Introduction of guanidinoacetate methyltransferase deficiency to the Dutch newborn screening program -feasibility assessment

Introduction

Early diagnosis and lifelong treatment of guanidinoacetate methyltransferase deficiency (GAMT-D), starting in the neonatal period, is the key in preventing irreversible brain damage and its consequences, ie. developmental delay, epilepsy, and behavioral problems related to the disease. Newborn screening of GAMT-D provides the most potential tool for this purpose.

Newborn screening of GAMT-D in the Netherlands has been requested by the Dutch Ministry of Health, Welfare, and Sport on 9 July 2015, based on the advice of the Health Council to extend the newborn blood spot screening (NBS) with several newborn disorders.

The implementation of GAMT-D screening in the NBS program has two major challenges. First, no commercial kit is currently available for the screening of GAMT-D and there is no extensive experience of in-house tests for population screening in the Netherlands as first tier tests. Second, the reports from the international laboratories screening GAMT-D have been based on derivatized samples, but the current Dutch NBS first tier program is performed by using underivatized samples. In this project, we investigated the possibilities and potential pitfalls involved in the incorporation of GAMT-D screening in the newborn screening.

The aims of this study were to

- discover the most optimal screening scenario
- provide solutions for the challenges and the risks resulting from the unavailability of a commercial kit for GAMT-D
- develop a validated first tier screening method for GAMT-D

The results of the study and the discussion of the aspects related to the implementation of GAMT-D screening are presented in this report, categorized into sections of each work package of the study and followed by an advice on how to proceed.

April, 2021

Dr. D.E.C. Smith Prof. Dr. G.S. Salomons

Dr. S.L. van der Beek
Prof. Dr. A. Boelen
Drs. E.H.B.M. Dekkers
Dr. T. van Dijk
Dr. A.G.M. van Gorp
Prof. Dr. L. Henneman
Dr. J.M. van de Kamp
Dr. M.F. Mulder
Dr. J.M.E. Männistö
Dr. W.F. Visser
Dr. M. Williams

Table of contents

Abbreviations

AGAT	Argininie:glycine amidino transferase
BSP	Bloodspot punch
Cr	Creatine
FIA-MS/MS	Flow injection analysis tandem mass spectrometry
FPR	False positive rate
GAMT-D	Guanidinoacetate methyltransferase deficiency
GAMT	Guanidinoacetate methyltransferase
GAA	Guanidinoacetate
GC-MS/MS	Gas-chromatography tandem mass spectrometry
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LOQ	Limit of quantification
Metabolic pediatrician	Pediatrician specialized / working on metabolic diseases
MMA	Methylmalonic acedemia
MRI/MRS	Magnetic resonance imaging/spectroscopy
MS/MS	Tandem mass spectrometry
NBS	Newborn screening
PA	Propionic acidemia
SAM	S-adenosylmethionine
UPLC-MS/MS	Ultra Performance Liquid Chromatography tandem mass spectrometry

1 GAMT-D screening scenarios -a literature review

1.1 GAMT-D

Guanidinoacetate methyltransferase deficiency (GAMT-D, MIM 601240, ORPHA 382) is an autosomal recessive disorder of creatine biosynthesis causing developmental delay / intellectual disability and the almost absence of cerebral creatine concentration. The biochemical hallmarks of GAMT deficiency in body fluids are increased levels of guanidinoacetic acid (GAA) and reduced levels of creatine (Cr). Treatment consists of high doses Cr to replenish cerebral Cr deficiency, ornithine supplementation and an arginine restricted diet. Normal neurodevelopmental outcome has been reported only in patients treated from the neonatal period, highlighting the importance of early treatment.

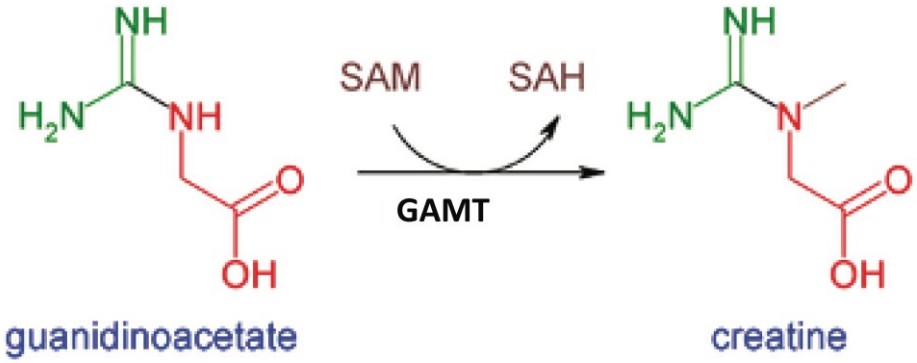


Figure 1.1. Guanidinoacetate methyltransferase reaction. Guanidinoacetate methyltransferase (GAMT, EC 2.1.1.2) is an enzyme which facilitates the formation of creatine by donating the methyl group of S-adenosylmehtionine (SAM) to guanidinoacetate.

1.2 Incidence of GAMT-D

Determining the true incidence in ultra-rare diseases as GAMT-D is a common problem, since the contributing alleles themselves are extremely rare. This causes particularily methodological challenges, e.g. need for especially large population studies or long follow-up time. Only few studies have estimated the incidence of GAMT-D based on a small number of diagnosed patients or calculated mutation carrier frequence.

An estimated incidence based on the number of clinically diagnosed patients was 1:114,000 newborns in Utah, U.S. (Viau et al., 2013). In the Netherlands, the estimated incidence was 1:250,000 newborns in a study conducted in AmsterdamUMC using *GAMT* gene sequencing in a population of 500 newborns giving a carrier frequency of 1:250 (Mercimek-Mahmutoglu et al. 2016).

Recently, the first GAMT-D screening positive infant was identified in Utah, U.S. after the screening of 275,000 babies during 5.5 years, but until more screen positive babies will be identified by NBS, the true incidence in the population remains unknown. (Utah

Department of Health, 2020). With respect to the challenges in determining the incidence, it is not that surprising that the NBS has not detected any true positives among over 770,000 screened infants in Australia (Pitt et al. 2014), although the incidence may even largely vary between different areas.

In the Netherlands, altogether 20 GAMT-D patients from 16 families have been diagnosed by biochemical and molecular analysis at the Metabolic Laboratory in AmsterdamUMC. Of these, two patients were born in 1993 and two in 2004 (12 years period before any laboratory had started NBS for GAMT-D).

As a conclusion, the *true* incidence of GAMT-D is not known and it can only be determined after significantly longer follow-up time compared to the published studies so far and by implementation of NBS. The incidence in the Netherlands is expected to be similar to other countries and 0-2 new GAMT-D patients would be expected per year in NBS (RIVM Rapport 2017-0041).

1.3 Previous studies on GAMT-D screening

Since the early 2000s, several newborn screening programs for GAMT-D have been conducted worldwide: State of Victoria, Australia (commenced in 2002); British Colombia, Canada (continued after the pilot study 2012-2015); State of Utah, U.S. (routine NBS started in 2015); and more recently two other states in the U.S.

Several groups have shared their experience of the implementation and performance of their pilot studies or screening programs in peer reviewed scientific publications (Table 1.1). The reports included altogether approximately 1 million children (range, 3,000-770,000). The studies were based either on ongoing NBS for GAMT-D (Pitt et al. 2014, Sinclair et al. 2016) or on development of an optimal NBS method for GAMT-D (results of Sweetman and Ashcraft 2002 presented in Mercimek-Mahmutoglu et al. 2012, Mercimek-Mahmutoglu et al. 2016, Pasquali et al. 2014).

Methods. All previous studies used in-house methods to measure GAA. One group measured also Cr in order to evaluate whether adding Cr to GAA would lower the false positive rate (FPR). None of the laboratories had experience using a CE certified kit, since these are still not available for GAMT-D. The cut-off range for GAA varied from 2.44-6.0 µmol/L corresponding to 99.0-99.9th percentile of the reference population.

Notably, all laboratories butylated their samples prior to the first-tier FIA-MS/MS assay. This method differs from the current first-tier metabolic screening procedures by FIA-MS/MS in the Netherlands using underivatized samples.

Of note, plasma and urine test for GAA has been previously used in the diagnosis of GAMT-D, but several groups (Viau et al 2013; Bodamer et al 2009; Schulze 2006) have found that the urine test may give false negative results, when taken shortly after birth.

Screening scenarios. A two- or three-tier screening procedure was used in all studies to control the relatively high FPR in the initial screening. In the first-tier, GAA (and Cr) was measured by FIA-MS/MS. This was usually integrated as a part of the already existing amino acid and acylcarnitine assay. In the second-tier, the compound(s) measured in the first-tier were re-analysed by more sensitive (UP)LC-MS/MS or GC-MS/MS assays. *GAMT* gene sequencing or urine analysis was used as the third-tier test.

Some studies performed the second-tier test from a second punch of NBS blood spot, but the more recent studies used the original punch for both first- and second-tier tests. The high methanol content was evaporated and replaced by a more aqueous and thus compatible sample for LC-MS/MS.

In addition to these reports, Mercimek-Mahmutoglu et al. (2016) studied primarily the incidence of GAMT-D by sequencing of *GAMT* in a cohort of 500 Dutch neonates, but they also reported the results of two-tier screening method including GAA and Cr analysis using LC-MS/MS and GC-MS/MS.

Results. *True positives.* In December 2020, Utah U.S. reported the world´s first true positive sample found in their GAMT-D newborn screening after the screening of 250,000 samples.

False positives. In the first-tier measurement of GAA by FIA-MS/MS, the FPRs ranged between 0.02% and 0.58% (Table 1.1). In the second-tier test analyzing the same compound by more sensitive GC-MS/MS, (UP)LC-MS/MS, the rate of the flagged samples decreased to 0.0-0.002%, referring to 0-3 flagged samples in each study. Subsequently, two groups performed DNA analysis either as a third-tier test or as a follow-up diagnostic test, and this reduced the number of false positives to zero. Pasquali et al. (2014) observed that the combination of a 99.5th percentile cut-off for GAA (2.44 umol/L) and 99.9th percentile cut-off for GAA/Cr ratio (0.011) resulted in more acceptable 0.08% false-positive rate.

False negatives. The studies have not identified any false negative results. Three laboratories are also the reference laboratories for their regions performing NBS of GAMT-D and they have not encountered any GAMT-D patients since the introduction of GAMT-D screening (follow-up time up to 11 years).

Table 1.1 Pilot studies and other study reports on GAMT-D screening methods.

Study	Location	In-house or CE certified		Method	Screened metabolites	*n*	GAA cut-off			False positives, n (%)	True positives (during)		False negatives (during)		Problems encountered
							umol/L	percentile							
Sinclair et al. 2016 (PMID: 27233226)	Canada	In-house	I	FIA-MS/MS	GAA	135,37	6.0 (18 mo); 3.5 (later)	99.9 (18 mo); 99.5 (later)	I	259 (0.19%)	0	(3 y)	NA		<1500g babies had isomeric interference in tier 1
			II	LC-MS/MS	GAA				II	3 (0.002%)					
			III	DNA analysis	-				III	0					
Pasquali et al. 2014 (PMID: 24276113)	U.S. (Utah)	In-house	I	FIA-MS/MS	GAA, GAA/Cr ratio	9,288 [a]	2.44	99.5	I	54 (0.58%)	0	(NA)	0	(5 y)	NA
			II	UPLC-MS/MS	GAA, GAA/Cr ratio				II	0					
Pitt et al. 2014 (PMID: 24477282)	Australia	In-house	I	FIA-MS/MS	GAA	771,35	5.0	NA	I	127 (0.02%)	0	(11 y)	NA		Premature babies had GAA disturbance in tier 1 --> checked with LC-MS/MS
			II	LC-MS/MS	GAA				II	3 (0.0004%)					
			III	LC-MS/MS	Urine GAA, Cr, Crn				III	0					
Mercimek-Mahmutoglu et al. 2012 (PMID: 23031365)	Canada	In-house	I	FIA-MS/MS	GAA	2,95	3,12	99.0	I	4 (0.1%)	0	(4 mo)	NA		NA
			II	UPLC-MS/MS	GAA				II	0					
			III	2 *GAMT* variants	-				III	0					
Sweetman&Ashcraft, presented in Merc.-Mahm. et al. 2012 (PMID: 23031365)	U.S. (Texas)	In-house	I	FIA-MS/MS	GAA	19,293	2.94	NA	I [c]	60 (0.5%)	0	(3 y 8 mo)	0	(4 y) [d]	NA
			II	UPLC-MS/MS	GAA		4.0-10.0 [b]		II [c]	0					

GAA, guanidinoaceate; Cr, creatine; Crn, creatinine; MS/MS, tandem mass spectromy; FIA-, flow injection analysis; LC-, liquid chromatography; UPLC-, ultra performance liquid

a) <7 days old, 4691; >7 days old, 4597

b) After the first year: value 4-10 umol/L resulted in second sample and if the value was >10 umol/L, second-tier test was performed

c) After the first year

d) Four years from the first sample to 1 year from the latest sample

Technical challenges encountered in GAMT-D pilot studies. The main challenge observed in the GAMT-D NBS studies was the high FPR in the first-tier FIA-MS/MS analysis. This was due to interfering compounds (an isomeric interference) of yet unknown origin. The interference was especially common in premature and low birth weight newborns.

Sinclair et al. (2016) found that the interference was especially common in babies with very low birth weight infants (<1500 g), accounting for nearly half of the samples flagged for second-tier test. In their study, the interference was observed only in their repeat NBS samples, which were routinely taken because of the high risk of false negative congenital hypothyroidism screening in newborns with low birthweight. Hence, it was presumed to be an exogenous compound from therapeutic interventions.

Discussion: As previously mentioned, the samples in the previous studies were butylated (i.e. adding a butyl group to the original molecule) prior to measurement by MS/MS. The advantage of butylation is that it creates primarily one intense fragment for detection by MS/MS and hence, improves sensitivity. The disadvantages of butylation are that it is more time consuming and that the group that will be removed from the molecule during fragmentation is the butyl group, leaving the original molecule intact. This makes interference with other molecules with identical masses (especially with low mass molecules like GAA and Cr) more likely. When molecules are not derivatized, the leaving fragment will be more specific and hence increases specificity. In the Netherlands, first-tier metabolic screening with FIA-MS/MS in combination with the Neobase II kit from PerkinElmer is performed on underivatized samples.

The study groups have not reported any other significant technical problems in the implementation of GAMT-D in the NBS. On contrary, Pitt et al. (2014) stated that GAMT-D was easily implemented in the existing NBS screening with an acceptably low false positive rate (after the second-tier).

Choosing the optimal cut-off values. For the determination of the most optimal cut-off values for GAMT-D screening, the published data of the GAA values in true GAMT-D patients within the first days of life are scarce, because the disease is typically diagnosed later. Reports of the GAA levels in the neonatal period exist of altogether eight patients (Table 1.2). Five patients had GAA levels of 7-18 μmol/L (mean 11.4) aged <72 hours and three patients had GAA levels of 12.1-28.7 μmol/L (mean, 18.8) aged <7 days.

Reference values of GAA, Cr and GAA/Cr ratio have been published in altogether 8257controls (Table 1.2). GAA levels of GAMT-D patients (12.1-28.7 μmol/L) were well above the highest GAA concentrations of the controls when measured by second tier tests (0.44-4.9 μmol/L).

One group published Cr levels and GAA/Cr ratios of three GAMT-D patients and two groups have reported these on healthy newborns (Table 1.2). Cr levels of the patients show clear overlapping with the controls, but the GAA/Cr ratios did not.

Carriers have not shown elevated GAA or lowered Cr levels (Mercimek-Mahmutoglu et al., 2016).

Discussion: Despite of the limited data and high variation of the GAA levels measured from the patients soon after birth the possibility of a false negative result of GAMT-D patient in the NBS is rather theoretical. GAA and Cr levels of all the currently published patients have exceeded the 99.9th percentile of the population. The possibility of

detecting carriers does not have to be taken into account when determining the cut-off values for GAMT-D NBS.

Table 1.2. GAA and Cr levels in GAMT-D patients and controls analyzed from butylated BSP samples collected at <7 days of age.

	GAMT-D patients range (mean±SD)				Controls range (mean±SD)			
	n	GAA (µmol/L)	Cr (µmol/L)	GAA/Cr ratio (µmol/mmol)	n	GAA (µmol/L)	Cr (µmol/L)	GAA/Cr ratio (µmol/mmol)
Merc.-Mahm. et al. 2016					500[a]	0.44-4.9 (1.14±0.45)	127-882 (408±106)	0.0001-0.014 (2.94±1.36)
Pasquali et al. 2014	3[b]	12.1-28.7 (18.8±7.5)	208-392 270±83)	0.051-0.116 (0.071±0.030)	4691[b]	0.5-10.4 (1.21±0.35)	138-877 (313±80)	0.001–0.032 (0.004±0.001)
					195[c]	0.44-3.21 (1.42 0.54)		
El-Gharbawy et al. 2013	3[d]	7-12			66	1.2 ± 0.5		
Merc.-Mahm. et al. 2012	2[a]	9.1; 10.7			3000[a]	0.5-2.06		
Bodamer et al. 2009	1[d]	18				<10		

[a] BSPs collected aged < 72 hours, measured by second-tier GC-MS/MS or UPLC-MS/MS

[b] BSPs collected aged <7 days

[c] BSPs collected at <7 days of age, measured by second-tier GC-MS/MS

[d] BSP collected aged <72 hours, method not available

Measuring multiple metabolites in GAMT-D screening. The GAMT-D patients have shown normal Cr levels at birth, since it rather reflects the maternal creatine concentration transferred via placenta to the fetus than true Cr levels of the newborn. Cr levels of GAMT-D patients differed from the controls only at two weeks of age (Pasquali et al. (2014) along with the clearance of the maternal creatine from the circulation of the newborn.

Pasquali et al. (2014) examined, whether the addition of GAA/Cr ratio as a secondary marker would improve FPR. They calculated that using a combined cut-off of 99.5th percentile for GAA and 99.9th percentile for GAA/Cr ratio, only 0.08% of the normal newborns would give false positive result, while all GAMT-D patients would be identified (instead of their FPR of 0.19% using solely GAA cut-off of 99.0th percentile).

It has been discussed that the GAA/Cr ratio could improve the sensitivity of the screening, but caution must be regarded against the possibility of false low results of GAA/Cr ratio due to the transplacental creatine transport to the fetus, although the levels have been markedly higher in the published GAMT-D patients (Mercimek-Mahmutoglu et al. 2016).

Interpretation of GAMT gene variants. There are two common pathogenic *GAMT* mutations, c.327G>A and the Portugesian founder mutation c.59G>C [Mercimek-Mahmutoglu et al. 2012]. However, since only a limited number of patient have been diagnosed worldwide, only a limited number of pathogenic *GAMT* mutations has been detected. Therefore, it is likely that the evaluation of the pathogenecity of the identified *GAMT* variants in the NBS necessitate additional functional analysis, ie. the measurement of the enzyme activity.

Discussion: Functional analysis of the *GAMT* variants have already been internationally validated and stated applicable (Mercimek-Mahmutoglu, et al. 2016).

1.4 The strategies for reducing false positive rate in NBS

We also evaluated other possible methods for reducing FPR, focusing primarily on amino acid related diseases, since GAA and Cr share molecular similarities with amino acids). The potentially most effective strategies to control false positive rate in NBS include a) measuring novel or additional biomarkers, b) using a second tier strategy and/or mutation analysis, and c) using post-analytical tools (Hall et al., 2014; Malvagia et al. 2020):

- For GAMT-D, using the measurements of Cr and the GAA/Cr ratio as additional biomarkers have shown some benefits, but also potential pitfalls, as described above.
- Second tier testing has significantly reduced the false positive rate for GAMT-D. The second tier can be a low/medium throughput method with low/intermediate costs. A mutation analysis could be an excellent option for a second tier, but this will only be feasible if the numbers after the first tier are low.
- Post analytical tools are used to improve positive predicted value (PPV) in NBS by comparison to reference data of confirmed cases rather than deviation from the reference range. However, adopting this strategy is not possible in GAMT-D, since the number of available samples of true GAMT-D patients is insufficient for reference data (Hall et al. 2014).

1.5 Screening methods for GAMT-D – Expert consultation

To assess the risks and benefits of different analytical screening methods, the manufacturer of the test kit currently used in the NBS (PerkinElmer) was contacted for an update on expanding this kit with GAMT-D and the experts of the international screening laboratories with a long experience of NBS for GAMT-D were contacted with a questionnaire to discover and assess the risks involved in in-house tests. Additionally, GAMT-D screening was discussed with several experts participating in the Cerebral Creatine Deficiency Syndromes workshop, Rotterdam in 2019 (Supplement 1).

Summary of the contact with PerkinElmer

In march 2021 PerkinElmer PE has informed us they are interested in developing a kit for GAMT-D screening and have shown interest in our validation experiments on the PE QSight220. Earlier they have investigated the separate measurement of GAA for newborn screening purposes (Supplementary Figure 1) on a research basis only.

Summary of the expert consultations

The following laboratories were contacted:

1. ARUP Institute of Clinical & Experimental Pathology, Salt Lake City, USA
2. BC Children's Hospital, Vancouver, Canada
3. University of Melbourne, Victoria, Australia

Questions

- When did your laboratory add GAMT-D to NBS?
- Does your laboratory currently perform NBS using a certified kit (which) or are you using an in-house method?
- Which method are you using for GAMT-D in NBS? In-house method / adjusted kit?
- Do you measure GAA or GAA/Cr ratio? Why?
- Does GAMT-D measurement require an additional BSP in your laboratory?
- How many and which tiers does your laboratory perform? Are all tiers performed in own lab?
- Which standards are you using? Do you keep a stock and how much?
- Which proficiency samples are you using?
- What are your cut-off values? Which percentile?
- How many neonates did your lab test in total?
- How many true and false positive have you encountered?
- Have you investigated false negatives and if yes did you encounter any?
- Have you encountered problems not using an CE-IVD certified kit or method (legal, analytical, IT)
- How much did adding GAMT-D to your NBS program cost per sample?

The laboratories had started GAMT-D screening in 2002, 2012 and 2016 either as a population-based pilot study or integrated into NBS program, and screened 148,000-771,000 infants. GAMT-D screening was performed from butylated samples and added to an already existing 'in-house' test (Pitt et al. 2014, Pasquali et al., 2014, Sinclair et al. 2016). The laboratories used different screening scenarios of two- or three-tier testing presented in Table 1.1. All tiers were performed in one laboratory. As a first-tier test, one laboratory measured GAA and GAA/Cr ratio to decrease FPR, but the two others measured GAA only, since they had not found that Cr would improve the test performance (and second-tier GAA was more effective).

GAMT-D screening was included in the routine NBS program in all the laboratories. In three-tier programs, the first two tiers were completed from the initial BSP, and the tier 3 DNA analysis required an additional BSP.

All laboratorios used 99.5th percentile cut-offs, ranging from 2.44 to 5.0 µmol/L (the experts had observed that the cut-off values were still comparable to the previously

published values presented in Table 1). The cut-off values are similar for all newborns regardless of sexes and being born term or premature. No special cut-offs for patients having had blood transfers or being adopted exist. FPRs after two tiers were 0.001-0.0003% (n=2-3) and after DNA analysis, no false positives were found. None of the laboratories had found true positives. False negatives had been investigated and were not found: at the time of the enquire, these laboratories had not encountered any new GAMT-D patients. Notably, the laboratories had screened GAMT-D for 3-17 years, and all of them are the expertise centers of their regions, so they would have been assumed to become aware of any diagnosed patients.

The laboratories had not encountered problems when not using an IVD-CE certified kit or method. All of them used CDC´s proficiency samples (one had an additional in-house proficiency testing), and purchased GAA and Cr from Sigma and labelled standards from different sources. The costs of GAMT-D screening were considered almost neligible, because it was integrated in the existing NBS. The costs were due to the purchase of standards and the DNA sequencing (1-2 per year).

Discussion and conclusions. Based on the expert consultations, in-house screening tests for GAMT-D are successfully implemented in NBS programs in several countries, showing an efficient performance with respect to FPR. Taken into account the variation in the screening scenarios between the centers, the benefits and disadvantages of adding Cr into screening must be evaluated in the Netherlands. Notably, within the first days of life, Cr concentration reflects rather the levels of the mother than the newborn (transplacental transduction).

1.6 Conclusions

On the basis of the results from the pilot studies on GAMT-D screening, the international research groups unanimously recommend at least two-tier testing in order to minimalize the FPR. The first two tiers could be performed from one BSP. Third tier DNA analysis would necessitate a new punch, but it reduced the FPR to zero. Using Cr/GAA ratio and GAA may not be efficient, but will be examined also in this study.

Based on the expected 170,000 births per year and the largest pilot study published (comprising of 771,345 with a FPR of 0,02%) (Pitt et al., 2014) 3 positive samples per month after tier 1 are expected in the Netherlands. Based on the highest reported FPR in a pilot of only 500 BSPs (0.6%), 20 (false) positive samples per week would be expected.

All other study groups used derivatized (butylated) samples for their first tier by FIA-MS/MS and hence, in the evaluation of the optimal screening scenario for GAMT-D in the Netherlands, we pay particular attention to the expected FPR when measuring *un*derivatized samples (especially in the premature group), which is the current method in the Neobase II kit from PerkinElmer used in the first tier of the Dutch NBS program, where GAMT-D screening could potentially be added to.

Adding GAMT-D screening to the currently used procedures may require an adaptation of the commercial assay currently used to screen over ten metabolic diseases. However, this will void the CE-IVD certificate for that assay. Therefore, the possibility of developing and validating an in-house screening test was studied as well. In addition, continuity of the measurements will be an important point of attention.

The in-house GAMT test will require an additional blood punch. The necessity and investment in additional mass spectrometers need to be evaluated, quality assurance of new tests needs to comply with current standards, new reference material needs to be available to ensure fidelity of test results, communication channels need to be updated, logistics have to be optimized, a risk analysis needs to be performed, legal issues (with special reference to European tender legislation and CE-IVD rules for in-house (non-commercial) tests) and additional costs need to be evaluated.

While immediate availability of the screening results is not eminent for GAMT-D screening, the turnaround time will fit within the current program (<5 weeks).

2 METHODS FOR GAMT-D NBS IN THE NETHERLANDS

Based on the current knowledge of GAMT-D NBS and the current international practises (Chapter 1), we investigated, whether a three-tier strategy with a high throughput MS/MS assay as a first tier, a more specific MS/MS assay as a second tier, and DNA analysis as a third tier (to reduce the FPR to practically zero), is the most optimal method for the Dutch NBS program.

In May 2022, European legislation will become effective in obliging laboratories to use CE-in-vitro diagnostic (CE-IVD) certified tests. However, currently there is no commercial CE-IVD certified test kit available for the detection of the most important markers for GAMT-D screening, GAA and Cr, which we can use as a first tier test. Therefore, the option of adding GAA and Cr to the Neobase ™ II kit of PerkinElmer, which is currently in use by the Dutch NBS program, was investigated in this study together with the possibility of developing an in-house tier 1 method for GAMT-D screening.

2.1 Scenario's for GAMT-D screening methods

We identified several possible methods for GAMT-D screening based on literature and current international practices.

<u>Tier 1 GAA (and Cr) assay</u>

No commercial tests for GAMT-D screening are available. All in-house methods used are based on a FIA-MS/MS detection of derivitized samples, but underivatized samples are currently used with the Neobase ™ II kit. Although, derivatization of samples is very sensitive, it is less time efficient and more costly. In order to fit the in house test into the Dutch Program a method was developed that resembles the work-up of the current screening and is more high throughput.

<u>Tier 2 GAA (and Cr) assay</u>

- Using an already described LC-MS/MS method (Mercimek-Mahmutoglu et al., 2016)

<u>Tier 3</u>

- DNA analysis of *GAMT* gene by exon sequencing (Mercimek-Mahmutoglu et al. 2016)

2.2 Aspects to consider when choosing the first-tier screening method

GAMT-D does not significantly differ from other metabolic diseases currently included in the Dutch NBS. For instance, the molecular structure of GAA is very similar to amino acids. The treatment and the perspectives for patients show similarities with, for instance, PKU, although the estimated prevalence of GAMT-D is markedly lower than PKU (1:250,000 vs. 1:18,000, respectively) (Verkerk et al., 1995). As also noted in The Health Council´s advisory report 'Neonatale screening: nieuwe aanbevelingen' on April 8th 2015, the previous results of GAMT-D NBS have showed low false positive rate and no false negatives after the second test.

Test method. Adding the first-tier test to the commercial kit would technically have the benefit of a limited extra resources (machines, personnel, know-how, etc.), but it would void the certification of the existing kit and this would pose a problem considering the upcoming European legislation. An in-house test gives more flexibility (exchange FIA-MS/MS for LC-MS/MS to enhance sensitivity), but risks the continuity of the assay. Tier 2 and 3 were already described in Mercimek-Mahmutoglu et al., 2016.

Analytical challenges of GAMT-D newborn screening are one of the most significant factors to examine/consider when evaluating different (logistic) scenarios. The major challenges arise from the following.

- *Similar molecular mass of GAA and Valine.* GAA has the same molecular mass as valine, a compound that is already measured using the Neobase ™ II kit. Valine and GAA share 2 fragment ions, m/z 72 and m/z 55 (see Supplemental Figure 2-4). The cut-off value of Valine is currently ≥ 340 μM. GAA has reference values of 0.5-10.4 μM. GAMT-D patients have shown values up to 28.7 μM, which is ~7% of the cut-off value of Valine and thus, not likely to contribute to additional false positives for valine related diseases. However, using either of the 2 shared fragment ions will result in falsely elevated levels of GAA. Another fragment should be used in order to accurately measure GAA. The results regarding these investigations are discussed in paragraph 3.5. The chosen m/z 76 fragment is specific (loss of guanidine group, see Supplemental Figure 5).
- For quantification purposes, either a calibration curve or calculation using the internal standard can be used (see Chapter 3). In case a calibration curve is chosen, the standard of GAA cannot be added to the standard mix of the Neobase™ II kit because it will result in wrong values of valine. Choosing another fragment for Val might be a solution but Val does not have other high intensity fragments (see Supplemental Figure 4). In addition, valine is already part of the screening for MSUD which will then be influenced. A solution could be the addition of a separate standard for GAA. Choosing the IS for GAA smartly will not result in problems for the addition to the kit. There are several options for isotopically labeled GAA: [D_2], [$^{13}C_2$], [$^{13}C_2$, 3-^{15}N]. Only the latter could interfere with [$^{13}C_5$]-Proline of the Neobase ™ II kit that shares the same molecular mass. Therefore only [D_2] and [$^{13}C_2$]-GAA are evaluated in this study.

For the development of such an 'in-house' method to be applied in an ISO15189 certified laboratory, the major considerable points are:

- *Efficacy of the method to detect GAMT-D patients.* Considerations whether the in-house method is good enough to perform GAMT-D NBS screening was part of this project and is evaluated at the end of Chapter 3, which turned out to be adequately to detect GAMT-D patients.
- *Continuity of the assay.* To ensure the continuity of GAMT-D screening, the (internal) standards ought to have proper availability, interchangeability, and stability in storage.
- *Variation between standard batches.* Preparation of any (internal) standard runs the risk of error and, if made in all laboratories separately it can lead to bias. Therefore, all newly prepared standard batches need to be evaluated prior to use (added to the SOP, Appendix 2).
- *Quality of the method.*
- *Proficiency samples* (4 levels) for measurement of GAA and Cr in BSPs are provided by the DCD, Centers of disease control and prevention (Paragraph 3.2.15).
- Development of an 'in-house' method will open up the flexibility to add additional diseases in the future that are not part of a screening kit.

3 Test qualities of GAMT screening pilot

During optimization we discovered that the limit of quantification for GAA was too low to adeqautely measure control samples due to matrix suppression (see Validation report), a phenomenon inherent to FIA-MS/MS. Therefore, GAMT-D screening is not feasible using FIA-MS/MS and hence, it was not possible to examine, whether it could be added to the FIA-MS/MS-based Neobase™ II kit. Therefore, we validated a LC-MS/MS-based in-house method, which remained as the only possible valid option. The following summarizes the results of the validation, which are presented more detailed in the separate Validation report (Appendix 1).

3.1 Validation of the first-tier test

For the validation of the first-tier in-house method, GAA and Cr were measured using comparable equipment as currently used in the screening laboratories, the PerkinElmer Qsight 220 mass spectrometer. Additionally, part of the analysis were performed using Sciex API5000 to evaluate, whether another brand would be an option. The current method for quantifying GAA and Cr has previously been validated in blood at the VUMC Laboratory and it was now validated for the analysis of dried blood spots on filter paper.

1/8 Inch BSP punches of 4 level CDC profiency samples were used when possible. Additionally, newborn screening samples from the current Dutch NBS were used for intra- and inter-assay evaluation and spiked volunteer adult samples (prepared as newborn screening samples) to assess stability of Cr and GAA.

Since GAA and Cr are not currently included in any commercially available newborn screening kit, the following were necessitated in the assessment and validation of the in-house first-tier screening method:

1. optimization of parameters for the Qsight
2. preparation, stability and availability of (internal) standards
3. assessment of extraction efficiency of Cr and GAA from dried bloodspots
4. performance parameters precision, accuracy, robustness, linearity, LOQ (limit of quantification)
5. the availability of proficiency samples

3.1.1 Optimization of parameters for Qsight

Optimization of the parameters in the mass spectrometer is the first step in testing a new method. It is performed to assess whether measurements of desired compounds will be feasible to detect using specific machines and if so, to determine optimal setting for these compounds.

During the optimization of the parameters in the FIA-MS/MS, using the same reagents that are currently used by the Neobase™ II kit in the Dutch newborn screening, it turned out to be not feasible using the PerkinElmer´s QSight 220 mass spectrometer due to a severe suppression of the GAA signal by the BSP matrix and low concentrations of GAA (Validation report 1.1). For this reason, the addition of Cr and GAA to the PerkinElmer´s FIA-MS/MS-based Neobase ™ II kit is not possible, and only the in-house method was validated in this study. FIA-MS/MS could be possible for the Dutch newborn screening, if another machine would be purchased: using the SciexAPI5000 the LOQ for GAA was adequate enough for measurement, which would make this a more sensitive alternative to the QSight 220. Since currently PerkinElmer QSight mass spectrometers are available in the screening laboratories, we proceeded with this machine for further validation of the in-house method by switching the first tier from FIA-MS/MS to LC-MS/MS method using a fast non-specific separation.

3.1.2 Preparation, stability and availability of (internal) standards

To ensure the continuity of the screening in cases of supply issues, the standards in use should either have available alternatives or prolonged shelflife. We found that multiple standards are available (Supplementary Table 1-4), serving as alternatives for each other (Validation report 2.1). Even the distinct labeled internal standards of GAA are comparable (Validation rapport 2.14). Based on the results of this study, [D_3]-Creatine and [$^{13}C_2$]-GAA were chosen as internal standards.

The results of this study also showed that GAA and Cr have over 10 years of shelflife in powder form, and that Cr and GAA remain stable under possible different transport conditions and during the whole target turnaround time (Validation report 2.2).

Storaging 100 mg of each standard serves as an adequate backup resource, covering more than 2 million samples. For (internal) standards, 3 options are available, of which freeze dried batches would be the most prefarable: no freezer capacity necessary and they were shown to be stable for >10 years. As an alternative, stock solutions having > 5 years shelf life could be an option.

3.1.3 Assessment of extraction efficiency of Cr and GAA from dried bloodspots

The extraction solution should adequately extract GAA and Cr from the BSPs. Except for the lowest addition (the added amount approximately 10% of the total concentration, which is basically the variation of the method), for both GAA and Cr, the method showed good extraction efficiency (86% for GAA and 100% for Cr) (Validation report 2.9).

3.1.4 Performance parameters

In order to accurately identify GAMT-D patients in the newborn screening, the method has to comply with certain performance criteria (Validation report 2.7-2.16). The expected control ranges are 138-877 µmol/L for creatine and 0.5-10.4 µmol/L for GAA. For patients, the expected GAA range is 12.1–28.7 µM and Cr range is 208–392 µmol/L (Pasquali 2014).

Linearity	Cr and GAA displayed a linear response in the expected control ranges.
LOQ	The LOQ for Cr was well below (<48.6 µmol/L) the expected lowest Cr concentration in controls. The LOQ for GAA was only just above (<0.73 µmol/L) the expected lowest control GAA concentration, but this is adequate to detect GAMT-D patients. A benefit of this LOQ is that patients with Arginine:glycine amidino transferase (AGAT) defiency, who have low GAA values will not be found with this method.
Accuracy	Concentrations of proficiency samples (range for Cr was 229-721 µmol/L and for GAA 1.1-19.3) were within 1SD of other laboratories.
Precision	All CV results are <9% except for the lowest CDC sample for GAA (1.1 µmol/L, cut-off value 3.06 µmol/L)), which is to be expected since it is near the LOQ for this method.
Carry-over	Samples with high GAA or Cr concentrations will not contribute to the result of the following samples.
Drift	No drift in the concentrations in time was found.
Robustness	Extraction solutions having different lot numbers did not result in deviation of the measured concentrations.

3.1.5 Proficiency samples

Analytically valid methods require that proficiency samples need to be either available or prepared to ensure the accuracy of the determined concentrations. The Centers for disease control and prevention (CDC) offer laboratories a quality ensurance program with 4 levels of GAA or Cr BSP proficiency samples (CDC Newborn Screening Quality Assurance Program). Every half year new controls are available to run alongside each batch. These samples were used for the validation of this method (Paragraph 2.15).

3.2 Reference values & cut-off values

For the determination of the reference values and subsequent cut-off values, 469 bloodspot punches from the current screening program were anonymized and divided into 4 groups of 120 samples:

- Group 1: Term female

- Group 2: Term male
- Group 3: Preterm female (≤36+0 weeks of pregnancy)
- Group 4: Preterm male (≤36+0 weeks of pregnancy)

3.2.1 Comparision between groups

GAA and GAA/Cr ratios of the samples were measured and the results were compared between the groups (Figure 3.1 and Table 3.1). Cr and GAA were significantly higher in preterm males compared to term males. No other statistical differences were found.

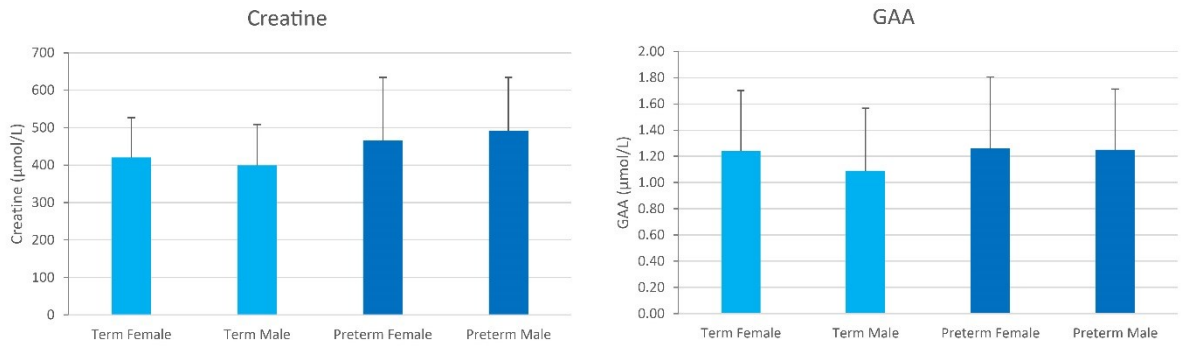


Figure 3.1 Average Cr and GAA levels in BSP of term and preterm children (bar represents SD, n~120 per group)

Table 3.1 Differences in Cr, GAA and GAA/Cr ratios between the groups (Independent Student T-test, SPSS).

	Sig. (2-tailed)		
	Cr	GAA	GAA/Cr ratio
Female Term versus Male Term	0.133	0.015	0.233
Female Term versus Female Preterm	0.020	0.811	0.554
Female Preterm versus Male Preterm	0.180	0.976	0.317
Male Term versus Male Preterm	0.000	0.008	0.640

3.2.2 Gestational age

On basis of the results in paragraph 3.2.1, we further examined by scatterplots if there is a relationship between Cr and GAA concentrations or GAA/Cr ratio and the gestational age (i.e. duration of pregnancy). Cr increases slightly with gestational age, also observed in the previous paragraph for males. No correlation was found between gestational age and GAA.

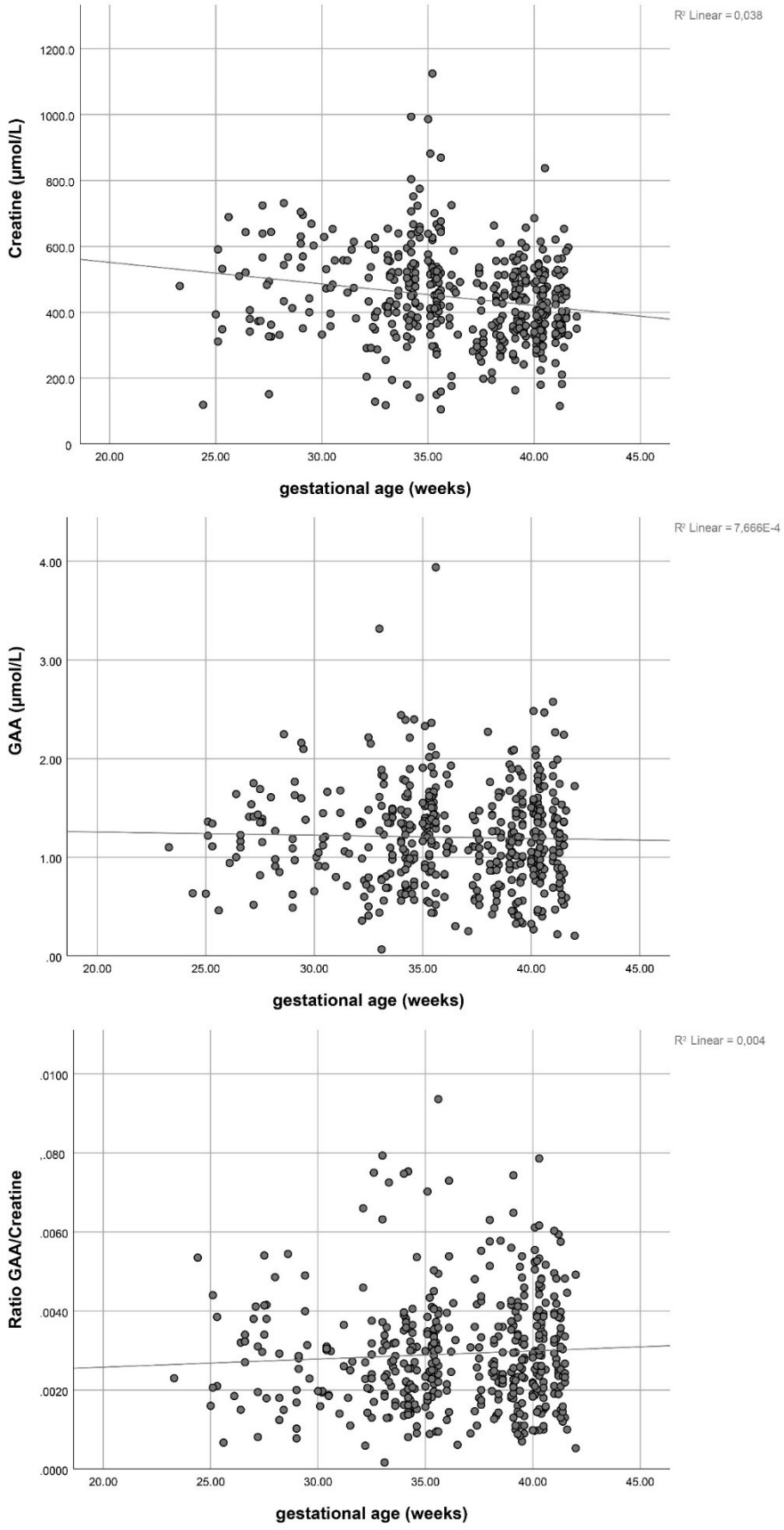


Figure 3.2 Scatterplots of Cr and GAA concentrations and GAA/Cr ratio versus gestational age, n=469.

3.2.3 Distribution

Since no difference between Cr and GAA and gender was found and no trend between GAA and gestational age was seen in the scatterplots, all groups were combined and showed a normal distribution is observed for Cr, GAA and GAA/Cr ratio (Figure 3.3).

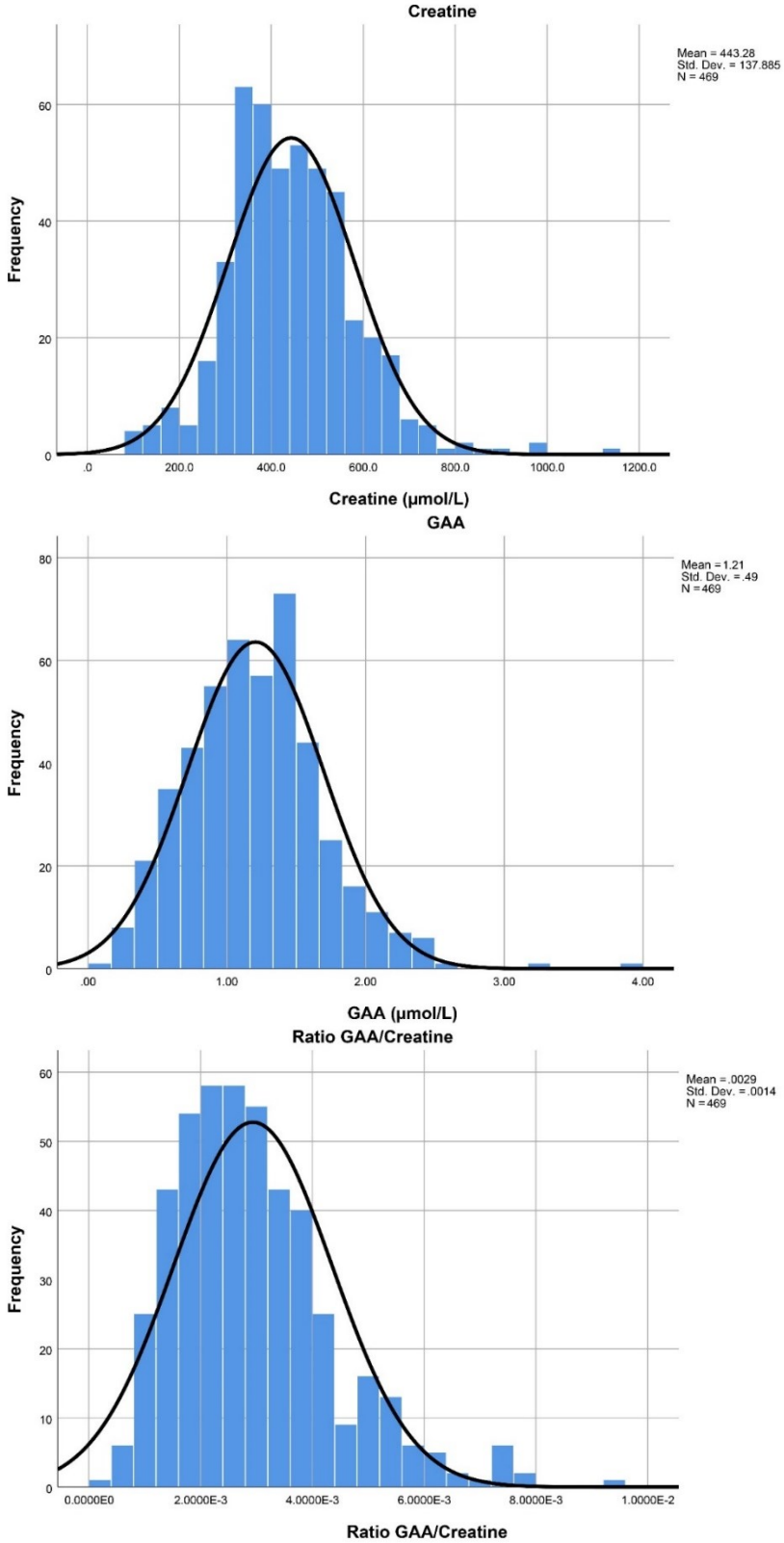


Figure 3.3. Histograms of Cr concentrations, GAA concentrations and the GAA/Creatine ratio in newborn screening BSP samples (n=469).

The combined average of the four groups (Table 3.2) are in agreement with previous studies (Pasquali et al., 2014; Sinclairet al., 2016).

	Average	Std Dev
Cr (µmol/L)	443.3	137.9
GAA (µmol/L)	1.21	0.49
GAA/Cr	0.0029	0.0014

Table 3.2 Average concentrations of Cr, GAA and GAA/Cr ratio of all groups combined (n=469).

3.2.4 Cut-off values

We determined gender- and age-independent cut-off values for the GAMT-D screening, since no differences were found between different groups.

Percentile	Cr	GAA	GAA/Cr	False positives
0.5	116.1			
1	125.4			
99		2.47	0.0075	0.85%
99.5		3.06	0.0079	0.43%

Table 3.3 Proposed GAA and Cr cut-off values for GAMT-D newborn screening.

The cut-off values found were comparable to those presented in Table 1.1. Both 99.0th and 99.5th percentile cut-off values resulted in markedly lower GAA concentrations than previously reported in GAMT-D patients (lowest 4.93 µmol/L; Pasquali et al.; 2014, Table 1.2).

Using cut-offs of 99.5th percentile (3.06 µmol/L), two reference samples exceeded the cut-off limit with GAA consentrations of 3.32 and 3.94 µmol/L, indicating a false positive rate of 0.4% in the first tier assay.

The addition of the GAA/Cr ratio in the analysis did not reduce the number of false positives: the same two false positive samples were detected with 99.5th percentile cut-off, and using 99th percentile cut-off resulted in discrepancies between GAA and GAA/Cr false positives (some samples showed abnormal GAA, some GAA/Cr ratio and some both).

Conclusions

We interprete that the 99.5th percentile cut-off value is the most optimal for the detection of GAMT-D in the newborn screening. It is in line with the cut-off values used in previously reported pilot studies (2.94–5 µmol/L) and is expected to result in acceptable false positive rate. The GAA/Cr ratio has no additional value in this method, but Cr could serve both as extraction efficiency control (low control GAA concentrations are below the LOQ) and as an indicator for that a detected elevated GAA concentration is due to GAMT-D (the patients have low Cr concentrations).

The validity of the GAA/Cr ratio could be evaluated more extensively after screening more samples. For now, we recommend that in addition for GAA, Cr is measured only when GAA exceeds the cut-off value. Taking into account the limitations (Cr

concentrations in the newborn can sometimes be affected by maternal Cr levels) and benefits (the potential in decreasing FPR), this would be most efficient method.

3.2.5 Positive control

An archived (40 months) newborn bloodspot card of a biochemically and genetically confirmed GAMT-D patient was used to verify, whether the new method would screen this bloodspot punch as positive. For creatine a value of 214.8 µmol/L was found and for GAA 13.64 µmol/L was found. The GAA/Cr ratio was 0.064. Both GAA and GAA/Cr are well above the cut-off values when using the 99.5th percentile cut-off. Creatine was in the lower range of the controls as expected. This confirms that the method used is able to detect GAMT deficient neonates.

Conclusions: All parameters fulfilled the criteria and the designed in-house method in this study detects GAMT-D patients adequately. This validated in-house method is expected to be ISO15189 compliant in all screening laboratories. A 99.5th percentile cut-off value for GAA is recommended and Cr, for now, does not seem to help lower the false positive rate.

3.3 Tier 2

The two screen positive samples exceeding the 99.5th percentile (3.06 µmol/L) GAA cut-off limit values were measured by the second tier assay using LC-MS/MS method decribed as first tier method in Mercimek-Mahmutoglu et al., 2016. Both samples turned out to be false positives.

	GAA (µmol/L)	
	Tier 1	Tier 2
Sample 1	3.32	0.97
Sample 2	3.944	2.95

Table 3.4 Results for GAA for tier 1 and 2 for samples above 99.5th percentile cut-off value.

3.4 Tier 3

The procedure of direct sequencing of the *GAMT* gene has been implemented on BSP´s of Dutch neonatal blood spot cards (Mercimek-Mahmutoglu et al., 2016). DNA will be extracted from dried blood spot punch by Generation DNA Purification kit. Sequencing of the *GAMT* gene (NM_000156.5) could be performed in the Metabolic Laboratories of Amsterdam UMC. The analysis can be performed within 2 working days. In the majority of cases this will reveal a true positive or negative result, presence of mutations YES or NO, respectively. In a few cases, variants of unknown significance may be detected that need further diagnostic workup.

3.5 Specificity and sensitivity

For the pilot a preliminary specificity could be established for tier 1 using the 99.5th percentile cut-off:

True negatives / (True negatives + False positives) x 100% = 469 / (469+2) = 99.6%

The sensitivity could not be established with this pilot since no true postives where found, but published studies already show that with the proposed 3 tier screening method no false positives were found and no GAMT-D patients were identified in the screened region.

3.6 Expected incidental findings

Elevated GAA is a unique parameter for this disease. No other disease could unintentionally be discovered screening for elevated GAA. The only other disease where GAA concentrations are abnormal is AGAT, but for this disorder GAA levels are lowered. With this method the lower range of reference values is just below the detection limit and therefore AGAT patients cannot be found using this method. Also the GAA/Cr ratio will not uncover AGAT patients.

Carriers of *GAMT* variants have not shown elevated GAA or lowered Cr levels (Mercimek-Mahmutoglu et al. 2016) and therefore will not be detected by the screening.

We recommend that in addition for GAA, Cr is measured only for establishing GAA/Cr ratio when GAA exceeds the cut-off value and hence, mother´s with high Cr levels will not be detected (and even if they would be, it wouldn't refer to any disease).

4 GAMT-D screening scenarios

All possible scenarios for GAMT-D NBS in the Netherlands were initially determined by combining the options for three-tier testing either by adding GAMT-D to the commercial kit (Scenario 1) or developing an in-house test (Scenarios 2a-c), and considering all the possibilities to perform the tier 1 and 2 assays in either regional or centralized laboratories (Figure 4.1).

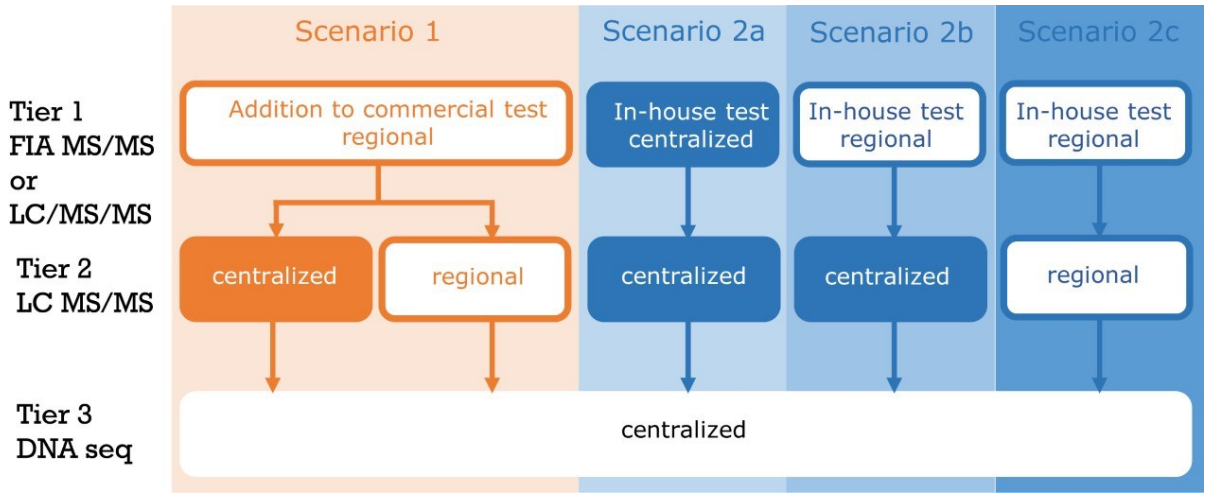


Figure 4.1 GAMT-D Newborn screening scenarios.

As presented in the results of this study (Chapter 3), the optimization of the the sensitivity of the FIA-MS/MS for GAA was too low (Validation report). Hence, the addition of the GAMT-D screening to the existing Neobase ™ II kit (PerkinElmer) (Scenario 1) showed not be a valid option and Scenario 1 had to be rejected.

All three tiers could be potentially performed either in every 5 regional screening laboratories or centralized in one screening laboratory having at least one screening laboratory as a back-up. However, in Scenario 1, the first tier could only be performed regionally, since it should then follow the current procedure. Additionally, for tier 3 DNA sequencing, the expected number of samples is low and only one centralized laboratory is sufficient.

4.1 Pros and cons of the scenarios

With the expected 170,000 births per year in The Netherlands, altogether 3200 samples per week are expected for tier 1. Based on the results of the previous international pilot studies (Table 1.1), the expected workload for tier 2 is 1 to 20 (false) positive samples from tier 1 per week, which is comparable to the current screening for propionic acidemia (PA) and methylmalonic acedemia (MMA) in the Dutch NBS program. This means approximately 0 to 4 samples per laboratory per week, if all the screening laboratories perform tier 2. For tier 3, only 0 to 3 (false) positive tier 2 samples are expected per year in total.

Recently, all screening laboratories aquired PerkinElmer Qsight 225MD mass spectrometers for tier 1 screening of Mucopolysaccharidose type I (performed 4 times per week) and tier 2 screening of PA and MMA (performed 2 times a week by Bilthoven

and once a month as instrumental back-up in Amsterdam). Logistic perspectives were considered for the different different scenarios (Figure 2.1):

Scenario 1: Addition to the Neobase ™ II kit. During optimization of the method, this option turned out not to be feasible. Logistically it would follow scenario 2b.

Scenario 2a: All tiers are performed centralised.

Tier 1

- An additional BSP from the heelprick card will be needed for tier 1, designated specifically for this measurement and sent to the central laboratory.
- Expected samples: 3200 per week.
- With a run time of 2 min per sample, 110 hours of additional LC-MS/MS will be required per week. None of the screenings laboratories has the analytical capacity to perform this amount of samples on their current machines (even if including the previously mentioned new mass spectrometers), which necessitates purchasing of at least two LC-MS/MS systems in the central laboratory.
- Samples should be send to the central laboratory on a daily basis in order to be able to process all the samples. Samples could potentially get lost on their way.
- Measuring a large number of samples only in one laboratory is vulnerable in case of malfunction (no back-up available), which would require the purchase of even more mass spectrometers for the back-up laboratory.

Tier 2

- 1 to 20 (false) positive samples per week are expected. With a run time of 10 minutes per sample, approximately 5 hours of LC-MS/MS will be required weekly (including controls and start-up).
- Transport of the samples to the central laboratory needs to be arranged (once a week is sufficient: taking in account estimated false positives and disease course).
- Sample preparation differs from the current NBS program and requires an evaporation device. Considering that the conversion to the LC-MS/MS of the tier 2 method and the preparation of the required solutions takes approximately 1 hour, measuring tier 2 in one central laboratory (with at least one laboratory as back-up) is more efficient and necessitates less personnel to be trained. For PA/MMA, measurement of tier 2 is currently centralized with another laboratory as back-up.

Tier 3

- For the third tier DNA analysis, an additional BSP will be needed and shipped to the central laboratory.
- Expected number of samples: 0 to 3 per year.

Scenario 2b: Tier 1 measured in all screening laboratories, tiers 2 and 3 are centralised.

Tier 1

- An additional BSP will be needed for tier 1 of GAMT-D which can be taken out of the heel prick card. Since sample preparation is almost identical to the current

screening, limited training of personnel is needed which makes it possible to measure tier 1 in all laboratories.

- Expected samples: 650 per week per screening laboratory.
- Each measurement takes 2 minutes and thus, 22 hours of additional LC-MS/MS is required per week per laboratory. So, in addition to a doubling of sample pretreatment workload (effect for amount of personnel necessary), each laboratory needs to have at least 22 hours of LC-MS/MS available weekly. Except for Bilthoven, these measurement would fit on the previously mentioned new PerkinElmer mass spectrometers. For Bilthoven, an additional mass spectrometer needs to be purchased.

Tier 2

- Transportation of the sample to the central laboratory needs to be arranged weekly (according to the estimated number of screen positive samples after tier 1).
- Expected samples: 1 to 20 (false positive) samples per week.
- With a run time of 10 minutes per sample, approximately 5 hours of LC-MS/MS needs to be available per week, which would fit on the current mass spectrometers.
- Sample preparation differs from the current NBS program and requires an evaporation device. Considering that the conversion to the LC-MS/MS of the tier 2 method and the preparation of the required solutions takes approximately 1 hour, measuring tier 2 in one central laboratory (with at least one laboratory as back-up) is more efficient and necessitates less personnel to be trained. For cystic fibrosis, measurement of tier 2 is currently centralized with another laboratory serving as a back-up.

Tier 3

- For the third tier DNA analysis, an additional BSP will be needed and shipped to the central laboratory.
- Expected number of samples: 0 to 3 per year.

Scenario 2c: Tier 1 & 2 measured in all screening laboratories and tier 3 in centralized laboratory.

Tier 1

- An additional BSP will be needed for tier 1 of GAMT-D which can be taken out of the heel prick card. Since sample preparation is almost identical to the current screening, limited training of personnel is needed which makes it possible to measure tier 1 in all laboratories.
- Expected samples: 650 per week per screening laboratory.
- Each measurement takes 2 minutes and thus, 22 hours of additional LC-MS/MS is required per week per laboratory. So, in addition to a doubling of sample pretreatment workload (effect for amount of personnel necessary), each laboratory needs to have at least 22 hours of LC-MS/MS available weekly. Except for Bilthoven, these measurement would fit on the previously mentioned new

PerkinElmer mass spectrometers. For Bilthoven, an additional mass spectrometer needs to be purchased.

Tier 2

- Expected number of samples: 0 to 4 (false) positive samples per week per laboratory.
- With a run time of 10 minutes per sample, approximately 2 hours of LC-MS/MS needs to be available per week per laboratory. All series require quality control samples to run along with the series. When the amount of sample is low, the contribution of the QC samples is relatively high.
- In addition, the LC-MS/MS needs to be converted to tier 2 method, solutions need to be prepared (conversion to the tier 2 method & preparation of the necessary solution will require approximately 1 hour). Sample pretreatment differs from the current screening and requires an evaporation device. More personnel needs to be trained to perform this analysis but this creates more flexibility to serve as back-up location for other laboratories. If tier 2 is performed in the same laboratory, no transport needs to take place for this tier.

Tier 3

- For the third tier DNA analysis, an additional BSP will be needed and shipped to the central laboratory.
- Expected number of samples: 0 to 3 per year.

Table 4.1 Summary of pros and cons of the different GAMT-D newborn screening scenarios.

		Scenario 2a	Scenario 2b	Scenario 2c
Analytical	Pros	• All expertise at one place • No variation between the labs	• Back-up labs available for tier 1	• Back-up lab available for tier 1
	Cons	• Back-up lab needed for tier 1 • Variation between machines		• Time- and resource-consuming to measure a relatively small amount of samples
Logistics	Pros		• Bulk of samples remain in own laboratory	• No loss of samples in shipping for tier 1&2. • Only 0-3 samples shipped to central lab per year
	Cons	• Samples need to be sent daily to central lab • Loss of samples in mail	• Samples need to be sent once a week to central lab	
IT	Pros	• Only 2 laboratories needs system update for GAMT-D measurement		All results are regionally collected and processed
	Cons	• Sending all results to regional offices	• Sending the tier 2 results to regional offices	• All laboratories need system update for GAMT-D measurement
Educational	Pros	• Only 2 laboratories need additional education	• Labs can serve as each others´ back-up for tier 1	• Labs can serve as each others´ back-up for tier 1
	Cons	• Fragile in case of trouble	• Limited additional instructions for tier 1 in all laboratories	• All laboratories need additional education
Quality	Pros	• No variation between laboratories		
	Cons		• Variation between laboratories	• Variation between laboratories
Costs (see 2.3)	Pros	• Only 1 additional LC-MS/MS systems necessary + less extra personnel + only 1 lab ICT update	Limited additional costs for mailing of samples	No additional costs for mailing samples
	Cons	Additional costs for daily mailing of samples	• Requires 1 additional LC-MS/MS systems necessary, extra personnel and additional ICT	• Requires 1 additional LC-MS/MS systems, extra personnel and additional ICT • More waste of resources for limited amount of samples tier 2

Red font, major cons; Green font, major pros.

4.2 Laboratory logistics

In order to map the logistic flows for the various scenarios, we identified 3 courses:

1. Measurement of samples
2. Transfer of samples to a central laboratory
3. Transfer of results to the regional laboratory

Scenario 1 (Figure 2.1) where the Neobase ™ II kit is extended with GAA and Cr turned out to be analytically unachievable. Logistics wise it would follow the same path as scenario 2b.

In the Netherlands, samples are send to regional offices based on their zip code. For scenarios 2a, 2b and 2c (Figure 2.1), separate samples need to be plated and either measured or transferred to the central laboratory.

Scenario 2a would involve the weekly transfer of 640 samples (for eastimation see previous paragraph) by regular post from the 4 regional laboratory to one central laboratory with all the associated risks described in the previous paragraph (Table 2.1). Subsequently, all samples are measured in the central laboratory and the results are transferred back to the regional laboratories. For tier 2, additional punches of the positive samples (approximately 0 to 4 samples, for estimation see previous paragraph) per laboratory per week need to be send to the central laboratory where they are measured and the result are transferred back to the regional laboratories again.

For scenario 2b, no logistic steps need to be taken for tier 1 and tier 2 will follow the same steps as scenario 2a.

Scenario 2c does not require any logistic steps until tier 3.

Tier 3 is the same for all scenarios and involves the transfer of punches of to an external laboratory from the relevant laboratory.

Since measurements in only one laboratory poses a threat to the continuity of the screening, a back-up laboratory needs to be available and measure samples on occasion.

Conclusions

Taken into account all arguments mentioned before, the Scenario 2a will not be feasible due to the sheer amount of samples that need to be transported to central laboratories. For the Scenario 2c, the availability of back-up laboratories is beneficial but the limited expected samples for tier 2 makes this option not cost-effective. Hence, the Scenario 2b is the preferred option, and is already currently applied for other diseases in the Dutch NBS.

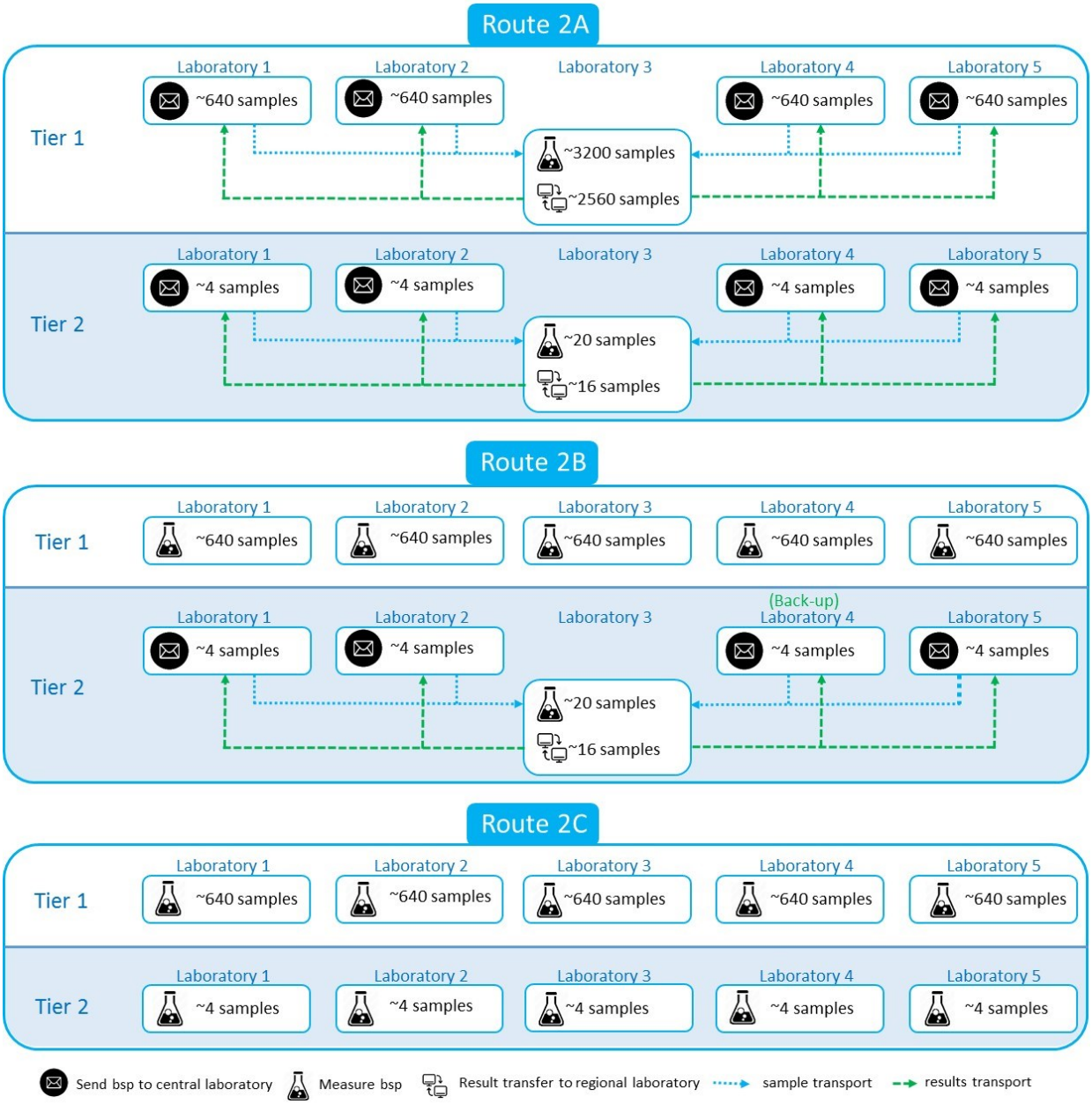


Figure 2.3 Weekly laboratory logistics for scenarios 2a, 2a and 2c based on 170,000 births a year and the maximum FPR found in literature (larger studies show lower FPR). Blue lines, transportation of the samples; green lines, transportation of the result data. Tier 3 will be the same for all scenarios and will involve 0-3 samples per year that need to be send to a central laboratory.

5 Legal, ethical and cost perspective

5.1 Legal aspects: impact of European directive 2017/746 on the use of in-house tests for NBS

European regulation 2017/746 on in vitro diagnostic medical devices applies to in-vitro diagnostic and in-vitro population screening tests including all tests proposed in this study, and will go into effect on the 26th of May 2022. It regulates the manufacture and use of Lab-Developed Tests (LDT or "in-house" tests), providing an exemption to some of the requirements in the regulation if certain conditions are met by the 'device' (i.e. the test in this case) and by the institutions manufacturing and using it, as described in article 5, section 5 and Annex I of the regulation. Because the proposed screening strategy for GAMT-D relies on in-house tests for all tiers, due to the unavailability of commercial tests, we have examined the implications of this regulation for GAMT-D screening, and for the use of in-house tests for neonatal screening in general. Below, the impact of each of the respective criteria listed in article 5 of European regulation 2017/746 are discussed one by one. The text from article 5 of European Regulation 2017/746 is reproduced (boxed, in blue).

> 5. With the exception of the relevant general safety and performance requirements set out in Annex I, the requirements of this Regulation shall not apply to devices manufactured and used only within health institutions established in the Union, provided that all of the following conditions are met:

General considerations

Firstly, it is important to note that article 5 (5) states explicitly that the 'relevant general safety and performance requirements set out in Annex I' (GSPR) must be met. Neonatal screening tests must therefore comply with this requirement, exempt or not. Annex I contains a list of safety and performance requirements that must be met and documented. The majority of these requirements will be met by tests that are developed and validated in an NEN-EN-ISO15189:2012 compliant laboratory, but annex I should be carefully reviewed and compliance with all requirements should be ensured and documented.

Considerations specific for GAMT screening

Documentation should be prepared that shows that the 'general safety and performance requirements set out in Annex I', including analytical and clinical performance of the test.

> (a) the devices are not transferred to another legal entity;

General considerations

The Dutch Neonatal screening is currently performed by five laboratories across the country. All institutions involved (4 hospitals and the RIVM) should be considered 'health institutions' and different legal entities, as defined by the regulation.

To ensure that comparable screening results are obtained by the different screening laboratories, identical reagents, methods, instruments and software are used as much as possible by the five laboratories that perform the neonatal screening. As an example, even changes between kit lots is synchronized between the laboratories. For an in-house test, it would be highly preferable for the screening laboratories to use the same batch/lot of critical reagents. The most straightforward way to accomplish this would be to buy the reagents in bulk in one of the laboratories, perform quality testing if necessary, and distribute them to the other four screening laboratories. However, this might be considered in violation with criterium 5(a) quoted above. Note that this criterium only relates to the distribution of reagents; protocols and test results can be shared freely between the laboratories. Two approaches to this potential issue are listed below:

Firstly, the Dutch Health and Youth Care Inspectorate (IGJ) should be consulted to request their view on this issue. From the European Regulation, it is not clear whether the distribution of individual reagents between the screening laboratories should be considered transferring a 'device'. If the inspectorate allows it, the pragmatic solution described above can be adopted.

If not, to avoid the issue, the screening laboratories can coordinate their orders for critical reagents from a supplier, specifying the same batch/lot. Quality checks can be performed In a single laboratory (if necessary), and the results can be communicated to the other labs.

In conclusion, this criterium does not prevent the use of an in-house method for screening, but it is important to consult the IGJ inspectorate to avoid unnecessarily complicating the distribution of critical reagents between the screening laboratories.

Considerations specific for GAMT screening

<u>1st tier</u>

The following reagents required for the 1st tier of GAMT screening are critical, (i.e. it is important to use the same (batch of) reagents to ensure that comparable results are obtained in the different laboratories:

- The Internal standards [$^{13}C_2$]-guanidinoacetate, and [D_3]-creatine
- Standards for guanidinoacetate (GAA) and creatine
- Quality control samples, consisting of dried blood spots prepared from blood spiked with GAA and creatine.

The internal standards and standards are commercially available, and can be purchased from the same supplier by the different labs, specifying a batch/lot to ensure that all laboratories receive the same reagents.

The quality control samples are not commercially available. However, they may be prepared in one of the screening laboratories and distributed for proficiency testing. Distributing samples between laboratories for quality control is a common and essential element of laboratory quality control management, such as described in NEN-EN-ISO15189:2012.

<u>2nd tier</u>

The second tier of GAMT screening tier will not necessarily be conducted in all five screening laboratories, depending among other things on the expected workload for second tier assays. However, it is likely that the 2nd tier test will be implemented in at least two of the screening laboratories (the redundancy allows for backup during instrument maintenance or unforeseen circumstances that might interrupt the service of one of the screening laboratories). Therefore, the considerations for tier 1 are applicable to tier 2 as well:

The following reagents required for 2nd tier of GAMT screening are critical, (ie: it is important to use the same reagents to ensure obtaining comparable results in the different laboratories):

- The Internal standards [$^{13}C_2$]-guanidinoacetate, and [D_3]-creatine
- Standards for guanidinoacetate (GAA) and creatine
- Quality control samples, consisting of dried blood spots prepared with blood spiked with GAA and creatine.

The internal standards and standards are commercially available, and can be purchased from the same supplier, specifying a batch/lot to ensure all laboratories receive the same reagents.

The quality control samples are not commercially available. However, they may be prepared in one of the screening laboratories and distributed for proficiency testing.

3rd tier

It is expected that only actual GAMT patients will progress to the 3rd tier, which consists of DNA sequencing of the GAMT gene. In light of the low prevalence of GAMT, it is likely that this test will be performed only at a single laboratory, and will therefore be in compliance with the above criterium 5(a) from the European Regulation.

(b) manufacture and use of the devices occur under appropriate quality management systems;

The test (i.e. the reagents that are required for performing the test) will be manufactured or purchased (and subjected to quality testing if necessary) in one or several of the screening laboratories. NEN-EN-ISO 13485, which represents the requirements for a quality management system for the manufacture of medical devices, would be an appropriate quality management system, but compliance with this NEN/ISO standard is not necessarily required.

(c) the laboratory of the health institution is compliant with standard EN ISO15189 or where applicable national provisions, including national provisions regarding accreditation;

All of the laboratories currently involved in neonatal screening are compliant with NEN-EN-ISO15189:2012.

(d) the health institution justifies in its documentation that the target patient group's specific needs cannot be met, or cannot be met at the appropriate level of performance by an equivalent device available on the market;

General considerations

To comply with this criterium, a thorough search of the market space needs to be conducted periodically (e.g. yearly) to identify any possibly suitable and certified (commercial) tests. If any are found, an assessment of their performance should be conducted, and compared against predefined criteria for 'appropriate level of performance'. If the 'appropriate level of performance' is achieved by the certified commercial test, an in-house test may not be used.

Considerations for GAMT screening

A search for CE-IVD-certified test for GAMT screening was conducted, by @. To the best of our knowledge, no certified tests are currently available on the market for any of the tiers proposed in this study for GAMT screening, nor will such tests become available in the near future. This result, and the search that was performed should be documented.

(e) the health institution provides information upon request on the use of such devices to its competent authority, which shall include a justification of their manufacturing, modification and use;

General considerations

A document should be prepared in which the justification of the manufacturing, modification and use of the tests are described, to be shared with the competent authority if requested.

Considerations for GAMT screening

These documents should be prepared for all tests necessary for GAMT screening, so that they can be shared with the IGJ inspectorate if requested.

(f) the health institution draws up a declaration which it shall make publicly available, including:

(i) the name and address of the manufacturing health institution,

(ii) the details necessary to identify the devices,

(iii) a declaration that the devices meet the general safety and performance requirements set out in Annex I to this Regulation and, where applicable, information on which requirements are not fully met with a reasoned justification therefor;

General considerations

A declaration as described in this article 5(f) should be prepared and made publicly available, for instance by publishing it on the RIVM website.

Considerations for GAMT screening

A declaration as described in this article must be prepared and published before GAMT screening can be started.

(g) as regards class D devices in accordance with the rules set out in Annex VIII, the health institution draws up documentation that makes it possible to have an

understanding of the manufacturing facility, the manufacturing process, the design and performance data of the devices, including the intended purpose, and that is sufficiently detailed to enable the competent authority to ascertain that the general safety and performance requirements set out in Annex I to this Regulation are met. Member States may apply this provision also to class A, B or C devices in accordance with the rules set out in Annex VIII;

This criterium applies to class D devices in all cases, but 'member states may apply this provision also to class A, B or C devices in accordance with the rules set out in Annex VIII'. According to the classification rules in 'annex VIII', pecifically, rule 3(m), neonatal screening tests are classified as 'class C'

(m) for screening for congenital disorders in new-born babies where failure to detect and treat such disorders could lead to life-threatening situations or severe disabilities.

It would be advisable to seek confirmation from the IGJ inspectorate for confirmation that they consider in-house neonatal screening tests as class C devices. And, if required by the inspectorate, it may be necessary to 'draw up documentation that makes it possible to have an understanding of the manufacturing facility, the manufacturing process, the design and performance data of the devices, including the intended purpose, and that is sufficiently detailed to enable the competent authority to ascertain that the general safety and performance requirements set out in Annex I to this Regulation are met', should the Dutch competent authority (i.e. IGJ) decide to 'ascertain that the general safety and performance requirements set out in Annex I to this Regulation are met'. Risk analysis should be performed and documented, for instance according to NEN-EN-ISO14971.

This paragraph shall not apply to devices that are manufactured on an industrial scale.

General considerations

Approximately 170,000 tests are performed annually for the Dutch neonatal screening. The five laboratories involved each perform on average 35,000 tests per year (the samples are not evenly distributed between the laboratories). It is unclear whether this should be considered 'industrial scale' (the legislation does not contain a definition of 'industrial scale'). This criterium in the legislation is primarily intended to prevent competition between in-house tests and commercial, certified tests. For 1st-tier tests, the Dutch competent authority (IGJ inspectorate) should be consulted, to check whether 170.000 tests annually would be considered 'industrial scale or not.

Considerations for GAMT screening

Because certified/commercial tests are not currently available on the market, for any of the tiers proposed in this study, no competition between in-house tests and commercial, certified tests exists. A periodic (e.g. yearly) re-evaluation of this criterium is necessary.

5.2 Perspectives of parents and health care providers

The parental and health-care provider´s perspectives on NBS for GAMT-D were examined:

- At several meetings with the PANDA study group (Psychosocial Aspects of Newborn Screening for Disorders Assessed; PANDA studie, Psychosociale Aspecten (uitbreiding) Neonatal Hielprikscreening) (Jun 2019, August 2019, Nov 2019, Sept 2020).
- By studying the ethical, legal, and social implications (ELSI) according to the recommended questions as proposed by Goldenberg et al. (2019) (Table 5.1 and 5.2) in collaboration with the PANDA study group.
- At the symposium for Cerebral Creatine Deficiency Syndromes on 6-7th November 2019 in Rotterdam (Supplement 1), organized by prof. Gajja Salomons, a group of international scientists in the field and three patient associations, VKS (Patiëntenvereniging VKS, Volwassenen en Kinderen met Stofwisselingsziekten), Xtraordinaire, and ACD (Association for Creatine Deficiencies) in conjuction with the SSIEM Annual Symposium 2019.
- Reviewing previous studies on the subjects.

Results: As a main conclusion of the assessments with the PANDA study group, GAMT-D was considered comparable to other metabolic diseases (e.g. PKU) in this respect. Hence, the possible impacts of the NBS for GAMT-D on parents and health-care providers are not expected to significantly differ from those already observed for other metabolic diseases in the Dutch NBS program.

No ELSI issues related to the NBS of GAMT-D were detected. One important ELSI issue related to the NBS system (Table 5.2) was identified at the start of the study in 2019, regarding the potentially prolonged turn-around time to complete all 3 tiers of GAMT-D screening compared to the target turn-around time of 5 weeks on other NBS disorders. This is not an issue, since the results will be obtained within 5 weeks. Some minor ELSI issues related to the NBS program were considered GAMT-D specific (Table 5.2), and these were either solved or investigated in this study.

At the symposium for Cerebral Creatine Deficiency Syndromes, Professor Dr. Marzia Pasquali (Professor of Pathology and Section Chief of Biochemical Genetics at ARUP Laboratories at the University of Utah, U.S.) gave a presentation on Feasibility of the NBS for GAMT-D. In the dialogue during the symposium, the families and the representatives of the patient organizations actively expressed their positive attitude towards NBS for GAMT-D. Taken into account the similarity of GAMT-D with other metabolic diseases, any surveys or questionnaires, which could even burden the parents, were not considered justified, as no new information could have been expected from their results.

There are no previous scientific publications of the potential impacts of adding GAMT-D screening in the NBS program on the families. Since the GAMT-D does not differ from other metabolic NBS diseases with this respect, the two most significant potential impacts are related to the following:

1. A significant stress can take place when the family is informed of the abnormal screening result of their newborn, as well as when they are referred, and while waiting for the confirmation of the diagnosis. Families experience significant distress and emotional insecurity especially prior to the first visit to the hospital (e.g. Blom et al.,

2021). This previous finding is expected not to be different in GAMT-D from other metabolic disorders that are identified.

2. False-positive screening results have also been associated with significant negative impact for the family, even if the newborn would be confirmed healthy at the follow-up diagnostics (Hewlett and Waisbren, 2006). Other studies have shown that most parents feel reassured after confirmation all is well (e.g. Vernooij-Langen et al., 2014). Moreover, it has been shown that parental knowledge reduces long-term anxiety by false-positive results (Vernooij-Langen et al., 2014). In 2021-2022, the PANDA study team will study the psychosocial impact and healthcare use of a false positive NBS result in the Netherlands.

Potential impacts of the different scenarios from the parental perspective. It does not differ, whether each tier is completed in regional or central laboratory. Three tier scenario with the DNA analysis is the most preferred as it already includes a confirmatory test (and only in very rare cases, undetermined results of tier 3 are expected).

Potential impacts of the different scenarios for the health care providers. None.

Conclusions. The potential psychosocial effects of GAMT-D screening do not differ from those related to the NBS of other diseases. The turn-around time of GAMT-D screening should fit into the current target time of NBS results. Additionally, the parents benefit from the minimalization of i) the false-positive rate, ii) the time period between the announcement of the abnormal test result and the first visit at follow-up clinic, and iii) in cases in which the functional analysis is required to confirm the diagnosis, the time while waiting for the confirmatory results. The three tier scenario with DNA analysis is the most optimal in this respect, since it has previously resulted in zero FPR.

Table 5.1 Evaluation of ELSI issues related to NBS for GAMT-D (questions based on Goldenberg et al., 2019).

ELSI questions		GAMT-D-specific	GAMT-D answers	REF
What are the potential ELSI of positive (abnormal) screening results related to GAMT-D?	• Do caregivers treat an infant differently when a presymptomatic diagnosis is made?	No		Salm et al., 2012
	• What are the potential harmful or beneficial effects of an NBS diagnosis on maternal–infant bonding or other family dynamics?	No		
	• Are there potential harms from subsequent diagnostic testing (which may be invasive) and treatment and how do these harms impact the net benefits of screening?	No		
What are the potential ELSI implications of *false positive* screening results related to GAMT-D?	• Do caregivers treat an infant differently as a result of receiving a false positive screen result? Are there long-lasting psychological consequences for a positive screening test in infants who do not have a condition? What is the effect of a false positive on maternal–infant bonding?	Unknown	The risk for false positive is extremly low, when tier 3 is completed by DNA analysis (the strict criteria ensures that the variants classified as likely pathonegic or pathogenic are disease-causing). In 2021-2022, the PANDA study team will study the psychosocial impact and healthcare use of a false positive NBS result in the Netherlands.	Beucher et al. 2010
	• Are there potential harms from subsequent diagnostic testing?	Yes	In rare cases, anesthesia is needed for MRS.	
What are the potential ELSI of *false negative* screening results related to a new condition?	• What is the preventable morbidity and mortality related to false negative screening results	No	When diagnosed in symptomatic phase, the neurological damaging is not be totally treatable	Stockler-Ipsiroglu, et al. 2014
	• Do normal NBS results provide false reassurance to parents (e.g., cause people to ignore symptoms of serious illness? or could cause a unnecessary diagnostic odyssey for families later in life?)	No		
What are the potential ELSI of obtaining and reporting *carrier status* related to GAMT-D?	• How does knowledge of carrier status impact the newborn/families? What is the cost/ benefit to the newborn? To the family? Of disclosing carrier status?	No	Carrier status of a newborn is not detected by GAMT-D screening.	El-Gharbawy et al., 2013
What are the potential ELSI of indeterminate results related to a condition?	• Does knowledge of potential illness provide families with reassurance that they will be able to intervene at the earliest possible moment? Does it lead to anxiety and concern about even minor symptoms?	No	Additional tests will lead to confirmation of the diagnosis within couple of days.	
	• Are there potential harms from subsequent diagnostic testing and follow-up?	Yes	In rare cases, anesthesia is needed for MRS.	

Table 5.2. Evaluation of ELSI issues related to the NBS system (questions based on Goldenberg et al., 2019).

ELSI questions		GAMT-D-specific	GAMT-D answers	REF
What are the cost or resource allocation implications for adding GAMT-D to the NBS?	• Is NBS program ready to implement the new screening test, or does it require radically new procedures, equipment, or expertise?	Yes	In-house test vs. adding GAMT-D to commercial kit is investigated. The turn-around time must fit into the current NBS program (5 weeks).	RIVM-report Mercimek-Mahmutoglu et al. 2016
	• What are the opportunity costs, of expanding to GAMT-D?	Yes	Other diseases that are not included in Neobase II kit can be performed (in parallel) with the developed in-house test to reduce costs	
	• Is there a sufficient number of clinicians trained to treat GAMT-D? What is their geographic distribution?	Yes	A sufficient number of expert clinicians in the Creatine expertise center, AmsterdamUMC (location VUmc)	
	• What is the system-wide financial cost of diagnosis and treatment? Are the prevalence and impact of the condition sufficient to justify the cost? Are there plans for long-term follow-up to judge impact of programs?	Yes	System-wide costs are investigated in this study. Expected new diagnoses are 0-2 per year based on the incidence studies. There will be a follow-up after 1 year.	
What are the health disparities or equity considerations related to adding GAMT-D to the NBS?	• Do decisions about how to screen for a condition have implications for which populations are most likely to be diagnosed (e.g., CF screening)?	No		Almeida et al., 2007
	• Are population-level results of NBS likely to affect one population in particular (e.g., reveal high rates of infectious disease or stigmatizing condition)?	No	Portuguese founder mutation	funct.analysis: Berends et al., 2017
	• What factors will influence access to confirmatory testing and treatment (e.g., health insurance, geography, culture, race/ethnicity)?	No		
What are the potential implications for public/parental trust in the NBS system or health department that might arise because of adding a new condition?	• Do false negative/false positives weaken faith in NBS programs and the ability of health departments to provide accurate and helpful information?	No	Three tier system identified no false positives in previous studies. No false negatives have been identified in previous studies.	
	• Is there transparency in the process of adding a new condition to a panel, the implementation of screening tests, and approach to follow-up and treatment?	Yes	Investigated in this study	
Does a condition raise any concerns regarding parental permission or challenges to the ethical or social justification for requiring population-based screening?	• Does the condition have such a high benefit:cost ratio that the general public and nearly all families would agree that NBS should be universal? Or would many reasonable people choose to opt out (e.g., later-onset condition with ambiguous benefits of treatment)?	No	GAMT-D screening will be population-based and universal in the Netherlands. Early detection and promt treatment prevents severe health issues (brain damaging). This is analogical to other NBS disorders.	Stockler-Ipsiroglu, et al. 2014; Viau et al. 2013

5.3 Cost estimate

An estimate of the costs of performing GAMT screening was performed. The initial set-up costs, recurring yearly costs for performing screening are included in the analysis. The following aspects have not been included in this cost estimate:

- The costs associated with acquiring and shipping the initial bloodspot samples are not included, because it happens already for the current screening program, and will not change regardless whether GAMT-D screening is implemented or not.
- 'Overhead' costs of the labouratoria involved

Furthermore, the effect of uncertainties on the final cost estimate was not examined (ie, a sensitivity analysis was not performed)

Tier 1

The cost of materials (consumables) for the 1st tier assay, calculated per sample analysed is shown in the following table:

1st tier									
	materials (per sample)								
		internal standard	[13C2]-GAA (46,9 ng)					€	4.12E-06
			[D3]-Creatine (149,1 ng)					€	4.08E-05
		extraction	extraction solution (125 µL)					€	0.185069
			96-well plates					€	0.065
			seals					€	0.046545
		analysis	mobile phase A					€	0.022386
			mobile phase B					€	0.04125
			HPLC column + 3X guard column					€	0.1239
		cost of materials (per sample)						***€***	***0.484196***

Table 5.3.1. Breakdown of cost of materials for tier 1 (per sample)

These costs are the same regardless of how many labs implement 1st tier screening. For 170.000 samples yearly, the total yearly cost of materials for tier 1 is estimated to be €82.313.

The cost of equipment maintenance for tier 1 was not included, because maintenance is already performed for the current screening program.

The costs of labour necessary for performing tier 1 analyses are shown in the table below:

labour (per week, per lab)						Tarif		
	punching		60	min		middle	€	111
	extraction	preparing IS	10	min		middle	€	18.5
		Incuberen BSP + overpipetteren	120	min		middle	€	222
	HPLC preparation	preparing mobile phase	10	min		middle	€	18.5
		fit and equilibrate column	15	min		middle	€	27.75
	Process/check data	integrate/check peaks	30	min		middle	€	55.5
		enter data into Neonat/LIMS	30	min		middle	€	55.5
		supervision	48	min		high	€	101.6
	Supervision/coordination		60	min		high	€	127
	cost of labour (for one lab, one week)						€	**737.35**
	cost of labour (for 5 labs, one year)						**€**	**191711**
	cost of labour (per sample analysed)						**€**	**1.127712**

Table 5.3.2. Breakdown of cost of labour for tier 1 (per week, per lab)

Tier 2

The second tier may be conducted in two or more of the screening laboratory (a mimimum of two is assumed, to ensure backup in case of instrument down-time). The costs for scenarios in which 2, 3 or all 5 of the screening labouratories was estimated.

For the cost estimate, it was assumed that 0.6% (or less) of samples in tier 1 will be (false) positive, yielding a maximum workload of 0.6%*170.000 = 1020 samples per year for the 2nd tier, or 20 per week.

Equipment

Implementation of the 2nd tier assay in the reference labouratory in Bilthoven would require expansion of the LC-MS/MS capacity in that labouratory. The workload for LC-MS/MS assays in the other four screening labs is lower, and allows for implementation of the 2nd tier GAMT screening without additional equipement. For this cost estimate, it was assumed that the 2nd tier assay will be implemented in Bilthoven in all scenarios, so a rough estimate of the cost of an additional LC-MS/MS instrument is included. In interpreting the impact of the cost of this machine, it should be considered that the machine will likely be used for the screening on other diseases as well, and therefore assigning its costs should be distributed. However, because the exact cost of an LC-MS/MS and what it will be used for exactly are uncertain, this is therefore not taken into account into the cost estimate here.

Equipment (annual depreciation)								
	additional LC-MS/MS in reference lab	annual depreciation					€	60000
	LC-MS/MS equipment cost (per sample analysed)						€	***58.82353***
	LC-MS/MS equipment cost (per sample screened)						€	***0.352941***

Table 5.3.3: rough estimate of the cost of an additional LC-MS/MS instrument. The exact costs depend on many uncertain factors (eg. type of instrument, type of contract (lease/buy), etc), and should therefore be interpreted with caution.

Consumables

consumable materials (per sample)								
	Sample prep	Nitrogen					€	8E-08
		Mobile phase A (100 µL)					€	0.000666
	Analysis	Mobile phase A					€	0.05328
		Mobile phase B					€	0.019912
		HPLC column					€	0.04975
	cost of consumable materials (per sample)						€	***0.123608***

Table 5.3.4: Breakdown of the cost of materials for tier 2, per sample

These costs are essentially the same regardless of how many labs implement 2nd tier screening. For 1020 samples yearly, the total yearly cost of consumables for tier 2 is estimated to be €126.

Labour

labour (per run)						Tarif		
	Evaporate and redissolve	evaporation		10	min	middle	€	18.5
		redissolving in mobile phase		10	min	middle	€	18.5
	HPLC preparation	prepare mobile phase		30	min	middle	€	55.5
		install and equilibrate column		15	min	middle	€	27.75
		set up analysis		10	min	middle	€	18.5
	Process/check data	integrate/check peaks		20	min	middle	€	37
		Export to excel		30	min	middle	€	55.5
	Supervision/coordination			30	min	high	€	63.5
	cost of labour (per run)						**€**	**294.75**
	cost of labour (2 labs, one run per week)						€	**589.5**
	cost of labour (3 labs, one run per week)						€	**884.25**
	cost of labour (5 labs, one run per week)						€	**1473.75**

Table 5.3.5: Breakdown of the cost of labour for tier 2, per run. Furthermore, the total costs of labour for performing one run per week in 2, 3 or 5 labs are shown.

The number of samples included in an analysis run has only very minor effect on the time (labour) necessary for performing 2nd tier analysis, and it was therefore considered negligible in this cost estimate. In contrast, for calculating the total costs of labour for 2nd tier analyses (on a national level), it is important to consider how many 2nd tier runs will be performed weekly. For instance, performing two analysis runs per week cost twice as much as doing only a single run per week, even if the total number of samples analyzed in that week is the same. Similarly, distributing the samples between four labs instead of two would require all four labs to perform analysis, (each lab analyzing fewer samples), again multiplying the cost per sample. Therefore, the table includes cost estimates for three different scenarios in which respectively 2, 3, and 5 labs perform 2nd tier testing.

Tier 3

The third tier consists of sequencing the *GAMT* gene. The second tier is not expected to yield false-positive results, and the yearly amount of of 3rd analyses is therefore expected to be equal to the incidence of GAMT-D. Estimates for GAMT-D incidence range between 1:120.000 and 1:770.000 , en is estimated at 1:250.000 by the AMC-VU for the Dutch population. In the past, up to two patients have been diagnosed in the Netherlands in a

single year, but this is unusual. For this cost estimate, a rate of 1 patient per year was assumed.

3e tier									
	Cost of analysis in external lab (per sa	expected: 1 or less per year						€	800
	cost of third tier (per year)							**€**	**800**

Table 5.3.6: costs of 3rd tier analysis (by an external laboratory)

Referral, Follow-up diagnostics and treatment

The costs for referral of the screen-positive neonate is shown in table 5.3.7

cost of refereal								
DVP			30	min	low	€	49
medical adviser			60	min	high	€	127
	cost of referals (per year)					**€**	**176**

Table 5.3.7: costs of referring a screen-positive neonate

The costs of follow-up diagnostics and treatment are unknown.

Startup costs

Implementation of GAMT-D screening will also involve some costs that are only made once (ie they do not recur weekly/yearly). The table below provides a cost estimate for these 'setup costs'

startup costs								
NEONAT								
	aanpassing NEONAT				Tarief	€	13.440,00	(ex btw)
	ondersteuning wetenschappelijk medewerker GZB	30	uur	hoog	€	3810	
	ICT ondersteuning	4	uur	hoog	€	508	
Praeventis								
	aanpassing Praeventis					€	30.323,84	(ex btw)
	uren DVP medewerkers					€	30000	
Neorah								
	ervaren ICT beheerder	40-50	uur	hoog	€	6350	(50 uur)
	uren RIVM en TNO	20-40	uur	hoog	€	5080	(40 uur)

Table 5.3.8: setup costs

Overview

An overview of the total costs of each tier is summarized in table 5.3.9:

		1 tier	2nd tier			3rd tier
		5 labs	*2 labs*	*3 labs*	*5 labs*	*1 lab (external)*
equipment (depreciation)	per year		€ 60,000	€ 60,000	€ 60,000	
materials (consumables)	per year	€ 82,313	€ 126	€ 126	€ 126	
labour	per year	€ 191,711	€ 30,654	€ 45,981	€ 76,635	
subtotal	per year	**€ 274,024**	**€ 90,780**	**€ 106,107**	**€ 136,761**	**€ 800**
estimated number of samples	per year	170000	1020	1020	1020	1
total per sample analysed	per sample	**€ 1.61**	**€ 89.00**	**€ 104.03**	**€ 134.08**	**800**
total per sample screened	per sample	**€ 1.61**	**€ 0.53**	**€ 0.62**	**€ 0.80**	**€ 0.0047**

Table 5.3.9: overview of total costs for each tier

6 Optimal diagnostic work-up for aberrant screening results and treatment and follow-up for a neonate with a confirmed diagnosis

6.1 Referral procedure for newborns with a positive screening result

Newborns identified to have an abnormal screening result after the third-tier test are screen positive. This refers to a detection of a presumed biallelic pathogenic or likely pathogenic *GAMT* gene variants in the mutation analysis.

Screen positive infants are referred to the nearest Center for metabolic Diseases (in Amsterdam, Groningen, Maastricht, Nijmegen, Rotterdam, and Utrecht) for the diagnostic work-up and follow-up. The referral should be completed on the next working day (in line with the protocol for MPS I) as the positive screening results in the third-tier test are detected in the screening laboratory.

The referral process proceeds according to the procedure for referral on Dutch NBS program. The screening laboratory first reports the abnormal result to the screening center/RIVM-DVP regional office, from which the medical advisor contacts the family GP and metabolic pediatrician in the nearest Center for metabolic Diseases. If the contact between the medical advisor and metabolic pediatrician takes place outside office hours, the referral goes through the GP on call (or the HAP huisartsenpost). GP will visit the family to inform about the referral on the same day. After this, the family will visit the metabolic pediatrician for follow-up diagnostics at the same or the following working day.

6.2 Procedure for diagnostic work-up after a positive screening result

There are two scenarios for the diagnostic work-up depending of the pathogenicity of the detected *GAMT* variants in the third-tier screening.

1. A detection of presumed homozygous or compound heterozygous pathogenic *GAMT* gene variant(s) is the confirmatory diagnostic test for GAMT-D. The child enters the GAMT-D treatment care path. Genetic testing of the parents is needed to confirm

homozygosity or compound heterozygosity. This can be done during the follow-up by taking blood samples of the parents in any of the centers and shipping those to AmsterdamUMC.

2. In case that *GAMT* variant(s) have yet undetermined pathogenicity, GAMT-D is highly suspected, but the diagnosis needs to be confirmed with additional examinations. Diagnostic follow-up samples will be collected of the infant (blood and urine samples as well as brain MRS, if available) to examine the clinical phenotype and functional analysis performed in AmsterdamUMC is warranted for the interpretation of the causation. All additional samples can be taken at the first visit with the metabolic pediatrician. Depending on the clinical characteristics of the patient, treatment can be started after collecting the samples or having the confirmatory results. Two options are possible for the functional analysis depending of the detected variants (the laboratory will choose):
   i) Functional analysis measuring GAMT enzyme activity in lymphocytes (Berends et al., 2017). This necessitates an additional blood sample of the screen positive infant, which is taken at the first policlinical visit and shipped to AmsterdamUMC. Abnormal enzyme activity confirms GAMT-D. Normal enzyme activity indicates false positive screening result.
   ii) Functional characterization using overexpression studies are possible for missense variants (Mercimek-Mahmutoglu et al., 2016). This does not require an additional blood samples of the patient.

In every case, diagnostic follow-up tests will be performed according to the carepath: urine and plasma GAA and Cr are measured at the first visit with the metabolic pediatric and the patient undergo brain MRI/MRS imaging (Appendix 3). Most academic centers will be able to measure GAA and Cr, but otherwise frozen urine and plasma samples can be shipped to AmsterdamUMC. The true positive child will be treated according to the GAMT-D care pathway.

6.3 Procedure for treatment plan for GAMT-D

The treatment plan of GAMT-D, "Zorgpad Guanidinoacetaat Methyltranferase (GAMT-D) deficiëntie" (December 2020, behandelarenversie) (Appendix 3) was developed in 2020 by a team of physicians specialized in metabolic diseases in collaboration with other GAMT-D experts and the patient organization. It is based on the most recent scientific research, whenever possible, in combination with the experience of the expert group.

Pediatric patients with GAMT-D are always treated at a university medical center by a pediatrician specialized in metabolic diseases, who has the primary responsibility for coordinating treatment and follow-up, and a multidisciplinary team consisting of clinical geneticist, pediatric neurologist, and several other practitioners, such as a nutritionist and physiotherapist. At the time of the diagnosis, appointments are arranged at outpatient policlinics of pediatric metabolic diseases, clinical genetics, and pediatric neurology.

Therapeutical treatment of GAMT-D consists of orally administered high dose creatine supplement (creatine monohydrate). Additionally, nutritional treatment with arginine-low diet, whether or not combined with oral ornithine supplement, can be considered to

decrease the *bodily* GAA concentration. In case of seizures, anti-epileptic medication may be required.

The follow-up takes place in outpatient policlinics. The frequency of the check-up appointments is adjusted according to age, medical situation, and the overall circumstance of the patient and his family, ranging 1-8 times per year. In general, the follow-up is carried out according to the schema presented in the treatment plan.

7 Discussion, final conclusions and proposed screening method

7.1 Discussion

This study was initiated to resolve several issues identified in a study on the feasibility of implementing screening for several new diseases, including GAMT-D, in the Netherlands 'uitvoeringstoets uitbreiding neonatale hielprikscreening'. In this study, we evaluated several alternative scenarios for newborn screening of GAMT-D in the Netherlands using the knowledge from previous pilot studies, experts from screening centers, and comprehensive analyses performed in AmsterdamUMC from BSPs of healthy newborns and a previously diagnosed patient.

Briefly, it was found that the use of in-house test(s) would present several challenges. Before the study, the screening algorithm was poorly defined, and it was unclear how GAMT-D screening could be practically best implemented in the five screening laboratories, in terms of logistics and cost effectiveness. In this discussion, we evaluate to what extent a go/no-go decision on implementing GAMT-D screening in the Netherlands may now be supported by the results of this study, and what uncertainties still remain.

Condition

GAMT-D, the target disease for screening is well-defined, and no secondary findings are expected. Carriers are not detected by the proposed screening strategy. Some uncertainty remains with respect to the prevalence of the disease, which cannot be resolved at present due to its rarity.

Testing method and predictive value

A three-tier testing strategy is proposed in this study, using in-house methods for all tiers.

First tier

We examined the risks and benefits of adding GAMT-D NBS to Neobase ™ II kit and the use of a specially developed in-house method. The results showed that the addition of GAMT-D first-tier screening with Neobase ™ II kit is not feasible with the machine currently used for it. Hence, we validated an in-house LC-MS/MS first-tier method, which is technically feasible and meets the CE-IVD requirements.

The first tier method was developed and validated as a part of this study. The validation results confirm that the method is capable of measuring the markers of interest (GAA and Cr) with sufficient accuracy, precision, sensitivity, and specificity. Compared to other published first tier methods for GAMT-D screening, the method developed in this study is unique in several respects. First, no butylation step is performed: it measures underivatized GAA and Cr. Second, it utilizes LC-MS/MS instead of FIA-MS/MS to attain the required sensitivity. We recommend measuring GAA/Cr ratio only when GAA concentration exceeds the cut-off value, but in the future, it will have to be analyzed on larger patient population, whether it is beneficial or not in decreasing FPR.

The determined cut-off values and reference values in this study were in line with the previously reported. Using underivatized samples (as currently used in the Dutch NBS) did not result in any higher FPR than previously reported but resulted in a more time efficient method.

This unique first tier assay introduces several new uncertainties. Only a relatively small number of heel prick cards were measured (for the validation) in this study. Therefore, additional work for validation is still required to establish the cut-off values using the new method, by measuring much larger numbers of heel prick cards (and patient cards, if available). Similarly, data from a larger number of samples is needed to establish accurate estimates for false positive and false negative rates, and positive predictive value. Furthermore, these numbers are also necessary to estimate the expected workload for the second tier test more accurately. A maximum rate of false-positive results from tier 1 was estimated at 0.6%. If the actual FPR is very different, it may impact the cost estimate and logistics.

Furthermore, the robustness of the new method over longer periods of time (months) has not been tested. It would be advisable to confirm that the new assay performs consistently over several months or longer, using heel prick cards and QC monitoring samples. QC monitoring samples for daily QC checks should be prepared (and QC criteria), which have not been described in this study.

Finally, considering that the GAMT-D first tier method, as used in this study, is conducted independently from the current screening, on a separate blood spot, it should be evaluated, whether this new (underivatized) method offers sufficient benefits over the internationally established (butylated) method for GAMT testing to warrant its use. Both methods are in-house methods, but both have advantages and drawbacks. Butylation dramatically increases personnel time and is very cost-ineffective. Before implementing GAMT-D screening, it should be evaluated, which assay would be the best option.

Second tier

For the second tier test, this study proposes to use a previously published method ([Mercimek-Mahmutoglu et al., 2016]), which has been used for several years in different screening programs internationally, and as such is a well-established method. Importantly, however, a notable difference between the published method and the testing strategy in this study is that the published second tier method starts with butylated material that remains from the first tier. To adopt the published method as the second tier method, a modification from the published protocol is not necessary and has been implemented in this study. It should be noted that already 500 BSPs from the Dutch NBS program have been tested with the second tier method (Mercimek-Mahmutoglu et al., 2016).

The (internal) standards for the tier 1 and 2 assays are available from more than one supplier and interchangeable and storaging 100 mg of each standard in powder form serves as an adequate back-up resource. The proficiency samples are available from the CDC.

Third tier

In the proposed screening strategy, the third tier consists of sequencing the *GAMT* gene. Considering that the specificity of the second tier was reported to be 100%, it may also be worthwhile to consider an algorithm without this third tier. It should be noted that already 500 BSPs from the Dutch NBS program have been tested with the third tier (Mercimek-Mahmutoglu et al., 2016) method.

Alternatively, a screening algorithm in which the second tier LC-MS/MS analysis is removed, and 1-st tier positive samples go directly to sequencing might be considered. This might be favourable if the number of false positive first tier tests and the costs of DNA sequencing are low. Considering the uncertainties in these numbers, however, it is at present difficult to evaluate these alternative scenarios.

Based on the previous studies, a three-tier method measuring underivatized GAA (and possibly also GAA/Cr ratio) with LC-MS/MS as a first tier, GAA (and possibly also GAA/Cr ratio) with LC-MS/MS as a second tier, and DNA analysis as a third tier was most efficient in minimalizing false positive rate and we recommend it for the Dutch NBS as well. If during a year's follow-up the predicted FPR of tier 1 turns out to be significantly lower than expected, the need for tier 2 could be re-evaluated.

In-house tests

The proposed screening algorithm relies on in-house test for all tiers. To inform a go/no-go decision on GAMT-D implementation, the inspectorate (IGJ) should be consulted to request their view on the following issues:

- Can reagents and QC samples be distributed between the screening laboratories? (article 5a)
- Would in-house neonatal screening tests be class C devices?
- Would performing 170.000 tests per year be considered 'industrial scale'?

If a commercial testing method becomes available for GAMT-D screening, it would be necessary to use the certified commercial testing method in the heelprick screening program due to current legislation (CE-IVD). At the end of this study, we have learned that at least one commercial party is in the process of developing a first tier test for GAMT-D screening (apparently based on our studies), for which they intend to seek IVD certification. It remains unclear in what schedule this test should be expected to become available. A market survey of commercial tests for GAMT-D screening should be conducted periodically (e.g. yearly).

Costs

A costestimate of screening was included in this study, but some uncertainties remains:

- The maximum expected FPR of 0.6% was assumed for the first tier test. As mentioned, however, because the proposed first tier test is different from internationally used tests, it is unclear whether this estimate is accurate. If a much larger FPR in the first tier will be observed, the costs of tier 2 may increase.
- The LC-MS/MS capacity in Bilthoven need to be expanded to accommodate GAMT-D screening. The cost of an additional LC-MS/MS instrument will probably

be shared between different diseases, but it is unclear between how many/which. As a result, the equipment cost of tier 2 is probably considerably overestimated.

Summary

Summarizing, our study has addressed most of the issues raised in the feasibility study on the implementation of GAMT-D screening. However, uncertainty remains regarding the issues described above. Further work is required to support a go/no go decision on the implementation of GAMT-D screening.

7.2 Final conclusions and proposed screening method

Based on the results of this study, we propose the GAMT-D screening to be integrated in the Dutch NBS program as presented in Table 7 and Figure 7. This scenario is legally and ethically justifiable and provides the most robust and time- and cost-efficient method for improving GAMT-D patients´ lives

From the perspective of the parents and health-care providers, GAMT-D is similar to many other screened inborn errors of metabolism and this study did not identify any unresolved issues.

The referral process proceeds according to the procedure for referral on Dutch NBS program. A GAMT-D positive screening result following the three-tier system is expected to confirm the diagnosis in the vast majority of the patients, but in rare cases, the confirmation (or exclusion) of the disease will be achieved by diagnostic follow-up samples and functional analysis of GAMT enzyme activity. The diagnostics, treatment, and follow-up are carried out with the GAMT-D care path. The metabolic pediatrician in one of the Dutch Centers for metabolic diseases will take the main responsibility in consultation with the Creatine Expert Center, AmsterdamUMC.

Table 7. The proposed screening scenario for GAMT-D NBS in the Netherlands.

TIER	Measured metabolites	Method	Screening laboratories	Number of expected samples per lab
1	GAA, GAA/Cr ratio	Validated in-house method, LC-MS/MS	All screening labs	640 per week
2	GAA, GAA/Cr ratio	Validated in-house method, LC-MS/MS	1 central lab and one backup machine or lab	0-20 per week [1]
3	DNA sequencing of the complete *GAMT*open reading frame	Described in: Mercimek-Mahmutoglu et al., 2016.	1 central lab	0-3 per year

[1] based on the results of the FPRs after the first tier in previous (0.02-0.6%) and this study (0.4%).

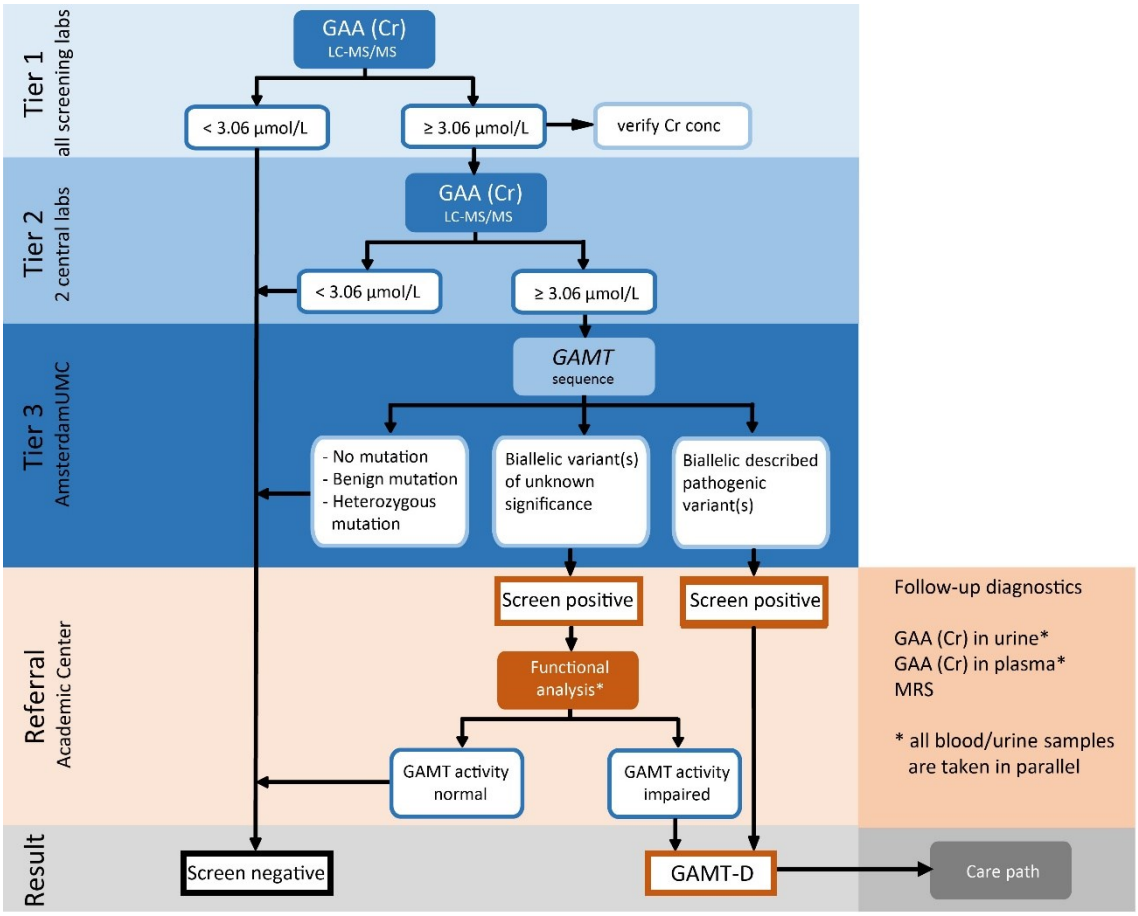


Figure 7. Flowchart of the proposed three-tier GAMT-D screening procedure, diagnostics and treatment. Tier 2 LC-MS/MS entails a more specific LC separation in comparison to tier 1.

References

Almeida LS, Vilarinho L, Darmin PS, Rosenberg EH, Martinez-Muñoz C, Jakobs C, Salomons GS. A prevalent pathogenic GAMT mutation (c.59G>C) in Portugal. Mol Genet Metab. 2007;91:1.

Berends LM, Struys EA, Roos B, Holwerda U, Jansen EEW, Salomons GS, Wamelink MMC. Guanidinoacetate Methyltransferase Activity in Lymphocytes, for a Fast Diagnosis. JIMD Rep. 2017;37:13.

Beucher J, Leray E, Deneuville E, et al. Psychological effects of false-positive results in cystic fibrosis newborn screening: a two-year follow-up. J Pediatr. 2010;156:771.

Blom, M., Bredius, R.G.M., Jansen, M.E. et al. Parents' Perspectives and Societal Acceptance of Implementation of Newborn Screening for SCID in the Netherlands. J Clin Immunol 2021;41:99.

Bodamer, O.A., Iqbal, F., Muhl, A., et al. Low creatinine: the diagnostic clue for a treatable neurologic disorder. Neurology 2009;72, 854.

CDC Newborn Screening Quality Assurance Program. https://www.cdc.gov/labstandards/nsqap_resources.html. Accessed in: 5th April 2021.

Cerebral Creatine Deficiency Sympisium 2019. https://creatineinfo.org/ccds-rotterdam-workshop-english/ Accessed in: 5th April 2021.

El-Gharbawy AH, Goldstein JL, Millington DS, et al. Elevation of guanidinoacetate in newborn dried blood spots and impact of early treatment in GAMT deficiency. Mol Genet Metab. 2013;109:215.

Goldenberg AJ, Lloyd-Puryear M, Brosco JP, et al. Bioethics and Legal Workgroup of the Newborn Screening Translational Research Network. Including ELSI research questions in newborn screening pilot studies. Genet Med. 2019;21:525.

Hall PL, Marquardt G, McHugh DM, Currier RJ, Tang H, Stoway SD, Rinaldo P. Postanalytical tools improve performance of newborn screening by tandem mass spectrometry. Genet Med. 2014 Dec;16(12):889-95. doi: 10.1038/gim.2014.62. Epub 2014 May 29. PMID: 24875301; PMCID: PMC4262759.

Hewlett J and Waisbren SE. A review of the psychosocial effects of false-positive results on parents and current communication practices in newborn screening J Inherit Metab Dis. 2006;29:677.

Malvagia S, Forni G, Ombrone D, la Marca G. Development of Strategies to Decrease False Positive Results in Newborn Screening. Int J Neonatal Screen. 2020;6:84.

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.

Mercimek-Mahmutoglu S, Sinclair G, van Dooren SJM, et al. Guanidinoacetate methyltransferase deficiency: first steps to newborn screening for a treatable neurometabolic disease. Mol Genet Metab. 2012;107:433.

Mercimek-Andrews S, Salomons GS. Creatine Deficiency Syndromes. 2009 Jan 15 [updated 2015 Dec 10]. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH,

Stephens K, Amemiya A, editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993–2020. PMID: 20301745.

Mercimek-Mahmutoglu S, Pop A, Kanhai W, Fernandez Ojeda M, Holwerda U, Smith D, Loeber JG, Schielen PC, Salomons GS. A pilot study to estimate incidence of guanidinoacetate methyltransferase deficiency in newborns by direct sequencing of the GAMT gene. Gene. 2016;575:127.

Pasquali M, Schwarz E, Jensen M, Yuzyuk T, DeBiase I, Randall H, Longo N. Feasibility of newborn screening for guanidinoacetate methyltransferase (GAMT) deficiency. J Inherit Metab Dis. 2014;37:231.

Pitt JJ, Tzanakos N and Nquyen T. Newborn screening for guanidinoacetate methyl transferase deficiency. Mol Genet Metab. 2014;111:303.

RIVM Feasibility study into expanding neonatal heel prick test. 2017

Salm N, Yetter E, Tluczek A. Informing parents about positive newborn screen results: parents' recommendations. J Child Health Care. 2012;16:367.

Schulze A, Hoffmann GF, Bachert P, et al. Presymptomatic treatment of neonatal guanidinoacetate methyltransferase deficiency. Neurology 2006;67:719.

Sinclair GB, van Karnebeek CDM, Ester M, et al. A three-tier algorithm for guanidinoacetate methyltransferase (GAMT) deficiency newborn screening. Mol Genet Metab. 2016;118:17.

Stockler-Ipsiroglu S, van Karnebeek C, Longo N, et al. Guanidinoacetate methyltransferase (GAMT) deficiency: outcomes in 48 individuals and recommendations for diagnosis, treatment and monitoring. Mol Genet Metab. 2014;111:16.

Utah Department of Health, Dec 30, 2020 https://health.utah.gov/featured-news/newborn-screening-in-utah-identifies-first-baby-born-with-gamt

Verkerk PH. Twintig jaar landelijke screening op fenylketonurie in Nederland. Landelijke Begeleidingscommissie PKU [20-year national screening for phenylketonuria in The Netherlands. National Guidance Commission PKU]. Ned Tijdschr Geneeskd. 1995;139:2302.

Vernooij-van Langen AM, van der Pal SM, Reijntjens AJ, Loeber JG, Dompeling E, Dankert-Roelse JE. Parental knowledge reduces long term anxiety induced by false-positive test results after newborn screening for cystic fibrosis. Mol Genet Metab Rep. 2014;1:334.

Viau KS, Ernst SL, Pasquali M, Botto LD, Hedlund G, Longo N. Evidence-based treatment of guanidinoacetate methyltransferase (GAMT) deficiency. Mol Genet Metab. 2013;110:255.

Supplementary Figure 1. Tier 1 screening for GAMT-D

Tier 1 screening for enzyme deficiencies associated with GAMT (guanidinoacetate methyltransferase) using flow-injection MS/MS

Collin Hill[1], Tsun Au Yeung[1], Joe Trometer[1]
[1]PerkinElmer, Waltham, MA

BACKGROUND and INTRODUCTION

- Newborn screening labs are responsible for the rapid testing of numerous samples a day for various inborn errors of metabolism.
- Typically, this is done by a flow-injection MS/MS approach (Tier 1), which allows for quick turn around time for disease diagnosis.
- One such disease is GAMT (guanidinoacetate methyltransferase) deficiency, which is caused a loss of function in the GAMT enzyme, leading to the inability to convert guanidinoacetate into creatine.
- GAMT deficiency is characterized by intellectual development difficulties, weak muscle tone and epileptic episodes.
- The treatment for the disease is oral creatine supplementation and manages the disease, speaking to the need for rapid identification in newborns.
- Here we describe a flow-injection based methodology for identification of contrived positive samples using the PerkinElmer QSight® 210 MD LC-MS/MS.

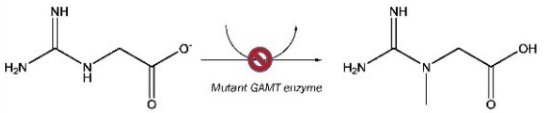


Figure 1 – Mechanism of GAMT deficiency

MATERIALS AND METHODS

- Samples that were contrived positive for GAMT deficiency were obtained from the CDC and tested along non-spiked DBS
- 3.1 mm punches were incubated with 125 µL of extraction solvent (84% ACN, 0.1% formic acid at 37°C for 30 minutes.
- 100 µL was transferred to a plate, dried down and reconstituted with water, 0.1% formic acid
- 20 µL of extract was injected using a flow-injection method, with 84% ACN, 0.1% formic acid as the mobile phase.
- PerkinElmer QSight® 210 MD LC-MS/MS was operated in positive mode for analysis of GAA and creatine.

RESULTS

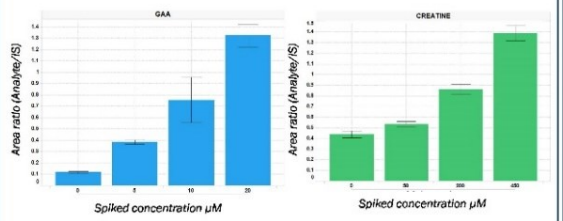


Figure 2 – Dose dependent increases of GAA and creatine, normalized to IS

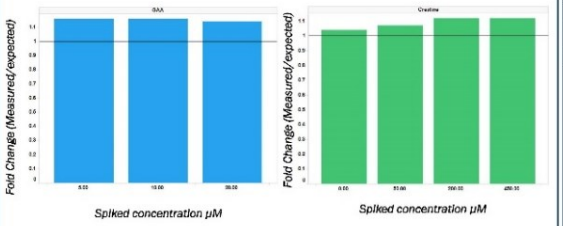


Figure 3 – Comparison of GAA and creatine against expected values

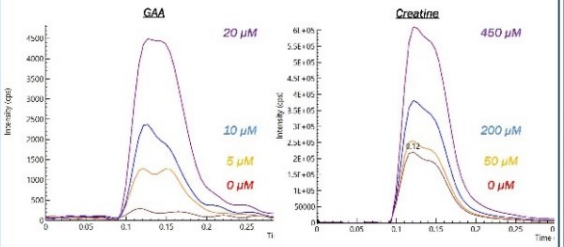


Figure 4 – Example XICs for GAA and creatine at the four spiked levels.

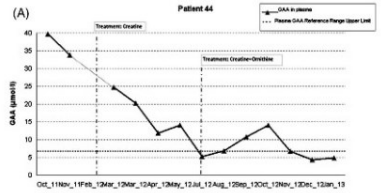


Figure 5 – An external literature article showing elevated GAA, and subsequent decrease when treated with creatine and ornithine[1]

- **Figure 1** demonstrates the mechanism of action of the GAMT enzyme, which is unable to catalyze the reaction GAA to creatine, resulting in GAA accumulation and creatine deficiency.
- **Figure 2** shows the concentration-dependent increase in spiked levels.
- **Figure 3** shows calculated values against expected values, demonstrating that the assay is accurate within 16% for both analytes, across all levels.
- **Figure 4** shows XICs from flow-injection for the analytes at all concentrations
- **Figure 5** shows external data from literature, demonstrating increased GAA and its subsequent decrease upon treatment

CONCLUSION

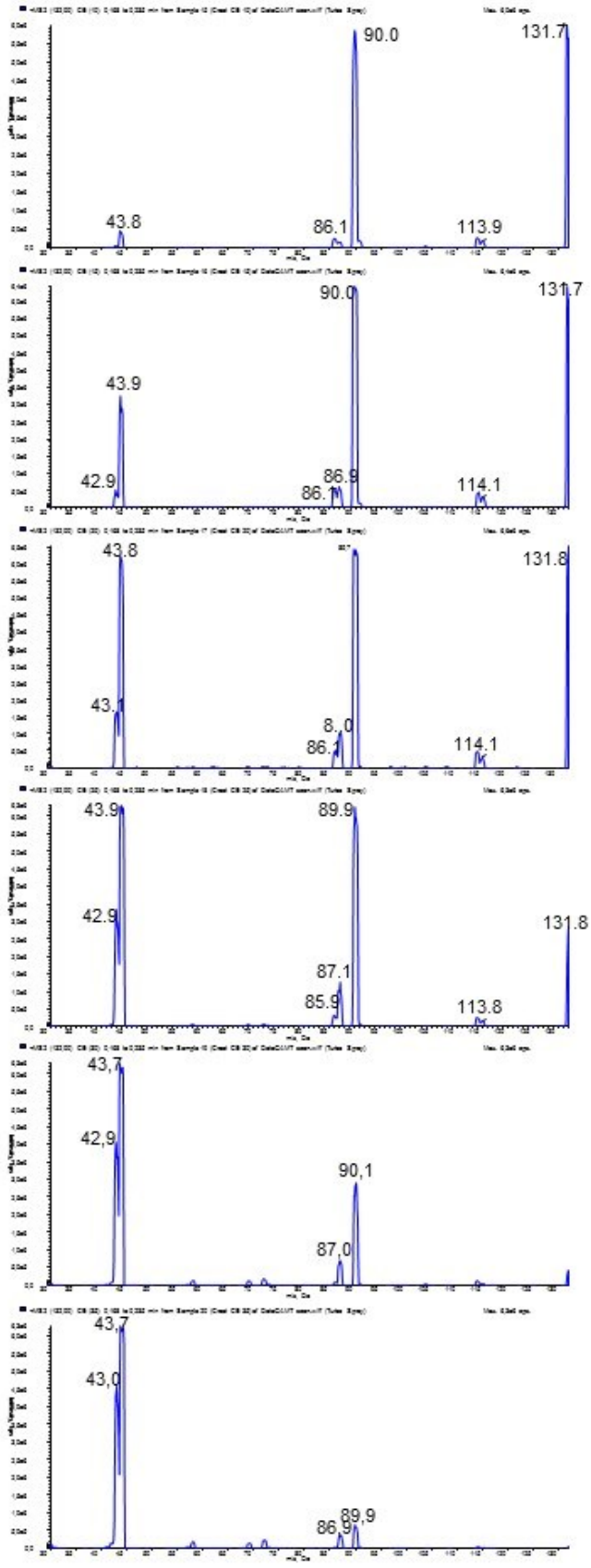
- Neonatal screening labs require quick and rapid identification of neonates who may harbor an inborn error of metabolism
- GAA and creatine measurements can be accurately done on a PerkinElmer QSight® 210 MD LC-MS/MS, and it is able to measure within clinically relevant ranges[1].

REFERENCES

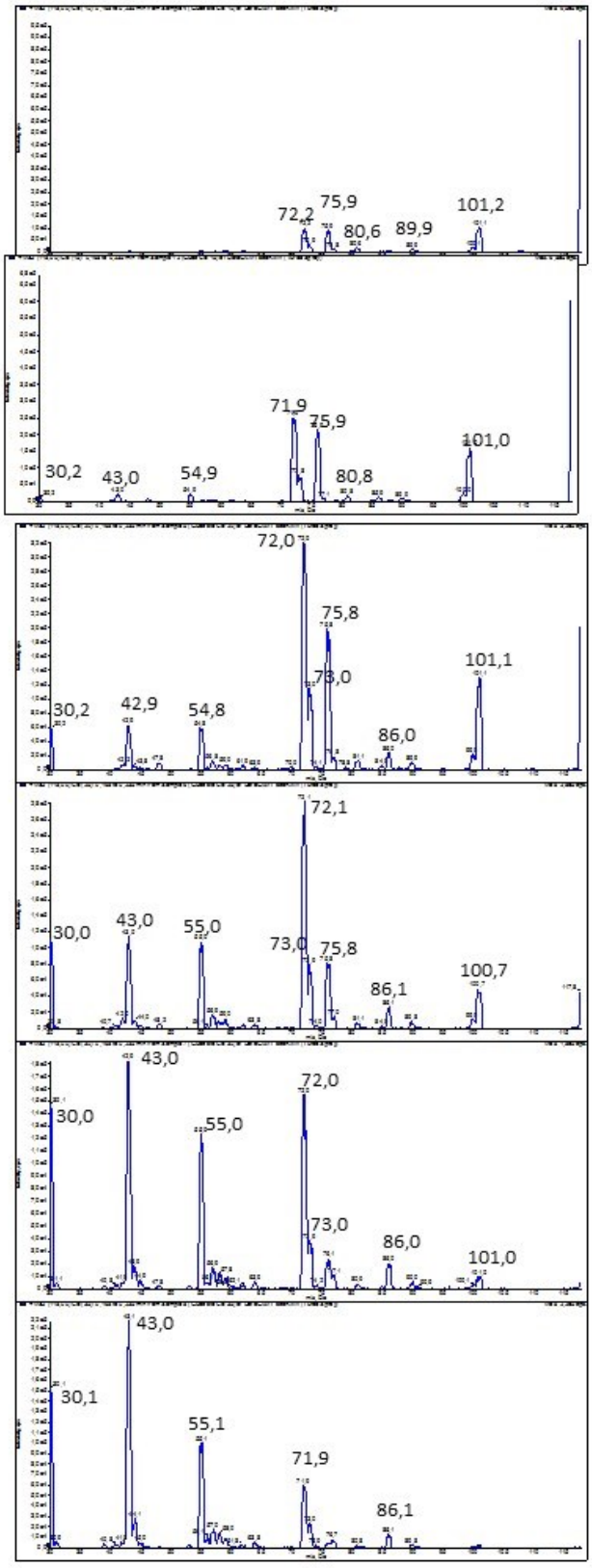
1.) Sylvia Stockler-Ipsiroglu et al. Guanidinoacetate methyltransferase (GAMT) deficiency: Outcomes in 48 individuals and recommendations for diagnosis, treatment and monitoring, Molecular Genetics and Metabolism,Volume 111, Issue 1,2014,Pages 16-25.

For research use only. Not for use in diagnostic procedures.

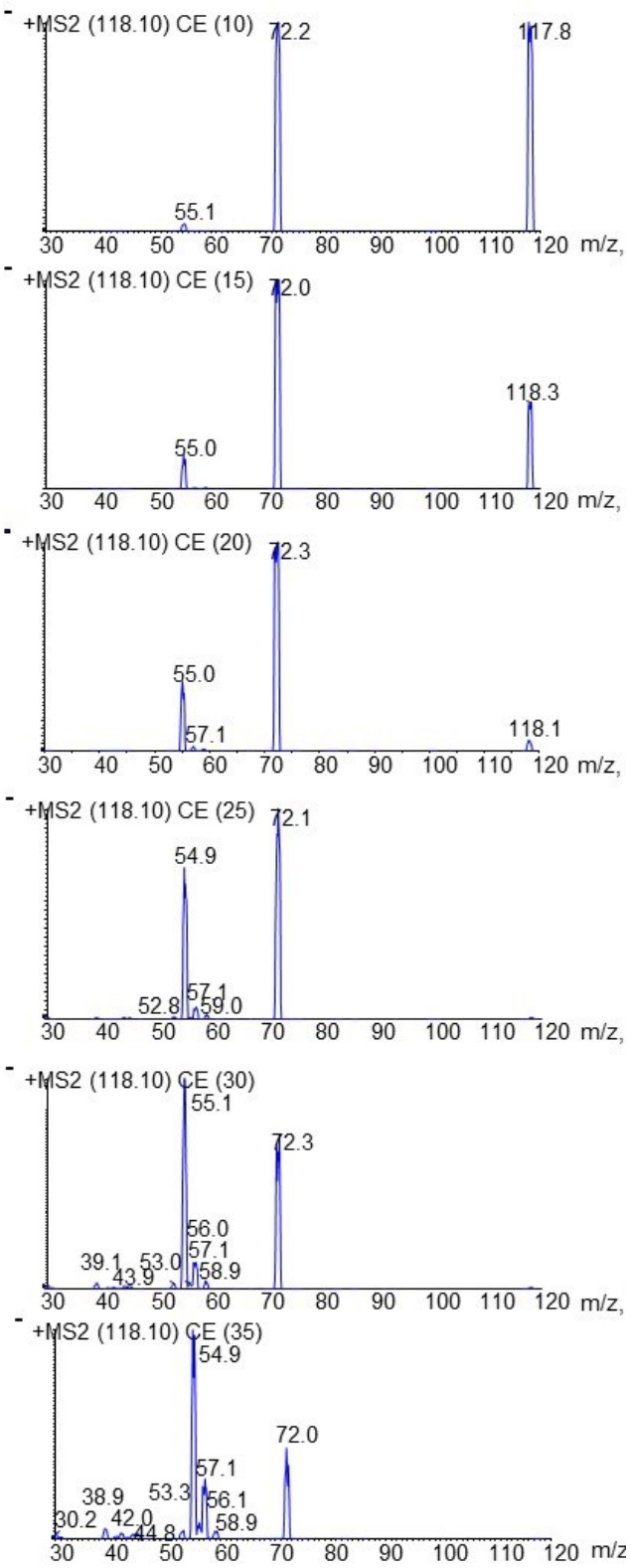
Supplementary Figure 2. Product spectra for creatine for various collision energies (positive electrospray LC-MS/MS)



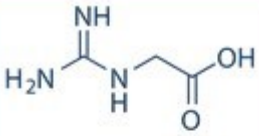
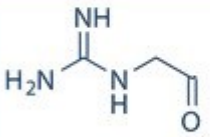
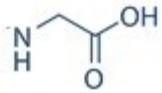
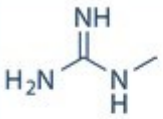
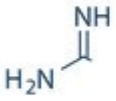
Supplementary Figure 3. Product spectra for GAA for various collision energies (positive electrospray LC-MS/MS)



Supplementary Figure 4. Product spectra for Valine for various collision energies (positive electrospray LC-MS/MS)



Supplementary Figure 5. Mass fragmentation of GAA (Positive electrospray LC-MS/MS)

GAA	
Fragments	
m/z 118 → 101	
m/z 118 → 76	
m/z 118 → 72	
m/z 118 → 43	
m/z 118 → 30	rearrangement

Supplementary Table 1. Suppliers of Cr for GAMT-D assay

Creatine CAS 57-00-1

Purity	Supplier	Code	Unit	Price
98 %	abcr	AB115087	25 g	€ 45.30
98 %	AK Scientific	F844 (70844)	5 g	$ 14
98 %	Alfa Aesar	A17477	25 g	€ 28.00
95+%	Ambeed	A518760	25 g	$ 9
95+%	Ark Pharm, Inc.	AK113717	25 g	$ 10
95+%	BLDPharm	BD122358	5 g	$ 8
	Carbosynth	FC01727	100 g	€ 36.00
	Enamine	EN300-1692957	0.1 g	$ 21
98 %	Fluorochem	240557	100 g	£ 13
≥98%	Glentham Life Sc.	GV1128	25 g	€ 20.60
95 %	Mcule	3625061513	1 mg	$ 13
98 %	Oakwood	240557	1 g	$ 10
>98%	TCI	C3610	25 g	£ 27
98 %	Energy Chemical	A04A17477	25 g	$ 47.38
NA	Toronto Res. Chem.	C781483	500 mg	$ 40
95 %	LabNetwork Selection	20016332	100 g	$ 19.96
95 %	LabNetwork Selection	20031723	250 g	$ 33.10
98 %	ChemScene	CS-W011104	5 g	$ 50
99 %	Energy Chemical	A01A011407	25 g	$ 5.04
98 %	Fluorochem	240557	1 g	$ 14
98 %	Yolne	SY457001	25 g	$ 19.15
98 %	Shanghai yuanye	ID: S48497	25 g	$ 14.62
99 %	Titan (Adamas)	66435	25 g	$ 16.13
98 %	Alfa Aesar China	A17477	25 g	$ 48.05

Creatine CAS 6020-87-7 (monohydrate)

Purity	Supplier	Code	Unit	Price
98 %	AK Scientific	F890 (70890)	5 g	$ 14
	Sigma Aldrich (MERCK)	C4255	50 g	€ 50.30
≥98%	Sigma Aldrich (MERCK)	C3630	100 g	€ 73.20
>98%	TCI	C0396	25 g	£ 16

Supplementary table 2: Suppliers of isotopically labelled Creatine for GAMT-D assay

Label	Purity	Supplier	Code	Unit	Price
Met-D3	> 98%	CDN Isotopes	D-1972	0.05 g	
Met-D3	97%	Buchem BV	NA	0.25 g	€ 300
Met-D3	98%	Sigma Aldrich (MERCK)	616249	1 g	€ 1340
Met-D3		Sigma Aldrich (MERCK)	604925		
Met-D3	97%	Eurisotop	DLM-1302-0.25	0.25 g	€ 284.80
Met-D3		Medical Isotopes Inc.	D1783	50 mg	$ 290.00
Gua-13C		Sigma Aldrich (MERCK)	569925		
Gua-13C		Buchem BV	NA	0.1 g	€ 891
Gua-13C	98%	Eurisotop	CLM-7933-0.1	0.1 g	€ 876.60
D5	> 98%?	CDN Isotopes	D-7706	0.05 g	
D5		Medical Isotopes Inc.	OD75234	50 mg	$ 710

Supplementary Table 3. Suppliers of GAA for GAMT-D assay

GAA CAS 352-97-6

Purity	Supplier	Code	Unit	Price
98 %	abcr	AB131511	5 g	€ 49.60
98 %	AK Scientific	A002	5 g	$ 13
95 %	AK Scientific	J92426	5 g	$ 10
95+%	Ambeed	A515873	5 g	$ 8
95+%	Apollo Scientific Ltd.	OR27969	100 g	$ 51
95+%	Ark Pharm, Inc.	AK144672	10 g	$ 8
>98%	BIONET Key Organics	AS-10760	500 g	€ 56.09
95+%	BLDPharm	BD132232	10 g	$ 5
	Carbosynth	FG23711	25 g	€ 57.50
95 %	Chemenu	CM186221	500 g	$ 250
	ChemScene	CS-0040590	5 g	$ 50
95 %	Enamine	MFCD00004278	0.1 g	$ 17
95 %	Fluorochem	450285	100 g	£ 11.00
	Mcule	MCULE-8327589872	1 mg	$ 6.00
99 %	Sigma Aldrich (MERCK)	G11608	25 g	€ 30.00
	Synquest Laboratories	4156-1-60	25 g	$ 35.00
>97%	TCI	G0167	25 g	£ 27.00
95 %	LabNetwork Selection	20022938	25 g	$ 23.18
95 %	LabNetwork Selection	10001479	25 g	$ 23.18
98 %	ChemScene	CS-0040590	5 g	$ 50
99 %	Energy Chemical	A01A010543	25 g	$ 9.24
98 %	Accela ChemBio Inc.	SY015122	5 g	$ 10
98 %	ASTATECH, INC	AC2229	100 g	$ 53
99 %	3A Chemicals Co., Ltd.	A44358	25 g	$ 10.42
98 %	Shangai yuanye	S64770	25 g	$ 12.10
98 %	Topscience Biochem CO. LTD	T4238	25 mg	$ 24.86
99 %	9 Ding Chemical	G012A	25 g	$ 10.42
97 %	9 Ding Chemical	G0167	25 g	$ 37.67
98 %	Titan (Adamas)	50928	25 g	$ 8.40
95 %	Fluorochem	450285	10 g	$ 14
95 %	Coolpharm	KH-36878	100 g	$ 31.92
NA	Spectrum Chemica	G3327	25 g	$ 46.75
99 %	Nanjing sunlida bio	SLD02792	25 g	$ 36.96
NA	Pfaltz & Bauer, Inc.	G05380	25 g	$ 69.35

Supplementary Table 4. Suppliers of isotopically labeled GAA for GAMT-D assay

Label	Purity	Supplier	Code	Unit	Price
1,2-13C2; 3-15N	97%	Cambridge Isotope Lab	CNLM-8300		
2,2-D2	98%	Cambridge Isotope Lab	DLM-9998	100 mg	€346
2,2-D2	> 98%	CDN Isotopes	D-6320	0.1	
13C2		Medical Isotopes Inc.	C4588	1 mg	$ 2000
13C2		Toronto Research Chemicals	G821252	1 mg	$155
13C2		Coompo Research Chemicals	C239476	1 mg	$140
13C2		AmsterdamUMC		25 mg	€350

Supplement 1. CCDS Symposium Program

Inborn Cerebral Creatine Deficiency Syndrome Symposium

September 6-7th 2019

Rotterdam, the Netherlands

Venue: De Boelen

Dit symposium, dat aansluit op SSIEM (Society for the study of inborn errors of metabolism), wordt georganiseerd door professor Gajja Salomons AMC/AMU), een groep internationale wetenschappers op het gebied van creatine en de patiëntenverenigingen VKS, Xtraordinaire en ACD. Het symposium geeft de kans aan patiënten, hun families, artsen en vetenschappers om de laatste en toekomstige ontwikkelingen en onderzoeksgegevens met elkaar te bespreken. Ook is het bedoeld om vooral direct contact met elkaar te hebben. Alle CCDS zullen aan de orde komen, met speciale aandacht voor Creatine Transporter Deficiëntie (CTD)., maar ook voor GAMT en AGAT.
https://www.stofwisselingsziekten.nl/evenement/cerebrale-creatine-deficientie-syndroom-ccds-workshop/

AGENDA

Friday, September, 6th:

15.45 Patient Perspectives

Xtraordinaire / ACD

Session 1: Signs and Symptoms of Cerebral Creatine Deficiency Syndromes (Chair: Carole Chehowah)

16.00 Observational Study / Databases / Prevalence

Vigilan: Observational Study of Male CTD Patients.

Judith Miller

Databases and Family Perspectives

Xtraordinaire, ACD

Prevalence in France & Netherlands

Aurore Curie / Gajja Salomons

16.45 The Basics of Behaviour: On Cognition, Development and Functioning

Sylvia Huisman

17.30 Break

17.50 Importance of Early Diagnosis: Feasibility of Newborn Screening for GAMT Deficiency

Marzia Pasquali

18.15 AGAT - The Therapeutical Target for GAMT Deficiency (and more?)
Andreas Schulze

18.40 Brain Magnetic Resonance Spectroscopy and Creatine Measurements.
Petra Pouwels

19.05 Q&As
19.30 Family Gathering and Discussion Time

Saturday, September, 7th:

9.00 Creatine Transporter Deficiency in Females
Jiddeke van de Kamp

Session 2: Animal Models for Cerebral Creatine Deficiency Syndromes (Chair: Jiddeke van de Kamp)

9.15 Creatine Deficiency Mouse Models (AGAT, GAMT)
Arend Heerschap

9.40 A New Knock-in Rat Model of Creatine Transporter Deficiency
Olivier Braissant

10.15 Translational Phenotypes and Biomarkers of Brain Function for Creatine Transporter Deficiency
Laura Baroncelli

10.40 Coffee break

11.00 The Sum of All Parts? Effects of Neurotransmitter-Specific Crt Knockouts on Learning.
Matthew Skelton

11.25 CTD Mouse Model Experiences and Therapeutic Options.
Ton de Grauw

11.50 Lunch break

Session 3: Supplementation Treatment in Cerebral Creatine Deficiency and NBS (Chair: Monique Williams)

13.00 Treatment Outcomes of Cerebral Creatine Deficiency Syndromes.
Saadet Andrews

Epilepsy in Cerebral Creatine Deficiency Syndromes
Saadet Andrews

13.40 ***Q&As***

Session 4: Development of New Treatment in Creatine Transporter Deficiency (Chair:Olivier Braissant)

14.00 Drug Development for Neurodevelopmental Disorders: Lessons Learned from Fragile X Syndrome
Vincent des Portes / Aurore Curie

Clinical Trials for Rare Disease: Current Pitfalls and New Perspectives.
Vincent des Portes / Aurore Curie

14.40 Novel Molecules for the Therapy of Creatine Transporter Deficiency
Maurizio Balestrino

15.05 Dodecyl Creatine Ester-Loaded Nanoemulsion as a Promising Therapy for Creatine Transporter Deficiency
Aloise Mabondzo

15.30 Rescue by 4-Phenylbutyrate of Several Misfolded Creatine Transporter-1 Variants Linked to Creatine Transporter Deficiency Syndrome
Sonja Sucic

15.55 Panel Discussion / Q&As

16.30 Families Debriefing

APPENDIX 1. Validation rapport Tier 1 screening for GAMT-D using Perkin Elmer Qsight 220

APPENDIX 2. Care path for GAMT-D 2020 -Zorgpad Guanidinoacetaat methyltransferase (GAMT-)deficiëntie 2020