

Evaluation of the genetic screening processor (GSP™) for newborn screening

Cite this: *Anal. Methods*, 2013, **5**, 4769

Ralph Fingerhut* and Toni Torresani

The new genetic screening processor for newborn screening (GSP™) from PerkinElmer was extensively tested under routine conditions. The GSP is intended to fully process all newborn screening tests, apart from tandem mass spectrometry. For our evaluation we used all so far available tests for the GSP (TSH, 17-OHP, IRT, total-T4 and GALT). For all 5 tests we have determined specificity, limit of detection (LOD), limit of quantitation (LOQ), intra- and inter-assay variation, recovery, influence of EDTA and on-board stability of the reagents. Results were also compared with AutoDelfia and the Astoria Pacific Spot Check System (GALT). LOD and LOQ were 0.38 and 0.45 mU L⁻¹ (blood) for TSH, 1.30 and 3.25 nmol L⁻¹ (serum) for T4, 0.35 and 0.55 nmol L⁻¹ (blood) for 17-OHP, 0.85 and 1.58 ng mL⁻¹ (blood) for IRT and 2.6 and 3.6 U dL⁻¹ (blood) for GALT. Mean recovery was 97.8–107.1%; intra-assay CVs 2.5–8.9; inter-assay CVs 5.7–11.0. On board stability was >37 days for the immunoassay. The dissolved GALT reagent is not stable at 4 °C, however if the reagent is removed from the GSP and stored at –18 °C, it is stable for 11 days. On board stability of the inducer is >140 days. The GSP is suitable for routine newborn screening. Compared to AutoDelfia, additional control procedures were included which increase the reliability of the system.

Received 9th April 2013

Accepted 11th July 2013

DOI: 10.1039/c3ay40593a

www.rsc.org/methods

Introduction

Newborn screening (NBS) for inborn errors of metabolism and endocrinopathies as performed worldwide is based on whole blood samples dried on filter paper, the so-called dried blood spots (DBS). The general assay format is the 96 well micro-titerplate (MTP), either coated with antibodies for immunoassays for the determination of thyroid stimulating hormone (TSH), total thyroxine (T4), 17-hydroxyprogesterone (17-OHP), and immunoreactive trypsin (IRT). Additional plain plates are used for wet chemistry assays such as the determination of amino acids and acylcarnitines by electrospray ionization tandem mass spectrometry (ESI-MS/MS), determination of total galactose, galactose-1-phosphate uridylyltransferase activity (GALT) and biotinidase activity. Of the immunoassays, fluorescence immunoassays (FIA) are the most sensitive apart from radio immunoassays (RIA), and DELFIA™ technology with time-resolved immunofluorescence combines sensitivity with extremely low interference with autofluorescence from biological samples.¹ AutoDELFIATM also includes total automation of pipetting, incubation, and measurement but there are also some drawbacks. First, DELFIA and AutoDELFIATM kits are sensitive to high concentrations of complexones such as EDTA and citrate that may cause false positive (FP) or false negative

(FN) results.^{2,3} High concentrations of metal ions may also interfere with the assays; however to our knowledge, such interference has not been reported so far in DBS. Second, AutoDELFIATM is a batch processor whereby a second batch can only be loaded after the first batch is finished, and third it is not capable of performing tests other than immunoassays.

The new integrated screening plate processor GSP™ from PerkinElmer overcomes most of these drawbacks. The GSP can be continuously loaded with up to 24 MTP for various assays (to date: TSH, T4, 17-OHP, IRT, and GALT). It has several integrated safety checks to avoid FP or FN results such as a “sample-elution check” in all assays and “check for floating disks” in the wet chemistry assays (*e.g.* GALT). A major and crucial improvement was achieved with the immunoassays. In the Delfia assays lanthanides (for NBS mainly europium) are bound to the respective antibodies or substrates with diethylenetriaminepentaacetate (DTTA) as Ln-DTTA chelates. After the final washing step these chelates are dissociated at pH 3.2 with the enhancement solution which also contains 2-naphthoyltrifluoroacetone (2-NTA), a β -diketone which will form a highly fluorescent Ln-2-NTA complex. However high concentrations of EDTA, citrate, Cu²⁺, Ca²⁺, and Zn²⁺ interfere with the Ln-DTTA chelates, releasing the Ln-label already in the primary reaction step (before the washer-step). The GSP-kits use different chelating agents. Diethylenetriaminepentaacetic acid is used to bind the Ln-label to the antibody as Ln-DTPA chelates. These complexes are more stable than the Ln-DTTA chelates and are not influenced by EDTA, citrate, Cu²⁺, Ca²⁺, or Zn²⁺. However

Swiss Newborn Screening Laboratory, Children's Research Center (CRC), University Children's Hospital of Zurich, Steinwiesstr. 75, CH-8032 Zürich, Switzerland.
E-mail: ralph.fingerhut@kispi.uzh.ch; Fax: +41 44 266 8110; Tel: +41 44 266 7732



these complexes dissociate only slowly at pH 3.2 which is impractical for rapid high throughput analysis. Dissociation of the Ln-DTPA chelates is quick at pH 2.3, but 2-NTA does not form stable complexes with the Ln at this pH. Therefore a different fluorescence enhancer (now called “inducer”), 1-(2-benzofuryl)-4,4,5,5,5-pentafluoro-1,3-pentadione (BFPP), is used.⁴ We had the opportunity to test all available kits on the GSP and compare them either with the AutoDELFIA-kits or with the Astoria Pacific continuous flow analyzer (GALT). In addition all test specifications that could be investigated under routine conditions were verified.

Besides the chemistry, the GSP has some more novelties concerning the handling of test-kits and reagents that might be unnecessary and perhaps even obstructive in the routine newborn screening process. Apart from the expiry dates of each

kit component every kit-lot itself has an expiry date, mainly that of the least durable kit component. Apart from that the expiry date changes dramatically once the reagents are loaded into the refrigerated reagent cassette of the GSP. Initially, this so-called on-board stability was 5 days, but was subsequently extended to 14 days for the immunoassays. Further, although every kit-component has certain specifications such as concentrations for the calibrators and controls *etc.*, it is not possible to interchange components from different kit-lots. These restrictions led us also to investigate their validity.

Materials and methods

The AutoDELFIA Model 1235 automatic immunoassay system, GSP model 2021, GSP Neonatal 17 α -OH-progesterone kit, GSP

Table 1 Limit of detection and limit of quantitation of the GSP-assays

Analyte	Values from kit-inserts				Values determined			
	LOB ^a	LOD ^b	LOQ ^c	Linearity	LOD ^d	LOQ ^e	LOQ ^f	LOQ ^g
TSH [mU L ⁻¹]	0.96	1.31	1.31	0.7–252	0.38	0.45	0.54	<0.09
T4 [nmol L ⁻¹]	2.97	6.44	10.47	1.89–204	1.30	3.25	5.13	<13.55
17-OHP [nmol L ⁻¹]	0.58	1.2	1.2	1.2–314	0.35	0.55	0.81	<0.58
IRT [ng mL ⁻¹]	0.76	2.2	2.2	9–500	0.85	1.58	2.56	<3.0
GALT [U dL ⁻¹]	1.6	2.5	2.5	0.7–16.5	2.6	3.6	6.0	<3.1

^a LOB = limit of blank; defined as the 95th centile of a distribution of blank samples. ^b LOD = limit of detection; according to NCCLS document EP17-A. ^c LOQ = limit of quantitation; defined as the lowest concentration with a total CV equal to or less than 20%. ^d LOD = limit of detection; defined as the mean of a sample without analyte + 3 SD. ^e LOQ = limit of quantitation; defined as the mean of a sample without analyte + 6 SD. ^f LOQ = limit of quantitation; defined as the mean of a sample without analyte + 10 SD. ^g LOQ = limit of quantitation; defined as the lowest concentration with a total CV equal to or less than 20%.

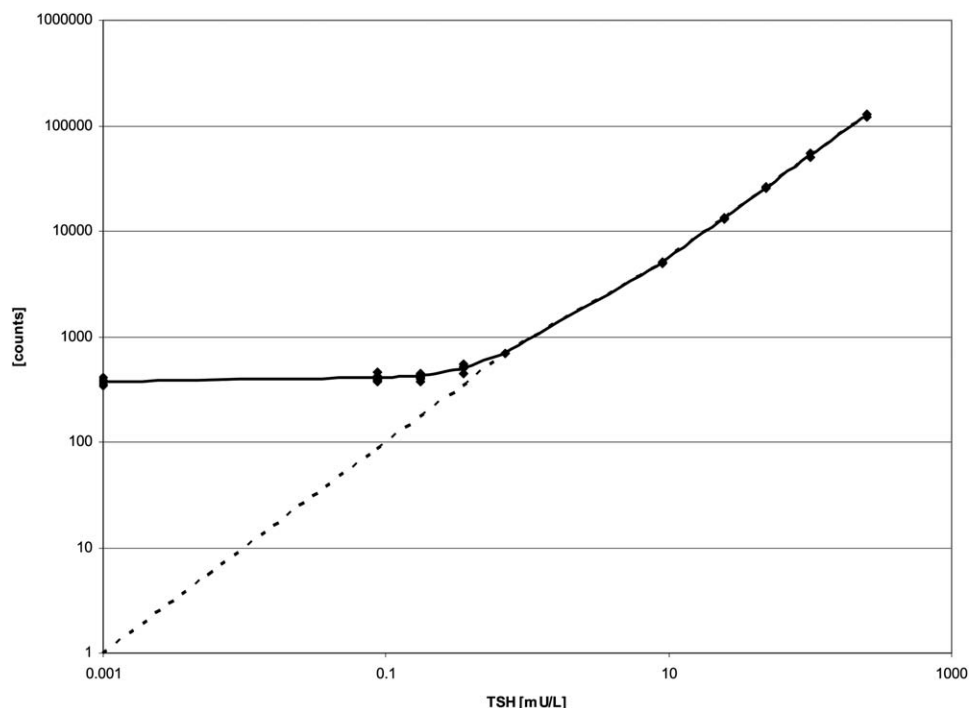


Fig. 1 Calibration curve of TSH with kit-calibrators A–F (0.7–252 mU L⁻¹), A/2 (0.35 mU L⁻¹), A/4 (0.175 mU L⁻¹), A/8 (0.0875 mU L⁻¹), and blank (0 mU L⁻¹). A/2 = kit-calibrator A cut into halves, A/4 = kit-calibrator A cut into quarters, A/8 = kit-calibrator A cut into eighths.

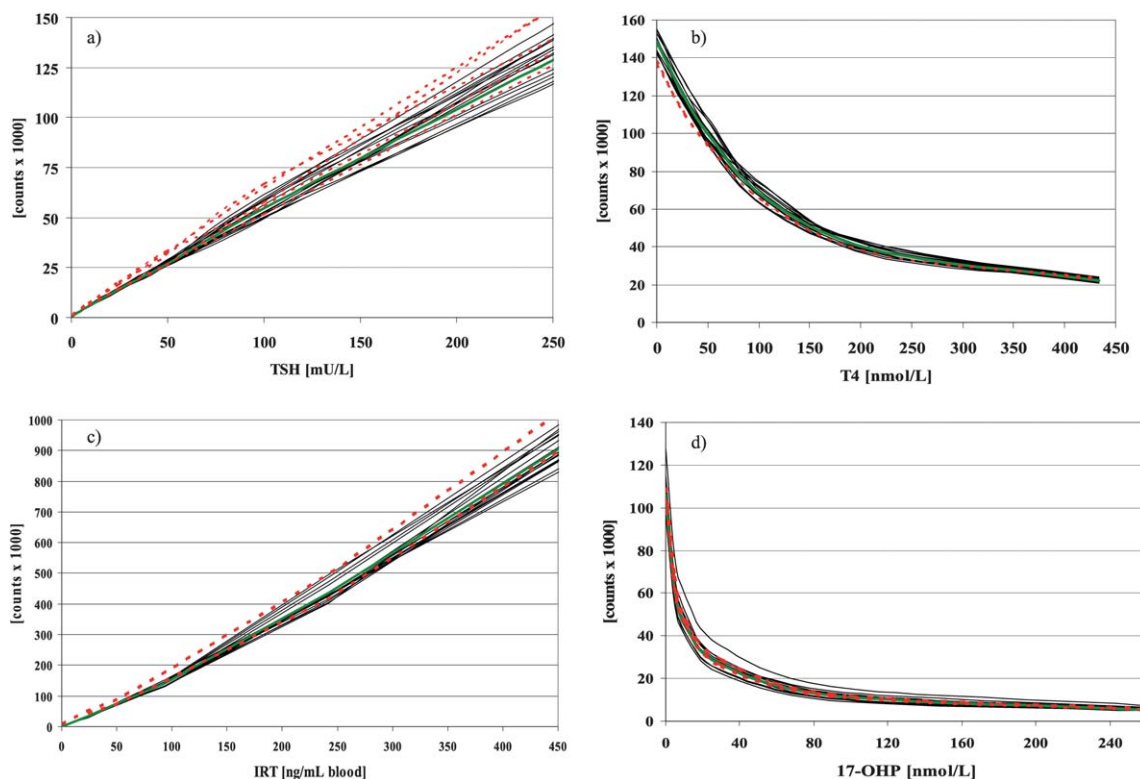


Table 2 Intra-assay variation, inter-assay variation and mean recovery of GSP-assays

Analyte	Intra-assay variation (CV)	Inter-assay variation (CV)	Mean recovery [%]
TSH	8.9 (at 15.0 mU L ⁻¹ ; <i>n</i> = 12) 4.9 (at 59.6 mU L ⁻¹ ; <i>n</i> = 12)	8.1 (at 15.0 mU L ⁻¹ ; <i>n</i> = 30) 8.3 (at 59.6 mU L ⁻¹ ; <i>n</i> = 30)	98.5
T4	8.2 (at 44.5 nmol L ⁻¹ ; <i>n</i> = 12) 4.8 (at 97.0 nmol L ⁻¹ ; <i>n</i> = 12) 3.4 (at 169.8 nmol L ⁻¹ ; <i>n</i> = 12)	11.0 (at 44.5 nmol L ⁻¹ ; <i>n</i> = 30) 8.9 (at 97.0 nmol L ⁻¹ ; <i>n</i> = 30) 8.9 (at 169.8 nmol L ⁻¹ ; <i>n</i> = 30)	101.7
17-OHP	3.1 (at 18.6 nmol L ⁻¹ ; <i>n</i> = 8) 4.9 (at 56.4 nmol L ⁻¹ ; <i>n</i> = 8) 5.0 (at 117.2 nmol L ⁻¹ ; <i>n</i> = 8)	8.7 (at 18.6 nmol L ⁻¹ ; <i>n</i> = 30) 7.5 (at 56.4 nmol L ⁻¹ ; <i>n</i> = 30) 5.9 (at 117.2 nmol L ⁻¹ ; <i>n</i> = 30)	107.1
IRT	3.5 (at 31.6 ng mL ⁻¹ ; <i>n</i> = 12) 5.0 (at 67.6 ng mL ⁻¹ ; <i>n</i> = 12) 2.7 (at 100.0 ng mL ⁻¹ ; <i>n</i> = 12)	6.1 (at 31.6 ng mL ⁻¹ ; <i>n</i> = 30) 5.8 (at 67.6 ng mL ⁻¹ ; <i>n</i> = 30) 5.7 (at 100.0 ng mL ⁻¹ ; <i>n</i> = 30)	97.8
GALT	5.1 (at 3.8 U dL ⁻¹ ; <i>n</i> = 12) 2.5 (at 15.5 U dL ⁻¹ ; <i>n</i> = 12)	8.1 (at 3.8 U dL ⁻¹ ; <i>n</i> = 30) 6.3 (at 15.5 U dL ⁻¹ ; <i>n</i> = 30)	101.9

Table 3 Mean counts and CV of kit calibrators over time

		Mean counts and (CV)					
		Calibrator A	Calibrator B	Calibrator C	Calibrator D	Calibrator E	Calibrator F
TSH	<i>n</i> = 16 over 96 days	718 (8.0)	5722 (6.1)	14 331 (6.5)	26 654 (5.8)	54 777 (8.5)	129 719 (7.8)
T4	<i>n</i> = 16 over 109 days	148 886 (3.0)	119 394 (4.0)	97 927 (5.0)	66 241 (7.4)	38 065 (6.9)	22 324 (6.7)
17-OHP	<i>n</i> = 16 over 91 days	106 204 (8.9)	63 856 (11.4)	44 339 (12.4)	26 584 (14.1)	11 935 (15.1)	5193 (16.5)
IRT	<i>n</i> = 16 over 37 days	1068 (32.7)	33 641 (6.2)	69 873 (5.0)	140 486 (6.0)	437 591 (7.2)	928 934 (5.9)

**Fig. 2** Calibration curves of 16 measurements from the same reagent vials, stored over 96 days (TSH and 17-OHP), 109 days (T4), and 37 days (IRT) inside the refrigerated cassette of the GSP (—), mean of these 16 calibration curves (—), and calibration curves, where GSP reagents were used with microtiter strips from AD kits (---).

Neonatal hTSH kit, GSP Neonatal T4 kit, GSP Neonatal GALT kit, GSP Neonatal IRT kit, AutoDelfia Neonatal 17 α -OH-progesterone test kit, AutoDelfia Neonatal hTSH test kit, AutoDelfia Neonatal T4 test kit, and AutoDelfia Neonatal IRT test kit were all from PerkinElmer, Turku, Finland. Spot Check analyzer model 321, and Spot Check Test kit total galactose (for total galactose and GALT activity) were from Astoria Pacific, Clackamas, USA.

All concentrations are related to whole blood except for total T4 which is related to serum. Units are as follows: TSH mU L⁻¹

Table 4 Expected concentration, mean concentration and CV of kit controls for *n*-measurements over time

Analyte	<i>N</i>	Time period	Expected	Mean	CV
TSH [mU L ⁻¹]	163	16 months	15.7	15.1	9.4
			64.4	60.5	8.1
T4 [nmol L ⁻¹]	65	11 months	43.9	39.7	12.2
			98.0	93.4	8.1
			166.4	161.0	7.7
17-OHP [nmol L ⁻¹]	82	7 months	18.6	20.9	10.3
			56.4	64.5	7.9
			117.2	135.9	8.6
IRT [ng mL ⁻¹]	85	3 months	30.8	27.5	6.7
			64.2	63.9	6.1
			96.4	96.4	4.8

blood; 17-OHP nmol L⁻¹ blood; IRT ng mL⁻¹ blood; GALT U dL⁻¹ blood; T4 nmol L⁻¹ serum.

The GSP Neonatal test kits contain all necessary components ready to use except for the GALT substrate reagent which has to be reconstituted with 2.8 mL of GALT-assay buffer. The only materials not provided with the kits are high- and low-volume pipette tips, wash concentrate and inducer.

For the determination of the limit of detection (LOD) the statistical approach was used. A sample without the analyte was measured repeatedly (*n* = 12) and the mean value + 3 SD was defined as the LOD. The determination of the limit of quantitation (LOQ) is not so straightforward. Statistical approaches using the mean value + 6 SD or mean value + 10 SD and an empirical approach whereby a sample of known analyte concentration is consecutively diluted and assayed until the SD exceeds 20% are described. That concentration for which an accuracy of less than 20% is still achieved is the empirical LOQ. The statistical approach often overestimates LOQs⁵ but even for different analytes measured with the same method, statistical and empirical LOQs may be nearly identical or differ by a factor of up to 10.⁶ We provide data on both statistical and empirical methods. For the statistical method either the lowest calibrator, if the respective analyte concentration was 0, was used (*i.e.* for 17-OHP, T4, IRT, GALT) or a blank filter paper was used (*i.e.* for TSH). Since dilution of DBS is difficult and the preparation of the whole series of DBS from a dilution series is extremely labor

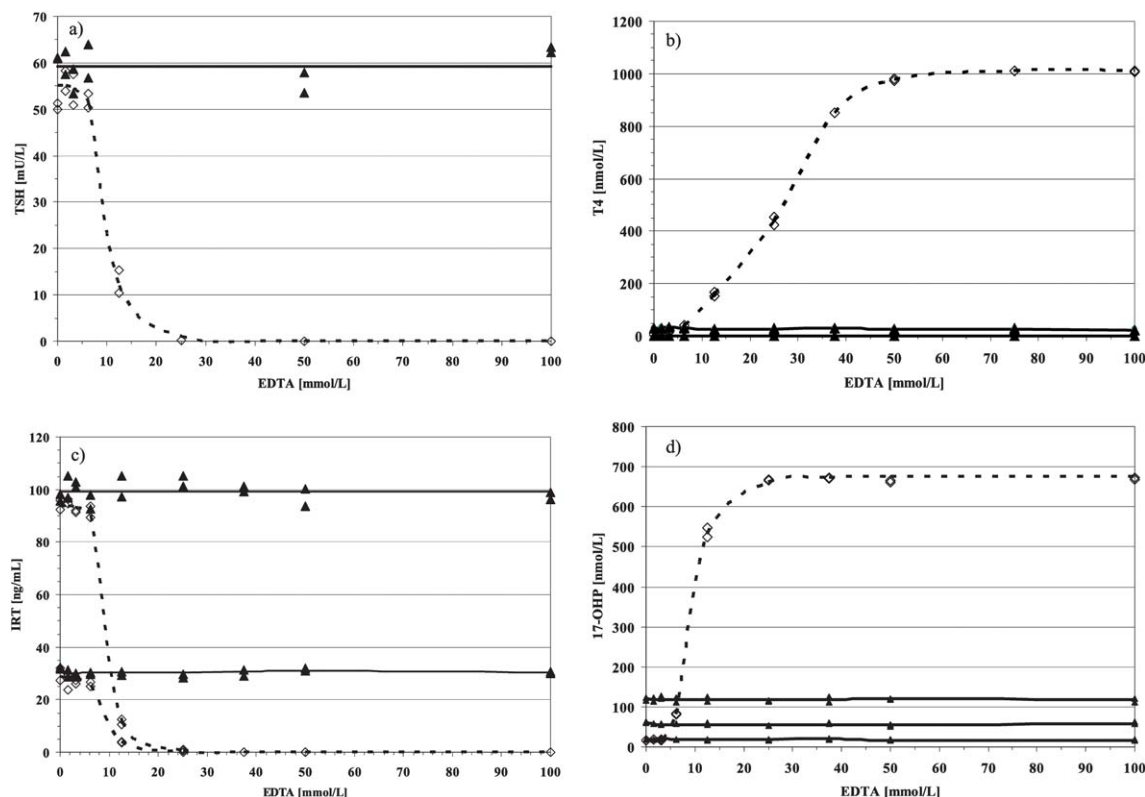


Fig. 3 Influence of EDTA contamination on TSH (a), T4 (b), IRT (c), 17-OHP (d), and GALT activity (e). For all tests, punches from the same control blood sample were used and spiked with EDTA to the indicated final concentration. For the immuno-assays (a–d) the influence of EDTA on the GSP-assays (—) was compared to the AD-assays (---).



Table 5 Comparison of GSP-immunoassays with the AD-immunoassays by linear regression ($cGSP = a \times cAD + b$). "cGSP" and "cAD" represent the concentrations of the respective analyte measured with the GSP or the AD

Analyte	<i>a</i>	<i>b</i>	<i>r</i> ²	<i>n</i>
17-OHP	0.978	0.298	0.975	9766
IRT	0.796	3.203	0.907	517
T4	0.809	15.302	0.816	701
TSH	0.787	0.374	0.961	5963

intensive, we used a rather unconventional approach. The respective kit-calibrators were taken with the lowest concentration different from 0. This calibrator was measured repeatedly ($n = 12$ to 48). If the SD was $<20\%$ then 12 DBS of this calibrator were cut into halves and these 24 halves were measured. If the SD was still $<20\%$, 12 DBS were then cut into quarters, or even into eighths. Since the cutting of 3 mm DBS into equal halves, quarters, or eighths is extremely difficult we have calculated the SD, if necessary, from the sum of 2 quarters of the same half of one DBS or 4 eighths of the same half of one DBS. For statistical analysis SPSS 16.0 was used. The *t*-test for independent samples was used to calculate the significance of the signal differences between blanks and analytes. We also used halves, quarters, and eighths of calibrator A of the TSH-kit,

and plain filter paper to evaluate the possibility of measuring decreased TSH concentrations. For the determination of recovery the respective kit calibrators were used. The assigned values from the quality control certificates were used as target values.

To measure the stability of the inducer at RT we diluted 10 μ L of the TSH tracer with 2.6 mL of the inducer stored at RT. Then 200 μ L of the diluted tracer was transferred into 12 wells of a microtiter plate and measurements made after shaking for 5 minutes using an orbital shaker. In addition the CVs of the kit-controls of the T4-assay were used to estimate the stability of the inducer at RT.

Results

The on-board and overall stability of the inducer is much better than that stated in the kit-inserts. Reloading of the inducer that had been in the GSP for longer than 7 days and use of the inducer longer than 3 months past the expiry date had no influence on the total counts of the calibrators of all 4 immunoassays or on the measured concentrations of the controls (data not shown). The CV of repeated measurements of the diluted TSH-tracer with the inducer stored at RT over a period of 154 days ($n = 90$) was 5.2. In addition the CVs of the T4-kit controls were calculated from fifteen T4-assays that were

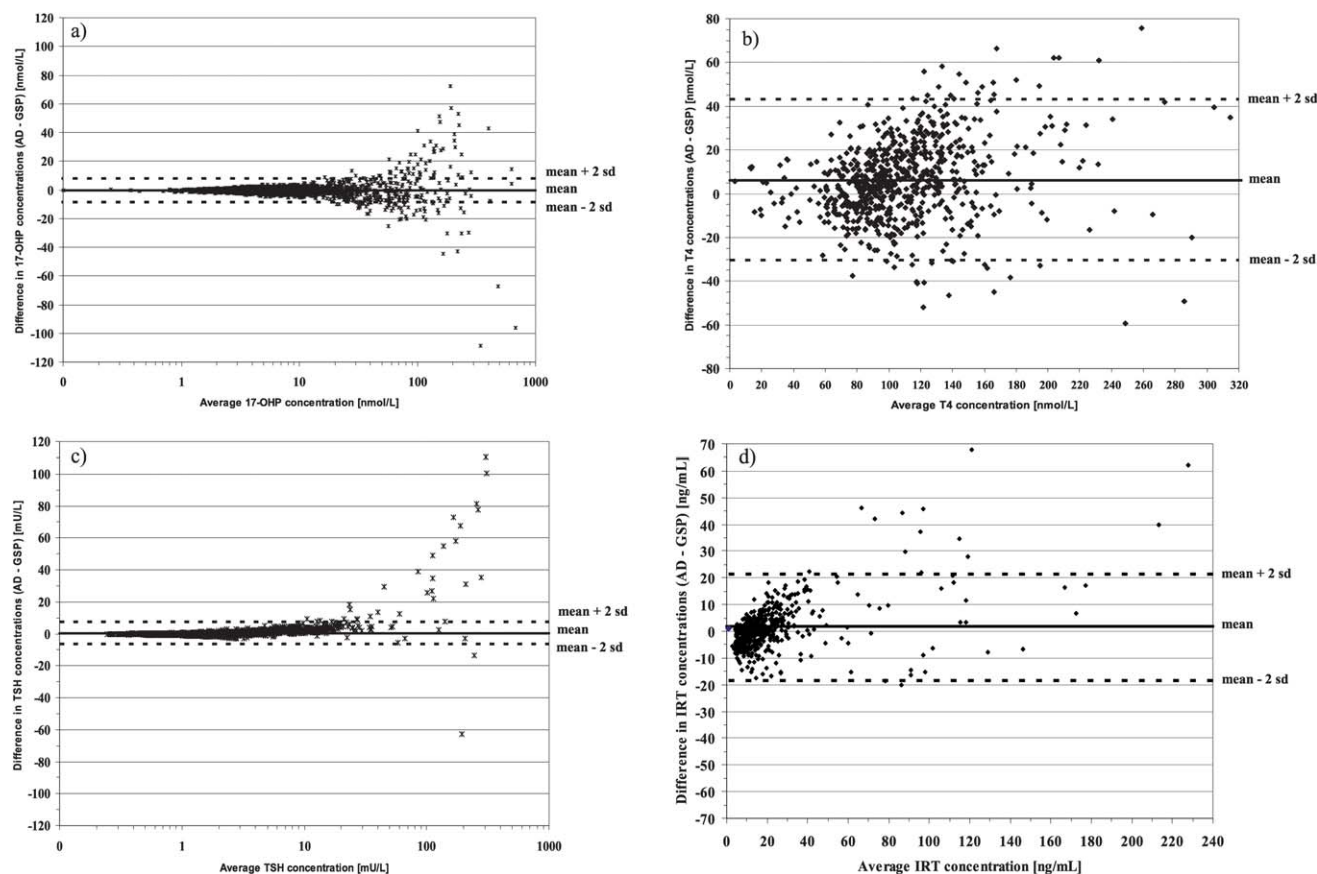
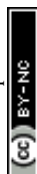


Fig. 4 Comparison of different analyte determinations by the GSP with AD. The difference in analyte concentration is plotted against the average analyte concentration determined by the two methods (17-OHP, (a); T4 (b); TSH, (c); IRT, (d)).



analysed with the inducer that had been stored for 140 days “on board” at RT. CVs were 11.0%, 7.0%, and 8.3%, at 42.9, 91.0, and 159.9 nmol L⁻¹, respectively. Values for limit of detection and limit of quantitation for all GSP-test kits are summarized in Table 1. For the TSH assay our data clearly show that quantitation of TSH below 0.7 mU L⁻¹ is not possible from DBS (Fig. 1), at least not with a linear extrapolation between the lowest calibrator (cal. A) and zero. In Table 2 intraassay variation, interassay variation and mean recovery are listed. To check the onboard stability of the kit-reagents we calculated the mean counts and CVs of the kit calibrators processed and measured with reagents from one vial stored inside the refrigerated reagent cassette of the GSP over 37 to 109 days (Table 3). In addition we plotted all calibration curves of the same time period over each other (Fig. 2a–d). As a second measure for on-board stability we recalculated the concentrations of the kit-controls with the first calibration curve measured with the respective kit-lot. Thus the counts of the kit-controls of measurement n were used with the counts of the calibrators of measurement ($n = 1$) to recalculate the concentrations of the kit-controls. Mean concentrations and CVs are shown in Table 4.

In contrast to the AD-assays, EDTA (up to 100 mmol L⁻¹) had no influence on the determination of TSH, T4, 17-OHP, and IRT (Fig. 3a–d). However the determination of GALT activity is still

influenced by EDTA (Fig. 3e), as previously described for a manual method.³

The performance of the GSP under routine conditions has also been compared to the AutoDelfia system for the immunoassays and to the SpotCheck system from AstoriaPacific for the GALT-assay. Results of linear regression for the different assays are listed in Table 5:

For TSH we were also able to compare two GSP kit lots: $\text{TSH}(503104) = 0.978 \times \text{TSH}(594747) - 0.116$, $r^2 = 0.980$, $n = 4159$. A good agreement was found for 17-OHP between GSP and AD over the whole concentration range (Fig. 4a). The assays for total T4 showed a higher variation although values were evenly distributed around the mean concentrations without suspicion of a systematic aberration (Fig. 4b). For IRT and TSH there was less agreement between GSP and AD. Although the mean difference of the concentrations was close to zero there are significant differences between these two methods. The GSP gives higher values in the low range, up to 2 mU L⁻¹ for TSH and 20 ng mL⁻¹ for IRT, respectively. Above these values the GSP measures lower values (Fig. 4c and d). For the GALT assay there was no good correlation between the GSP and the SpotCheck assay either with a linear regression [$\text{GALT}(\text{GSP}) = 0.022 \times \text{GALT}(\text{SC}) + 8.992$; $r^2 = 0.166$], or with a logarithmic curve fit [$\text{GALT}(\text{GSP}) = -7.67 + 4.03 \times \log(x)$; $r^2 = 0.192$] (Fig. 5).

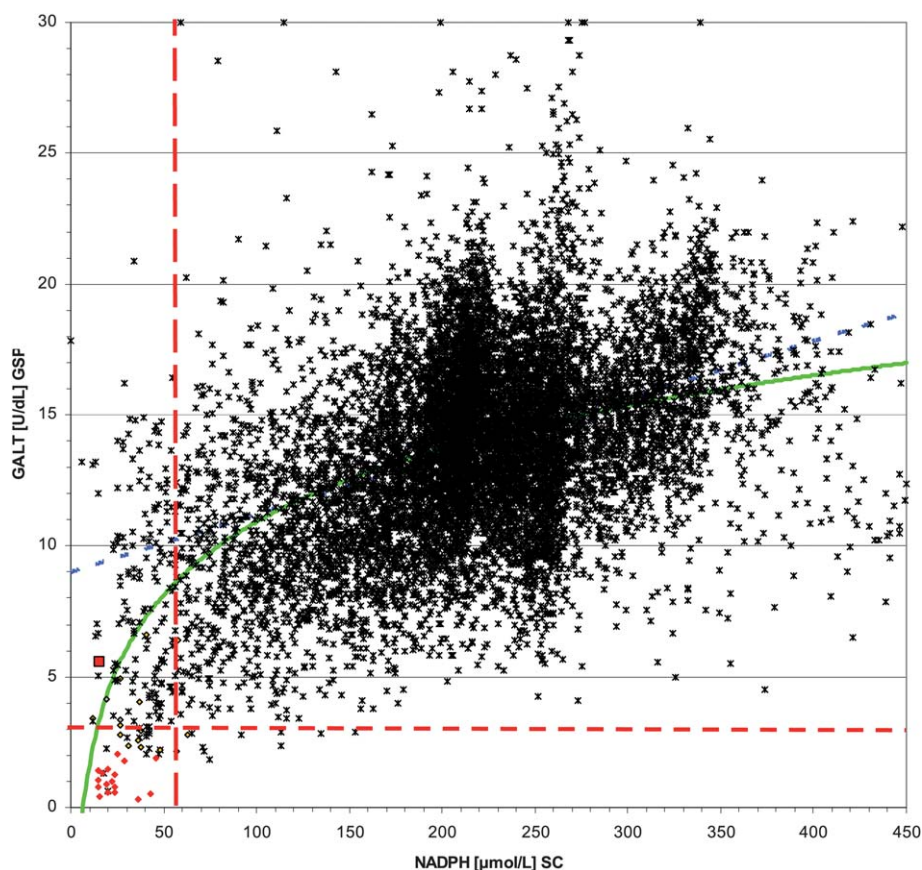


Fig. 5 Comparison of GALT determination by GSP (expressed as U dL⁻¹) and SpotCheck (expressed as NADPH formation in μmol L⁻¹). Cut-off for GSP, horizontal line (---); cut-off for SC vertical line (---); linear correlation (---); logarithmic correlation (---). Normal newborn sample (x), samples from babies with D2/G variants (♦), samples from patients with classical galactosaemia (♦), “floating disk”-sample from a patient with classical galactosaemia (■).



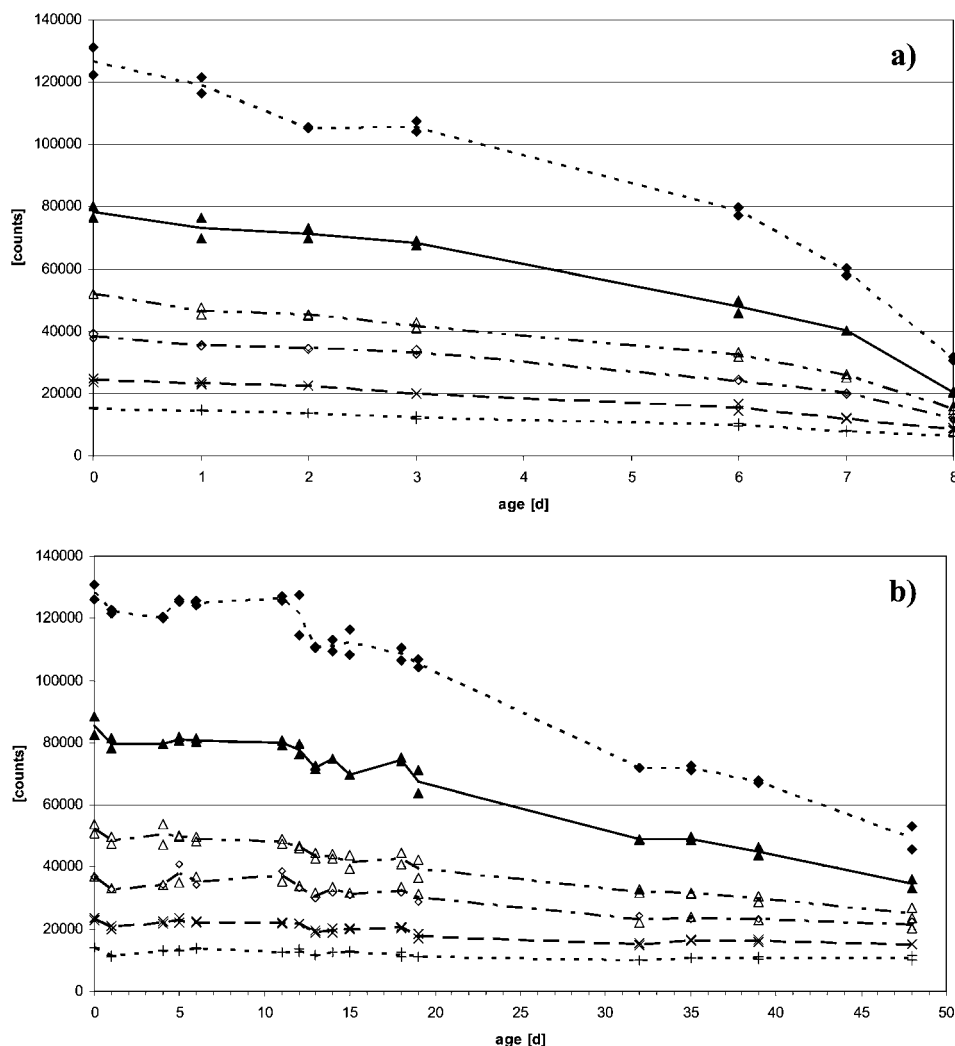


Fig. 6 Stability of GALT reagent after reconstitution, stored inside the GSP in the refrigerated cassette (a), and stored at -20°C between the runs (b). Counts of calibrators a–f are plotted against the time of storage after reconstitution.

Reconstituted GALT reagent is not stable in the refrigerated cassette (“on board”). There is already a decline of 5% per day during the first 2 days of shelf-life (Fig. 6a). However when the reconstituted GALT reagent was removed from the GSP after each run, stored at -20°C and thawed just before the next run, it was stable for 11 days (Fig. 6b).

Discussion

The new integrated plate processor (GSP™) from PerkinElmer combines many improvements for the performance of routine newborn screening. First the different chemistry of the immunoassays which is now insensitive to EDTA contamination of the samples prevents false negative results and reduces the number of false positives for the respective assays. Also the additional measurement for sample elution helps to avoid false negative results. The refrigerated reagent cassette facilitates reagent handling because reagents only have to be loaded once and can normally remain in the instrument until they are used

up. In fact the “on-board”-stability of the reagents for the immunoassays is much longer than that stated by the manufacturer but the usage of the reagents beyond the “on-board”-stability period is not possible, at least not with the default settings of the instrument. The same applies for the inducer. The stated “on board” stability of the GALT reagent is in our experience unacceptable. Since the GSP is obviously intended for use in medium to large size screening laboratories, we can presume that there are at least two instruments installed in these laboratories and they will be most probably run in an alternating manner. Thus GALT reagents will be reused after 48 h and if a batch of 2 or more plates is run after 48 h then the first plate, which will correspond to the calibration curve, will be processed with the GALT reagent stored for 48 h “on board”. If there is insufficient reagent for the second plate then this will be processed with a freshly reconstituted reagent. However this means that the calibration curve has approx. 10–12% lower counts than the newborn samples on the second plate. Therefore according to the manufacturer’s instructions GALT-activity



in nearly 20–50% of all newborn samples is systematically overestimated bearing the risk of false negative screening results for galactosemia. This can be overcome by our proposed modification of freezing the reagents after each run. However, this procedure is not designated in the standard software set-up.

The well known problems with the determination of LOB, LOD, and LOQ^{5,6} remain and could not be totally solved by our investigation. For the IRT and 17-OHP assays the data provided by the manufacturer and our data are quite plausible (see Table 1) and since decreased values of IRT and 17-OHP have no clinical relevance for NBS there is no need for further evaluation. For the other three assays the situation is not so clear-cut (see Table 1). Linearity of the assays, as stated by the manufacturer, exceeds the limit of blank (LOB). However our data are also incongruent for TSH and GALT. For the TSH assay this is explained by the fact that even plain white filter paper, measured with the routine GSP-kit using a linear curve fit, results in a finite TSH-concentration of 0.32 mU L⁻¹ with a CV of 6.9. Therefore the recommendation of the kit manufacturer that “Samples that result in values below 1.31 mU L⁻¹ blood are recommended to be reported as ‘<1.31 mU L⁻¹ blood’.” is supported by our data, although the limit of detection could better be set to the concentration of the lowest calibrator (≈ 0.7 mU L⁻¹). This implies that decreased levels of TSH are not detectable with this method. For total T4 the slightly discrepant results for linearity, LOD and LOQ, do not pose a problem since the lower cut-offs for newborns and even older children (114 nmol L⁻¹ and 65 nmol L⁻¹, respectively) are well within the linear range of the assay. The GALT assay however is a challenge since although the LOD is 2.5–2.6 U dL⁻¹, 18 samples of 4 different patients with confirmed classical galactosemia showed a mean GALT activity of 1.06 U dL⁻¹ (range: 0.50–2.03 U dL⁻¹). It is clear that it is uncertain whether these are true values or just erratic values below the LOD/LOQ. Despite these problems the GALT assay, especially in combination with the measurement of total galactose, is a suitable method for galactosemia newborn screening.

The GSP Neonatal 17-OHP assay and the GSP Neonatal total T4 assay showed very good correlation with the AutoDelfia assays and although every laboratory has to check its cut-offs regularly, it seems that the cut-offs for 17-OHP and total T4 can be readily assigned from the AutoDelfia method. For IRT and TSH, cut-offs have to be checked in advance since there seems to be a difference in calibration. For the TSH assay this is not

quite comprehensible since there is a WHO standard preparation available.

Otherwise the high stability of immunoassays on the GSP would allow assays with kit-lot specific master-curves which would save approximately 5% of the total reagent costs. We were also not able to confirm the manufacturer's constraints that components from different kit-lots of the same assay are not interchangeable. This was clearly demonstrated by the use of AutoDelfia microtiter strips, within the GSP-tests. Again the separate distribution of test-specific components (coated microtiter plates, elution buffer, tracer and antiserum), rather than test-kits with a large amount of surplus elution buffer, tracer, and antiserum would save monetary resources and would additionally protect the environment.

For smaller newborn screening laboratories that still run the manual Delfia assays the installation of the GSP-instrument might not be cost-effective. However they can still benefit from the advantages of the GSP-kit chemistry. The GSP-reagents can be easily used as a manual test-kit and employed with the same Multicalc-protocols on the Victor plate readers from PerkinElmer, as with the Delfia kits. For manually processed TSH, the interassay CVs ($n = 5$) were 5.19 and 5.05 at 15 and 59.6 mU L⁻¹ respectively.

Acknowledgements

We thank PerkinElmer for the provision of reagents for this study and Prof. Brian Fowler for his language advice and helpful suggestions during the preparation of the manuscript.

References

- 1 I. Hemmila, S. Dakubu, V.-M. Mikkala, H. Siitari and T. Lövgren, *Anal. Biochem.*, 1984, **137**, 335–343.
- 2 U. Holtkamp, J. Klein, J. Sander, M. Peter, N. Janzen, U. Steuerwald and O. Blankenstein, *Clin. Chem.*, 2008, **54**, 602–605.
- 3 R. Fingerhut, T. Dame and B. Olgemöller, *Eur. J. Pediatr.*, 2009, **168**, 553–558.
- 4 K. R. Blomberg, V.-M. Mikkala, H. H. O. Hakala, P. H. Mäkinen, M. U. Suonpää and I. A. Hemmilä, *Anal. Bioanal. Chem.*, 2011, **399**, 1677–1682.
- 5 D. A. Armbruster, M. D. Tillmann and L. M. Hubbs, *Clin. Chem.*, 1994, **40**, 1233–1238.
- 6 G. L. Long and J. D. Winefordner, *Anal. Chem.*, 1983, **55**, 712A–724A.

