentrations of various amino acids such as Leu and Phe but does not have high concentrations of Tyr (it is present as *N*-acetyl Tyr). By using both Phe/Tyr and Phe/Leu, a potential false-positive result can be reduced further if it is confirmed that the newborn was on TPN and there are elevations of dextrose markers in the acylcarnitine profile (which strongly suggest a contamination of blood with TPN amino acid intravenous solutions). Finally, MS/MS offers greater sensitivity than the other methods. In the study of DBS samples from early hospital-discharged newborns described above, it was shown that lower cutoff values and the Phe/Tyr ratio could be used to detect a positive PKU sample at as early as 6 to 12 hours after birth. Table 1 summarizes methods for the detection of PKU from single analyte through multiple metabolite methods.

Beyond PKU: Screening for Inherited and Acquired Metabolic Disorders

Tandem mass spectrometry is more than a screen for PKU. It is more than an analysis of Phe. It is a tool to measure the concentration of dozens of amino acids and acylcarnitines in a single analytical run from a tiny punch from a DBS. The panel of metabolites is used to detect more than 30 inherited disorders of metabolism (depending on how you count and differentiate them) in NBS.[43] It is fully automated and utilizes state-of-the-art stable isotopic internal standards and MSs. It utilizes fully automated data processing and interpretation systems. The essential method is so robust that a few different sample preparation or data acquisition methods can be chosen without altering the final results. It detects abnormalities that may be caused by acquired disorders or iatrogenic causes as observed for preterm infants and TPN exposure. Tandem mass spectrometry has been used to monitor the treatment of metabolic disorders if blood spots are chosen as the required matrix. However, liquid plasma is the preferred choice by most metabolic specialist in part because of the accuracy of measurement in a liquid matrix whether analyzed by MS/MS or other techniques. One study compared MS/MS analysis of Phe and Tyr in DBS with ion exchange of plasma samples for therapeutic monitoring of PKU.[44] As expected, the precision and accuracy of a DBS analysis was poorer than the analysis using liquid plasma. Although the study suggested it was a comparison of MS/MS versus ion exchange chromatography, it was a false comparison as it is well known that MS/MS analysis using IDMS with or without chromatography is considered the gold standard in clinical and pharmaceutical analysis. Care should be taken in differentiating the goals of an analysis and the desired precision, accuracy, and matrix choice needed.

For example, as a screening technique, MS/MS has been used for the postmortem analysis using DBS and other biological matrices such as dried plasma or dried bile samples. Tandem mass spectrometry is not just for NBS, as it has been used to analyze dried plasma spots for free and total carnitine in adult patients and in dialysis patients.[45] The method has potential to analyze amino acids and acylcarnitines in other fluids such as cerebrospinal fluid, saliva, tears, and even human breast milk for nutritional content. New methods that utilize another punch from the DBS sample were developed to confirm the presence of disease or metabolites and reduce false-result rates. Tandem mass spectrometry can work in parallel with molecular analysis such that one of these 2 analyses can serve as the confirmation of the other, phenotype and genotype. The MS/MS method can be adapted in diagnostic laboratories for confirmatory testing using plasma, urine, cell culture, and other matrices. Tandem mass spectrometry is the ultimate, flexible tool for analytical measurements regardless of the matrix and this is why the public health community witnessed the rapid expansion of NBS over the last decade and it continues as the method of choice for a variety of new disorders.

Summary and Concluding Remarks

The analysis of Phe for PKU detection across time was chosen as the model system to describe how advancements in MS/MS technology contributed to improved laboratory performance. Tandem mass spectrometry has made NBS programs better and has improved the health outcomes of the affected newborn through improvements in accurate early detection. A quicker, accurate result enables physician to arrive at a diagnosis quickly and assuredly so that treatment can be administered sooner and more effectively. More than 50 years of NBS for PKU have shown that technology can continue to improve and expand screening. Improvements were made because of a crucial health problem requiring a new creative approach to solve. Tandem mass spectrometry is a state-of-the-art technology that is still advanced but no longer new. It is the method of choice in almost all modern NBS facilities. Because MS/MS is an extremely versatile instrument, it continues to grow in its application for other metabolites and disorders such as lysosomal storage diseases. The NBS programs are taking advantage of this multiple purpose system to perform new assays for confirmatory testing (second-tier testing). The use of MS/MS will continue to expand, even with the anticipated introduction and expansion of molecular screening methods into NBS programs. The 50-year story of PKU detection illustrates that the solution to NBS problems is an early, more sensitive, accurate, and selective detection technology with far fewer false-positive results. Thereby, the clinician has better tools and greater confidence in providing affected newborns with the best possible health outcomes for coping with a lifelong genetic disorder. In the next 50 years of NBS, many new technologies and approaches will be introduced to enhance the detection and treatment of newborns affected by PKU and other inborn errors of metabolism.

Author's Note

Declaration of Conflicting Interests

Funding

References

1. Guthrie R, Susi A. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics*. 1963;32:338-343.
2. Garrod A. The Croonian Lectures on Inborn Errors of Metabolism. *Lancet*. 1908;172(4429):73-79.
3. Recommended Uniform Screening Panel, Advisory committee on heritable disorders in newborns and children. 2006. Web site. http://www.hrsa.gov/advisorycommittees/mchbadvisory/herita bledisorders/recommendedpanel/index.html. Updated March 2015. Accessed September 24, 2016.
4. Lee MS, Li W. *Dried Blood Spots: Applications and Techniques*. Hoboken, New Jersey: Wiley; 2014.
5. Clinical and Laboratory Standards Institute. *Blood Collection on Filter Paper for Newborn Screening Programs; Approved Standard*. CLSI Document NBS01-A6. 6th ed. Wayne, PA: CLSI; 2013.
6. Centerwall SA, Centerwall WR. The discovery of phenylketonuria: the story of a young couple, two retarded children, and a scientist. *Pediatrics*. 2000;105(1 pt 1):89-103.
7. Centerwall WR. Phenylketonuria. *J Am Med Assoc*. 1957;165(4):392.
8. Penrose L, Quastel JH. Metabolic studies in phenylketonuria. *Biochem J*. 1937;31(2):266-274.
9. Centerwall WR, Chinnock RF, Pusavat A. Phenylketonuria: screening programs and testing methods. *Am J Public Health Nations Health*. 1960;50:1667-1677.
10. Berry HK. Procedures for testing urine specimens dried on filter paper. *Clin Chem*. 1959;5:603-608.
11. Berry HK, Sutherland B, Guest GM, Warkany J. Simple method for detection of phenylketonuria. *J Am Med Assoc*. 1958;167(18):2189-2190.
12. La Du BN, Michael PJ. An enzymatic spectrophotometric method for the determination of phenylalanine in blood. *J Lab Clin Med*. 1960;55:491-496.
13. Wright SW. Mass screening for phenylketonuria. *J Pediatr*. 1962;61:651-652.
14. Koch R, Williamson ML, Donnell GN, et al. A cooperative study of two methods for phenylalanine determination: McCaman-Robins fluorimetric and microbiologic inhibition methods. *J Pediatr*. 1966;68(6):905-911.
15. Cechak P, Zidlicky A. Determination of phenylalanine in serum by a modified fluorimetric method in the diagnosis of phenylketonuria. *Acta Univ Carol Med Monogr*. 1977;(73 pt 3):53-55.
16. Guthrie R. Screening for phenylketonuria. *Triangle*. 1969;9(3):104-109.
17. Guthrie R. The origin of newborn screening. *Screening*. 1992;1:5-15.
18. Kirkman HN, Carroll CL, Moore EG, Matheson MS. Fifteen-year experience with screening for phenylketonuria with an automated fluorometric method. *Am J Hum Genet*. 1982;34(5):743-752.
19. Nitschke E. Thin layer chromatography separation and semi-quantitative determination of blood phenylalanine as procedure for the systematic screening of phenylketonuria in newborn infants [in German]. *Z Klin Chem Klin Biochem*. 1968;6(3):208-210.
20. Schon R, Thalhammer O. Routine thin-layer-chromatography in 25000 newborn infants. Comparison with the Guthrie-tests [in German]. *Z Kinderheilkd*. 1971;111(3):223-232.
21. Allen RJ, Frey HJ, Fleming LM, Owings CL. Semiquantitation of leucine, isoleucine, and valine by thin-layer chromatography in management of maple-syrup urine diseases. *Clin Chem*. 1972;18(5):413-416.
22. Deng C, Deng Y, Wang B, Yang X. Gas chromatography-mass spectrometry method for determination of phenylalanine and tyrosine in neonatal blood spots. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2002;780(2):407-413.

23. Deng C, Deng Y. Diagnosis of maple syrup urine disease by determination of L-valine, L-isoleucine, L-leucine and L-phenylalanine in neonatal blood spots by gas chromatography-mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2003; 792(2):261-268.

24. Levy HL. Newborn screening by tandem mass spectrometry: a new era. *Clin Chem*. 1998;44(12):2401-2402.

25. Chace DH, Kalas TA, Naylor EW. Use of tandem mass spectrometry for multianalyte screening of dried blood specimens from newborns. *Clin Chem*. 2003;49(11):1797-1817.

26. Lehotay DC. Chromatographic techniques in inborn errors of metabolism. *Biomed Chromatogr*. 1991;5(3):113-121.

27. Costa CG, Struys EA, Bootsma A, et al. Quantitative analysis of plasma acylcarnitines using gas chromatography chemical ionization mass fragmentography. *J Lipid Res*. 1997;38(1):173-182.

28. Millington DS, Kodo N, Norwood DL, Roe CR. Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. *J Inherit Metab Dis*. 1990;13(3):321-324.

29. Norwood DL, Kodo N, Millington DS. Application of continuous-flow liquid chromatography/fast-atom bombardment mass spectrometry to the analysis of diagnostic acylcarnitines in human urine. *Rapid Commun Mass Spectrom*. 1988;2(12): 269-272.

30. Millington DS, Roe CR, Maltby DA. Application of high resolution fast atom bombardment and constant B/E ratio linked scanning to the identification and analysis of acylcarnitines in metabolic disease. *Biomed Mass Spectrom*. 1984;11(5):236-241.

31. Rashed MS, Ozand PT, Harrison ME, Watkins PJF, Evans S. Electrospray tandem mass spectrometry in the analysis of organic acidemias. *Rapid Commun Mass Spectrom*. 1994;8(1):122-133.

32. Johnson AW, Mills K, Clayton PT. The use of automated electrospray ionization tandem MS for the diagnosis of inborn errors of metabolism from dried blood spots. *Biochem Soc Trans*. 1996; 24(3):932-938.

33. Chace DH, Naylor EW. Expansion of newborn screening programs using automated tandem mass spectrometry. *Ment Retard Dev Disabil Res Rev*. 1999;5:150-154.

34. Chace DH. Mass spectrometry in the clinical laboratory. *Chem Rev*. 2001;101(2):445-477.

35. Ding JH, Bross P, Yang BZ, et al. Genetic heterogeneity in MCAD deficiency: frequency of K329E allele and identification of three additional mutant alleles. *Prog Clin Biol Res*. 1992;375: 479-488.

36. Bennett MJ, Hale DE, Coates PM, Stanley CA. Postmortem recognition of fatty acid oxidation disorders. *Pediatr Pathol*. 1991; 11(3):365-370.

37. Bennett MJ, Rinaldo P, Millington DS, Tanaka K, Yokota I, Coates PM. Medium-chain acyl-CoA dehydrogenase deficiency: postmortem diagnosis in a case of sudden infant death and neonatal diagnosis of an affected sibling. *Pediatr Pathol*. 1991;11(6): 889-895.

38. Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. Geneva, Switzerland: World Health Organization; 1968.

39. Chace DH, Millington DS, Terada N, Kahler SG, Roe CR, Hofman LF. Rapid diagnosis of phenylketonuria by quantitative analysis for phenylalanine and tyrosine in neonatal blood spots by tandem mass spectrometry. *Clin Chem*. 1993;39(1):66-71.

40. Chace DH, Sherwin JE, Hillman SL, Lorey F, Cunningham GC. Use of phenylalanine-to-tyrosine ratio determined by tandem mass spectrometry to improve newborn screening for phenylketonuria of early discharge specimens collected in the first 24 hours. *Clin Chem*. 1998;44(12):2405-2409.

41. Chace DH, De Jesus VR, Lim TH, Hannon WH, Spitzer AR. Tandem mass spectrometric identification of dextrose markers in dried-blood spots from infants receiving total parenteral nutrition. *Clin Chim Acta*. 2010;411(21-22):1806-1816.

42. Chace DH, De Jesus VR, Lim TH, Hannon WH, Clark RH, Spitzer AR. Detection of TPN contamination of dried blood spots used in newborn and metabolic screening and its impact on quantitative measurement of amino acids. *Clin Chim Acta*. 2011; 412(15-16):1385-1390.

43. Sweetman L, Millington DS, Therrell BL, et al. Naming and counting disorders (conditions) included in newborn screening panels. *Pediatrics*. 2006;117(5 pt 2):S308-S314.

44. Stroup BM, Held PK, Williams P, et al. Clinical relevance of the discrepancy in phenylalanine concentrations analyzed using tandem mass spectrometry compared with ion-exchange chromatography in phenylketonuria. *Mol Genet Metab Rep*. 2016;6:21-26.

45. Reuter SE, Evans AM, Faull RJ, Chace DH, Fornasini G. Impact of haemodialysis on individual endogenous plasma acylcarnitine concentrations in end-stage renal disease. *Ann Clin Biochem*. 2005;42(pt 5):387-393.