BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Quantitative G6PD point-of-care test can be used reliably on cord blood to identify male and female newborns at increased risk of neonatal hyperbilirubinaemia: a mixed method study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-066529
Article Type:	Original research
Date Submitted by the Author:	12-Jul-2022
Complete List of Authors:	Bancone, Germana; Shoklo Malaria Research Unit; University of Oxford Centre for Tropical Medicine and Global Health Gilder, Mary Ellen; Department of Family Medicine, Faculty of Medicine, Chiang Mai University Win, Elsie; Shoklo Malaria Research Unit Gornsawun, Gornpan; Shoklo Malaria Research Unit Penpitchaporn, Penporn; Shoklo Malaria Research Unit Moo, Phaw Khu; Shoklo Malaria Research Unit Archasuksan, Laypaw; Shoklo Malaria Research Unit Wai, Nan San; Shoklo Malaria Research Unit Win, Sylverine; Shoklo Malaria Research Unit Aung, Ko Ko; Shoklo Malaria Research Unit Hashmi, Ahmar; Institute for Implementation Science, University of Texas Health Sciences Center (UTHealth), Houston, TX USA; Department of Health Promotion and Behavioral Sciences, School of Public Health, University of Texas Health Sciences Center (UTHealth) Hanboonkunupakarn, Borimas; Mahidol-Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University Nosten, Francois; Shoklo Malaria Research Unit; University of Oxford Centre for Tropical Medicine and Global Health Carrara, Verena Ilona; Shoklo Malaria Research Unit; Institute of Global Health, Faculty of Medicine, University of Geneva, Switzerland McGready, Rose; Shoklo Malaria Research Unit; University of Oxford Centre for Tropical Medicine and Global Health
Keywords:	PAEDIATRICS, NEONATOLOGY, MOLECULAR BIOLOGY

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

For peer review only

TITLE

Quantitative G6PD point-of-care test can be used reliably on cord blood to identify male and female newborns at increased risk of neonatal hyperbilirubinaemia: a mixed method study

Germana Bancone[1,2,*], Mary Ellen Gilder[3], Elsie Win[1], Gornpan Gornsawun[1], Penporn Penpitchaporn[1], Paw Khu Moo[1], Laypaw Archasuksan[1], Nan San Wai[1], Sylverine Win[1], Ko Ko Aung[1], Ahmar Hashmi[5,6], Borimas Hanboonkunupakarn[4], Francois Nosten[1,2], Verena I Carrara[1,2,7], Rose McGready[1,2]

1. Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Mae Sot, Thailand.

2. Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom.

3. Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai 50200, Thailand.

4. Mahidol-Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand.

5. Institute for Implementation Science, University of Texas Health Sciences Center (UTHealth), Houston, TX USA

6. Department of Health Promotion and Behavioral Sciences, School of Public Health, University of Texas Health Sciences Center (UTHealth) Houston, TX USA

7. Institute of Global Health, Faculty of Medicine, University of Geneva, Switzerland

* corresponding author: germana@tropmedres.ac

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

ABSTRACT

Objectives: New point-of-care (POC) quantitative G6PD testing devices developed to provide safe radical cure for P. vivax malaria may be used to diagnose G6PD deficiency in newborns at risk of severe neonatal hyperbilirubinaemia, improving clinical care, and preventing related morbidity and mortality.

Methods: We conducted a mixed-methods study analyzing technical performance and usability of the "STANDARD G6PD" Biosensor when used by trained midwives on cord blood samples at two rural clinics on the Thailand-Myanmar border.

Results: In 307 cord blood samples, the Biosensor had a sensitivity of 1.000 (95%CI 0.859-1.000) and a specificity of 0.993 (95% CI 0.971-0.999) as compared to gold standard spectrophotometry to diagnose G6PD deficient newborns using a receiving operator characteristic (ROC) analysis-derived threshold of ≤4.8IU/gHb. The Biosensor had a sensitivity of 0.727 (95%CI: 0.498-0.893) and specificity of 0.933 (95%CI: 0.876-0.969) for 30-70% activity range in females using ROC analysis-derived range of 4.9 to 9.9IU/gHb. These thresholds allowed identification of all G6PD deficient neonates and 80% of female neonates with intermediate phenotypes.

Need of phototherapy treatment for neonatal hyperbilirubinaemia was higher in neonates with deficient and intermediate phenotypes as diagnosed by either reference spectrophotometry or Biosensor.

Focus group discussions found high levels of learnability, willingness, satisfaction, and suitability for the Biosensor in this setting. The staff valued the capacity of the Biosensor to identify newborns with G6PD deficiency early ("We can know that early, we can counsel the parents about the chances of their children getting jaundice") and at the POC, including in more rural settings ("Because we can know the right result of the G6PD deficiency in a short time. Especially for the clinic which does not have a lab").

Conclusions: The Biosensor is a suitable tool in this resource-constrained setting to identify newborns with abnormal G6PD phenotypes at increased risk of neonatal hyperbilirubinaemia.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Strengths and limitations of this study

- The technical performance of the G6PD quantitative point-of-care diagnostic device was assessed against the current gold-standard spectrophotometric assay.
- Receiving operator characteristic analysis was used to identify the best diagnostic thresholds.
- Usability among clinical personnel from a resource-constrained setting was analysed using a conceptual framework developed for similar settings.
- Fewer than planned focus group discussions were conducted and they occurred in a single clinical site providing a possibly narrower point of view on the usability topics explored.

INTRODUCTION

Pathologically increased levels of bilirubin during the first week of life, i.e. neonatal hyperbilirubinaemia (NH), are common and dangerous for the developing brain. The most severe form of NH, kernicterus, causes neurological sequelae in >80% of neonates (56/100,000 live births globally, [1]). Every year, an estimated twenty-four million newborns are at risk of NH-related adverse outcomes with three-quarters of mortality occurring in sub-Saharan Africa and South Asia [1,2]. These preventable deaths and disabilities disproportionally affect neonates where universal health care and treatment options are scarce, if not absent [3].

Several genetic and clinical factors influence the timing and evolution of NH, including G6PD deficiency, ABO blood group incompatibility, prematurity/low birth weight and sepsis [4]. Early identification of these risk factors can dramatically improve neonatal clinical management during the first days of life [5].

The enzymatic defect of G6PD deficiency, caused by mutations on the X-linked G6PD gene, is a known risk factor for increased levels of bilirubin after birth and it is associated with susceptibility to drug-induced haemolysis [6]. Risk of severe NH is increased in both deficient and heterozygous newborns with abnormal phenotypes [7-9] and universal neonatal screening of G6PD deficiency is supported by WHO in populations where more than 3-5% of males are affected [10].

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1
2
3 78 G6PD deficiency is particularly prevalent among neonates from tropical regions [11], where clinical
4
5 79 care is often provided in a non-tertiary hospital or clinic context. Knowledge of G6PD status by medical
6
7 80 staff and parents can aid in avoiding potentially haemolytic antibiotics or other agents (such as
8
9 81 naphthalene), improved follow-up, and heightened awareness of signs and symptoms of severe NH.
10
11
12 82 G6PD deficiency is very common among the Karen and Burman population along the Thailand-
13
14 83 Myanmar border (9-18% in males, [12]) where it is associated with an increased risk to develop NH
15
16 84 requiring phototherapy both in G6PD deficient (over 4-fold [13]) and in heterozygous females (over 2-
17
18 85 fold [5]) as compared to wild type genotype neonates. In a recent study, screening of G6PD by
19
20 86 qualitative Fluorescent Spot Test (FST) on cord blood failed to identify almost 10% of G6PD deficient
21
22 87 neonates [14].
23
24
25 88 Demonstrating usability of new quantitative Point-Of-Care (POC) G6PD diagnostic tests by locally
26
27 89 trained clinical staff can inform clinical deployment in this setting and in other rural settings. This study
28
29 90 assessed the technical performance and usability of the “G6PD STANDARD” (SD Biosensor, Korea) test
30
31 91 when used by trained midwives in two clinics along the Thailand-Myanmar border.
32
33
34
35 92 MATERIALS AND METHODS
36
37
38 93 Study design
39
40 94 A mixed-methods study was conducted to evaluate both the technical performance of the “G6PD
41
42 95 STANDARD” (SD Biosensor, Korea) test (henceforth “Biosensor”) and its usability by midwives in a non-
43
44 96 tertiary setting. G6PD enzymatic activity and haemoglobin concentration measured by the device were
45
46 97 compared to the gold standard reference spectrophotometric assay and haematology analyser,
47
48 98 respectively. Performance of the G6PD fluorescent spot test (FST) currently used routinely at the point-
49
50 99 of-care, was also compared to the reference and new test.
51
52
53 100 Following local staff training, user proficiency was assessed before study start; usability was explored
54
55 101 using focus group discussions (FGD) at the end of the study.
56
57
58
59
60

For peer review only

Study setting and population

The study was conducted in SMRU clinics situated along the Thailand-Myanmar border in Tak province (Thailand) where free antenatal care and birthing services are provided for migrant women of predominantly Karen and Burman ethnicity.

SMRU midwives come from the same population as the pregnant women and patients seeking care at SMRU clinics. The majority of midwives have primary or secondary education and receive clinical training on-site. Pregnant women attending SMRU clinics at Wang Pha (WPA) and Maw Ker Thai (MKT) were informed about the study at regular antenatal care visits in the 3rd trimester. Informed consent procedures and eligibility assessments for mothers were completed before labour commenced. Eligibility of neonates was assessed immediately after delivery, and those born at an estimated gestational age (EGA) by ultrasound ≥35 weeks with no severe maternal complications at delivery and no severe neonatal illness were included. In order to allow laboratory analyses to be performed within 30 hours from collection, only neonates born during week days were included. For all neonates, indication for starting phototherapy treatment followed the recommendations of the UK NICE guidelines [15].

Blood analyses for technical evaluation of Biosensor

Two milliliters of cord blood were collected into EDTA from the umbilical cord using an established SMRU SOP. An aliquot of anticoagulated blood was used by the midwives in the delivery room for the Biosensor following manufacturer's instructions within one hour of collection (Appendix 1). Tests were repeated if the test result was an error or "HI" (a result obtained when G6PD activity is very high, outside the instrument analytic range). High-level and low-level Biosensor controls were run weekly or monthly (depending on availability) from April 2020 until May 2021.

An aliquot of anticoagulated blood was analysed by G6PD fluorescent spot test (FST) at the clinical laboratory. The remaining blood was stored at 4°C until shipment to the central SMRU laboratory on the same day.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Gold standard reference testing for G6PD and haemoglobin were performed by spectrophotometric assay and haematology analyser (with complete blood and reticulocyte counts), respectively, at the SMRU central laboratory.

G6PD spectrophotometric assay was performed using Pointe Scientific kits (assay kit # G7583-180, Lysis Buffer # G7583-LysSB). Kinetic determination of G6PD activity at 340 nm was performed using a SHIMAZU UV-1800 spectrophotometer with temperature controlled cuvette compartment (30°C). Samples were analysed in double and mean activity was expressed in IU/gHb using the Hb concentration obtained by complete blood count analysis. The final result was calculated using manufacturer's Temperature Control Factor of 1.37. Two controls (Normal, Intermediate or Deficient; Analytic Control Systems, Inc. USA) were analysed at every run and results compared to expected ranges provided by manufacturer. Complete blood count was performed using a CeltacF MEK-8222K haematology analyser (Nihon Kohden, Japan). Three-levels quality controls were run every day and device maintenance and calibration were performed regularly. Reticulocytes were analysed by microscopy after staining with supervital staining Crystal Violet.

Buffy coat recovered from whole blood after centrifugation was stored at -20°C for later DNA extraction using standard columns kit (Favorgen Biotech, Taiwan). Genotyping for G6PD common mutations was performed through established SOPs [16]. Mahidol mutation was analysed in all samples. Other mutations were only analysed in phenotypically deficient or intermediate samples (G6PD < 9.31IU/gHb by reference test) with wild type or heterozygote Mahidol genotypes. Viangchan, Chinese-4, Kaiping, Canton, Union and Mediterranean were analysed first and full gene sequence was performed if none of these mutations were found.

Biosensor training, user proficiency and usability assessment

Midwives of WPA and MKT SMRU clinics were trained for use of Biosensor and were eligible to participate in the usability component of the study following informed consent. Two to four training sessions were provided at each clinic in the local language by an experienced laboratory technician

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

(author LA). The sessions lasted from 1 to 2 hours and included a short introduction about the test, a practical demonstration using imitation blood, and supervised use of the biosensor by each midwife. Midwives were allowed to practice the procedure the week following the training prior to taking a user proficiency test. The proficiency test was administered by author LA in the local language and it consisted of a questionnaire (modified from a questionnaire developed by PATH (https://www.finddx.org/wp-content/uploads/2020/09/PATH_STANDARD-G6PD-User-Competency-Assessment-quiz_08oct19.pdf) and direct observation of two consecutive tests. Midwives were asked to explain out-loud their actions while performing the first test. The proficiency test was analysed by authors GB and GG and midwives who scored <85% were re-trained before study start. A visual aid with all critical steps of the procedure was printed and available in the delivery room during the study.

The usability component of the study followed the conceptual framework for acceptance and use of a rapid diagnostic test for malaria proposed by Asiimwe et al. [17] that evaluates 6 components: learnability, willingness, suitability, satisfaction, efficacy, and effectiveness. The focus group discussions (FGD) specifically focused on 4 main themes of learnability, willingness, satisfaction, and suitability. Due to COVID, only two of the planned six total FGD were conducted. The midwives were grouped by their seniority, with senior and junior midwives together, and midwife assistants in a separate group in order to encourage honest and open conversation. One researcher (KKA) facilitated the FGD while an experienced assistant took notes; both were fluent in Burmese and Karen languages used in the FGD. Immediately following the FGD, research staff debriefed and noted main themes of the discussion. FGDs were audio-recorded and subsequently translated and transcribed in English. Two researchers (MG and GB) independently analysed the transcript using thematic analysis based on the pre-set framework [17] using Taguette (a free and open access qualitative data analysis software, https://joss.theoj.org/papers/10.21105/joss.03522) and confirmed findings with the KKA. Face-to-face meeting and exchange of notes allowed for triangulation between the researchers.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1
2
3
4 177
Blood analysis for assessment of neonatal hyperbilirubinaemia

178 Routine clinical care for newborns included at least one total serum bilirubin (TSB) test before 179 discharge (around 48h of life) using capillary blood measured on-site by the rapid quantitative 180 bilirubinometer BR-501 (Apel Co. Ldt, Japan).

181

182
Sample size and statistical analyses

183 The expected prevalence of G6PD deficiency in the population living at the border is 9-18% in males 184 and 2-4% in female [12, 16] corresponding to approximately 20-30% heterozygous females, 60% of 185 whom have intermediate activity [18]. Assuming that the proportion of females and males in the 186 neonate population is 50%, 9% were expected to be G6PD deficient and 7% to be G6PD intermediate. 187 In order to obtain 95% CI of the limits of agreement within 0.5 SD of the difference, about 31 neonates 188 with deficiency and 25 with intermediate phenotypes were needed, with a minimum total sample size 189 of 350 samples.

190 Clinical data were double entered in MACRO and collated with laboratory data; data were analysed 191 using SPSSv27.

192 Male median (MM) was calculated in all males with wild type genotypes in both the references 193 spectrophotometric assay and the Biosensor. Deficiency was defined as enzymatic activity below 30% 194 of MM by reference spectrophotometry and receiving operator characteristic (ROC)-derived 30% 195 threshold by Biosensor; intermediate phenotypes were defined as enzymatic activity between 30% 196 and 70% of the MM or ROC-derived threshold.

197 Mean and standard deviation (SD) were reported for continuous variables. Categorical variables were 198 compared by Chi-squared test and ANOVA. Bland-Altman plot was used to inspect correspondence 199 between G6PD activity detected by Biosensor compared to the spectrophotometry assay [19]. 200 Correlation was assessed using Pearson's coefficient of correlation and Interclass Correlation 201 Coefficient (ICC). Area under the curve (AUC) of the ROC curve [20] was calculated at different activity 202 thresholds to analyse clinical performances (i.e. sensitivity and specificity) of the Biosensor. Cohen's

5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Kappa coefficient was calculated for categories of phenotypes identified by Biosensor and spectrophotometry.

For analysis of haematologic features and risk of neonatal hyperbilirubinaemia, neonates gestational ages assessed by ultrasound were categorized as ≤38 and >38 weeks according to epidemiologic studies conducted previously in the same population [21].

Statistical significance was assessed at the 5% level.

Patient and Public Involvement statement

At the outset of the study, the research team engaged the local population through a local ethics and research advisory committee, the Tak Province Community Advisory Board, Thailand. This group is comprised of community leaders, and were asked to advise on study design, process, and outcomes of interest, and subsequently approved the study (TCAB201904).

RESULTS

A total of 331 cord blood samples were collected between April 2020 and November 2021; six were clotted and excluded from all analysis. Of the remaining 325 samples, 257 (79%) were collected in MKT clinic and 68 in WPA clinic, in 166 (51%) female and 159 male neonates. Mean (SD) of estimated gestational age of newborns was 39.1 (1.0) weeks.

General haematologic characteristics

As expected for this specimen, haematological characteristics of cord blood (Table 1) show higher white blood cell count, haemoglobin concentrations, reticulocyte counts and larger cellular volumes compared to adult blood. Reticulocyte counts and red cell distribution width were higher in neonates <38 weeks gestational age (P=0.02 and P=0.01 respectively) while the other indexes did not differ by gestational age groups.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1. Haematologic characteristics of cord blood samples according to newborn gestational age. Results are shown as mean (SD)

EGA (weeks)	N*	WBC (10^3/ uL)	NEU (10^3/ uL)	LYM (10^3/ uL)	RBC (10^6/ uL)	HGB (g/dL)	HCT (%)	MCV (fL)	MCH (pg)	MCHC (g/dL)	RDW (%)	PLT (10^3/ uL)	Reticulocyte (%)
<38	19	13.1 (3.6)	9.6 (3.3)	2.7 (1.7)	4.3 (0.4)	14.4 (1.7)	48.0 (5.4)	110.9 (6.6)	33.2 (2.7)	29.9 (1.5)	16.8 (1.5)	259.2 (66.2)	2.8 (1.8)
≥38	298	14.3 (3.8)	10.8 (3.6)	2.8 (1.6)	4.5 (0.5)	14.5 (1.7)	49.0 (5.2)	109.0 (7.9)	32.3 (3.0)	29.6 (1.4)	16.0 (1.2)	261.4 (47.7)	2.1 (1.1)
P_{ANOVA}		*0.17*	*0.16*	*0.88*	*0.14*	*0.68*	*0.43*	*0.30*	*0.21*	*0.41*	*0.01*	*0.85*	*0.02*

* Number of samples analysed by haematology analyser was 317 out of 325; 7 samples were analysed by Hemocue and result used to calculate G6PD enzymatic activity.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

G6PD genotypes

A total of 26 hemizygous mutated males (21 Mahidol, 2 Kaiping, 1 Viangchan, 1 Coimbra, 1 Orissa), 3 homozygous mutated females (Mahidol), 34 heterozygous females (32 Mahidol, 1 Canton, 1 Viangchan) and 262 wild type (129 females and 133 males) were found. Overall allelic frequency of all mutated alleles was 13.4%. The distribution of G6PD activity by spectrophotometry and biosensor associated with different genotypes are shown in Figures 1 and Supplementary Tables 1 and 2.

Fluorescent spot test

The poor performances of the FST in cord blood were confirmed here, with the FST failing to identify 23% (7/30) of deficient neonates and 100% of the intermediate females (22/22; Table 2).

Technical evaluation of Biosensor

Male medians by reference spectrophotometric assay and Biosensor

MM G6PD activity by spectrophotometer was 13.3 IU/gHb giving a 30% threshold of 4.0 IU/gHb for diagnosis of deficiency; intermediate activity (30-70%) in females ranged between 4.1 and 9.3 IU/gHb. The cord blood-specific 30% spectrophotometric threshold identified all the hemizygous male and homozygous female newborns (Figure 1A).

MM of G6PD activity by Biosensor calculated on 307 samples was 14.4 IU/gHb giving a 30% threshold of 4.3 IU/gHb for diagnosis of deficiency. Intermediate activity (30-70%) in females ranged between 4.4 and 10.1 IU/gHb (Figure 1B).

In 7% of cases (23/325), the Biosensor provided an initial result of “HI” activity without a numeric value. Of the 19 samples retested, 14 had “HI” results again and 5 samples had an activity ranging from 17.3 to 20.0 IU/gHb; all samples with initial or confirmed “HI” results were normal by spectrophotometry and had a wild type genotype. Overall, 18 samples (5.5% of the total) did not have a final numeric result by Biosensor but would have been considered “normal”, according to the spectrophotometric assay.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Biosensor performance

Biosensor performance was assessed for 307/325 samples that yielded numeric results. The mean (±1.96SD) difference in enzymatic activity between Biosensor and spectrophotometry was 1.05 IU/gHb (LoA: -3.52 to 5.62 IU/gHb) as represented in the Bland-Altman plot in Figure 2A. A very strong correlation between enzymatic activity by Biosensor and reference spectrophotometry was observed (Pearson's r=0.855, p<0.001; ICC=0.905, p<0.001).

The mean (±1.96SD) difference in Hb between the Biosensor and haematology analyser was 0.70 g/dL (LoA: -2.83 to 4.23 g/dL) (Figure 2B). A moderate correlation between Hb levels by Biosensor and haematology analyser was observed (Pearson's r=0.637, p<0.001; ICC=0.728, p<0.001).

Area under the curve (AUC) of the ROC analysis (Figure 3A) of the 30% threshold was 0.999 (95%CI: 0.997-1.000); ROC analysis showed that 30% of Biosensor MM (4.3IU/gHb) was associated with sensitivity of 0.931 (95%CI: 0.758-0.988) and specificity of 0.989 (95%CI: 0.966-0.997) while a threshold of 4.8IU/gHb had a sensitivity of 1.000 (95%CI: 0.859-1.000) and a specificity of 0.993 (95% CI: 0.971-0.999). This second threshold was therefore used for the subsequent analyses.

AUC of the ROC analysis (Figure 3B) for the 70% threshold was 0.972 (95%CI: 0.949-0.994) and ROC analysis showed that a threshold of 9.9IU/gHb had a better sensitivity and specificity as compared to the 70% of Biosensor MM (10.1 IU/gHb). The ROC-derived threshold had a sensitivity of 0.842 (95%CI: 0.716-0.921) and specificity of 0.984 (95%CI: 0.957-0.995) to identify samples with ≤70% activity and was used for subsequent analyses.

AUC of the ROC analysis for the range 30-70% activity was 0.935 (95%CI: 0.887-0.983); sensitivity and specificity for intermediate phenotypes in females were 0.727 (95%CI 0.498-0.893) and 0.933 (95%CI: 0.876-0.969) respectively based on ROC-derived thresholds as compared to 0.592 (95%CI: 0.390-0.770) and 0.953 (95%CI: 0.897-0.980) using Biosensor MM thresholds.

When comparing phenotypes defined according to the 30% and 70% thresholds of spectrophotometry and ROC-derived threshold for Biosensor (Table 2), the Biosensor correctly identified all deficient and

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

normal males and all deficient females. In females, the Biosensor incorrectly identified 9% (2/22) of intermediate females (activity by spectrophotometry 33% and 62%) as deficient, and 7% (9/130) of phenotypically normal female neonates as intermediate (activity by spectrophotometer ranging from 71% to 113%). It also misdiagnosed 18% (4/22) of intermediate samples as normal. Of these 4 samples, 3 were Mahidol heterozygotes and 1 was a wild type and their enzymatic activity by spectrophotometry ranged from 54% to 64%. Cohen's kappa coefficient was 0.841, p<0.001. Overall, the majority of samples with discordant results (11/15) were identified by the Biosensor as having a "worse" phenotype. Characteristics of the 15 samples with discordant results are reported in Supplementary Table 3.

Table 2. Diagnostic performance of FST and Biosensor as compared to gold standard spectrophotometry.

		Spectrophotometry				
		Male		Female		
		Deficient	Normal	Deficient	Intermediate	Normal
FST	Deficient	20	0	2	0	0
	Normal	6*	133	2*	22	137
	Total	**26**	**133**	**4**	**22**	**137**
Biosensor	Deficient	26	0	4	2#	0
	Intermediate	NA	NA	0	16	9$
	Normal	0	125	0	4&	121
	Total	**26**	**125**	**4**	**22**	**130**

Phenotypes are based on 30% and 70% thresholds for spectrophotometry. For Biosensor, threshold for deficiency is ≤4.8IU/gHb and 4.9 to 9.9IU/gHb for intermediate, both obtained by ROC analysis. Total sample for Biosensor was 307; total sample for FST was 322 (3 samples were not analysed by FST at the clinic)

*Enzymatic activities ranging from 12% to 27% of spectrophotometry MM.
Two Mahidol heterozygotes with activity by spectrophotometry of 33% and 62% of MM.
$ Two Mahidol heterozygotes and 7 wild type samples with enzymatic activity by spectrophotometry ranging from 71% to 113%.
& Three Mahidol heterozygotes and 1 wild type samples with enzymatic activity by spectrophotometry ranging from 54% to 64%.
Characteristics of discordant samples are reported in Supplementary Table 1.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

No difference in results were observed by clinic (ICC=0.899, $p<0.001$ in MKT and ICC=0.930, $p<0.001$ in WPA) or user. In MKT clinic where the test was used over 20 months, a trend of larger absolute mean differences in activity (Biosensor - Spectrophotometry) were observed in the last 4-8 months of use as compared to the first 12 months (Supplementary Figure 1).

Risk of neonatal hyperbilirubinaemia

Risk of neonatal hyperbilirubinemia by phenotype (determined by spectrophotometry) was assessed in term neonates (EGA≥38weeks). A significantly larger proportion of G6PD deficient neonates (29%) underwent phototherapy for treatment of NH as compared to G6PD normal (6%, RR[95%CI] =4.9 [2.3-10.5]; P<0.001). A larger proportion of female neonates with intermediate phenotypes (90% of whom were heterozygotes) required phototherapy (15%), although in this small cohort the difference did not reach statistical significance (RR[95%CI] =2.6 [0.8-8.1]; P=0.13; supplementary Table 4. Relative risk by quantitative phenotypes were similar to those already established by genotypes in the same population [5].

Biosensor training, user proficiency and usability assessment

A total of 22 midwives in two clinics were initially trained and completed the users' proficiency test, including 7 senior, 10 junior and 5 assistant midwives. Median (min-max) observed score from the questionnaire (max 7 points) and observed tests (max 18 points) was 22.1 (18-24.5). The median score did not differ by seniority: assistant 21.4 (18.0-23.5), junior 22.0 (19.3-24.5), senior 22.8 (21.0-24.5); most midwives (72%) had a score >21 points (>85% of maximum score). The most common mistakes in the questionnaire were on how to mix the blood and the buffer (pipetting 10 times *vs* shaking the buffer tube) and on volume of blood mixture to transfer into the device. On observation, the most common mistakes were failure to check the date on Biosensor screen and failure to check test expiry date (rated as minor mistakes as expired test strips are automatically recognized by the Biosensor and rejected).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Two focus group discussions were held in December 2021 in MKT clinic, four weeks after completion of the sample collection at that site; one FGD included 6 senior and junior midwives, and one included 6 assistant midwives. Discussions on satisfaction, learnability, willingness, and suitability and future use are summarized in Table 3. Overall satisfaction was high, although staff were concerned with invalid results, and found it challenging to dedicate one member of the team to perform the biosensor test in the delivery room in the busy postpartum period. In terms of learnability, the midwife assistants reported learning the device more easily, though some were anxious about missing steps. The senior staff were anxious about mistakes and clotted blood, and reported the need to refer to the instructions as a problem. Contrary to the positive expressions to keep using the device at the clinic, the midwives' willingness to use the device was not high and they requested a dedicated staff to perform the test or the test to be done in the laboratory. In terms of suitability and future use, the midwives found the results clinically useful and a valuable diagnostic tool in both their setting and field clinics. However, they were concerned about neglecting clinical care while doing a laboratory test, the cost of the device, and emphasized the need for good training.

Table 3. Selected quotes by theme from focus group discussions.

Theme	Quotes
A. Satisfaction	"It is very good for the children. It is good to know if the child has G6PD deficiency or not from birth. The advantage of the device is that it can detect the children without having to do a heel stick on the baby. On the other hand, there is an increase in work…. But now that we are good at using it, it's fine." [FGD1] "Sometimes if someone is doing the test by using the device it means there are fewer staffs to be with mothers and babies which is not good." [FGD1]
B. Learnability	"After the one-time training, we had 1 or 2 times experiences practically. Then we can do it." [FGD2] "I am really scared I will forget the steps." [FGD2] "We have to look at the book very often, if not we forget the process of what to put and how to put it." [FGD1]
C. Willingness	"Facilitator: Yes. What do you think about keeping on using this device in the future? Participant: Of course. It is good. Participant: Yes, it is good. But if we can have a specific staff to do it then it will be better." [FGD2] "To make changes, take out the blood and send it to the lab. Then only lab staff have to

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	do that." [FGD1]
D. Suitability & Future Use	"Because we can know that early, we can have counseling with the parents about the chances of their children getting yellow skin. We can take time to counsel." [FGD1] "Because we can know the right result of the G6PD deficiency in a short time. Especially for the clinic which doesn't have a lab then it is difficult to know the G6PD status. But with this device, they will only need to take a little blood from the baby and they can know the result of G6PD." [FGD2]

DISCUSSION

This is the first study to assess clinical performance and usability by locally trained health workers of the "STANDARD G6PD" Biosensor test for identification of G6PD deficient and intermediate phenotypes in cord blood. Current data, together with previously collected evidence from clinical trials in the same population [5], clearly indicate that newborn heterozygous girls with G6PD intermediate phenotypes, who are not identified by the FST, are at increased risk of NH and require phototherapy [7, 8]. The availability of a validated POC quantitative test such as the Biosensor and its inclusion in diagnostics guidelines for neonatal care at birth will allow identification of this group of neonates and better clinical care in several settings [22-25]. Together with other easy-to-use non-invasive tools for diagnosis of NH (e.g. Transcutaneous bilirubinometers), this study provides evidence that Biosensor could be used in non-tertiary rural settings for identification of neonates who need referral to higher levels of care. In settings where phototherapy is available, this study indicates that the Biosensor is a better option than FST to support clinical management of neonates. Technical performance of the Biosensor using ROC-derived threshold was comparable to that observed in adult blood in laboratory and field studies [26-29].

The phenotypic classification provided by the Biosensor was superior to the currently available qualitative test (FST) both for deficient and for intermediate phenotypes. Among intermediate phenotypes, 80% were identified as either deficient or intermediate, allowing a better identification of neonates at potential jaundice risk as compared to the currently used FST-based diagnosis [14, 30]. Poor performance of FST can be explained by the higher G6PD enzymatic activity at birth as compared

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

to adulthood [31, 32]; this is probably the result of several haematological factors including younger red cell age, increased number of reticulocytes with higher G6PD activity [33, 34] and higher WBC count [28] as observed here. Importantly, because of higher enzymatic activity in cord blood, thresholds established in adult blood cannot be used to identify deficient or intermediate phenotypes by either spectrophotometry or Biosensor at birth and would have missed identification of 10% (3/29) deficient neonates (2/26 deficient males and 1/4 deficient females) and 86% (19/22) intermediate females.

Biosensor haemoglobin values had a moderate correlation with those assessed by automatic haematology analyser. Although cord (and neonatal) blood samples have higher haemoglobin levels and increased viscosity, Biosensor's performance in measuring G6PD activity was not worse at higher haemoglobin levels.

While the Biosensor provided a numeric result in 94.5% of cases, in few cases an "error" message or a "HI" result was obtained which, according to the protocol, required re-analysis of the sample. Samples that tested "HI" were confirmed to be normal, both phenotypically by spectrophotometry and by genotype (all wild type). In routine practice it will not be needed to repeat the test in samples showing "HI" result should the manufacturers include this information in the instructions for use.

The usability component highlighted important themes to be taken into consideration for future use of the Biosensor at birth. The midwives have been involved in previous research regarding neonatal jaundice and appreciated the importance of early G6PD diagnosis to identify newborns most at risk of neonatal hyperbilirubinaemia and to facilitate optimal clinical care and parental counselling. The non-invasive nature of cord blood analysis was considered an advantage. In this setting, the SMRU midwives recommended that the test be performed by dedicated staff or by the available laboratory to assure appropriate clinical care is provided to the newborns and mothers; nevertheless, they estimated that in more rural contexts it may be appropriate for trained birth attendants to perform the test. Of note, midwives considered their reliance on reading the visual aid while performing the test (which is standard practice in laboratories) a weakness and this aspect might need to be taken

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

into account when training clinic field staff. Usability results obtained here might not be generalizable to every other context but there are data being collected in several rural and community-based settings that corroborate ease of use of this device to guide malaria treatment after appropriate training [26, 35, 36] .

Although midwives felt uncertain about properly conducting the test at the beginning of the study, the laboratory data showed highly accurate results in the first 12 months of use and very good results in the latter 8 months, supporting suitability of the test among health care workers without prior experience in diagnostics. Follow up studies should explore the causes of this slight decrease in quality over time which could be attributed to environmental or users' factors as well as device durability over >1 year of use in tropical conditions.

Limitations

A practical limitation of Biosensor testing on cord blood is the extra step needed to collect the blood with a syringe from the cord. A sampling device that collects a fixed volume of blood directly from the cord would streamline the process.

It is very likely that performance and reference ranges observed here in cord blood could apply to neonatal capillary or venous blood collected within the first 24 hours of life but this was not evaluated during the study.

The study was conducted in a period critically influenced by the COVID-19 pandemic. Travel restrictions resulted in a delayed study start, reduced enrolment in one clinic (WPA), and a protracted enrolment duration of the study overall. Fewer than planned FGD were conducted—including planned discussions at key time points during the study—and they occurred in a single clinical site providing a possibly narrower point of view on the usability topics explored. Additional staff stressors and human resource limitations due to COVID-19 and the political unrest in 2021 were not assessed but may have influenced the results of both the technical and usability components of the study.

For peer review only

CONCLUSIONS

The "STANDARD G6PD" Biosensor is a reliable POC tool to support the perinatal care of newborns at higher risk of neonatal hyperbilirubinemia by demonstrating very high sensitivity in identification of deficient newborns and high sensitivity in identification of female newborns with intermediate activity. Its use by trained personnel in rural clinics and birthing centers with a high prevalence of G6PD deficiency, together with assessment of bilirubin levels before discharge, has the potential to avert disability and death from hyperbilirubinaemia.

Extending use of the Biosensor for newborn testing in countries where it is already deployed for malaria case management in resource-constrained settings [37], would provide a higher return on this investment. Use of Biosensor in populations with prevalent G6PD deficiency outside malaria endemic regions might increase the benefit-cost ratio of universal screening [38] in all settings [39].

ACKNOWLEDGMENTS

The authors wish to thank all the mothers for theier collaboration and understanding; the study would not have been possible without the hard work and dedication of all SMRU staff involved, especially during such a difficult time of political unrest and COVID-19 pandemic. Acknowledgments are also extended to SD Biosensor for donating the devices and the tests for the study.

ETHICAL APPROVAL

The study was approved by Oxford Tropical Research Ethics Committee, UK (OxTREC 532-19), the Mahidol University Faculty of Tropical Medicine Ethics Committee, Thailand (TMEC 19-048, MUTM 2019-080-02) and the Tak Province Border Community Ethics Advisory Board (TCAB201904). Written informed consent was obtained from literate mothers and midwives; a thumbprint was obtained in the presence of a literate witness for illiterate mothers.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

AUTHORS' CONTRIBUTION

Substantial contributions to the conception or design of the work: GB, AH, FN, VIV and RM. Acquisition, analysis or interpretation of data for the work: GB, MEG, EW, GG, PP, PKM, LA, NSW, SW, KKA, AH, BH, FN, VIC and RM. Drafting the work or revising it critically for important intellectual content: GB, MEG, EW, GG, PP, PKM, LA, NSW, SW, KKA, AH, BH, FN, VIC and RM. Final approval of the version to be published: all authors. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: all authors.

COMPETING INTERESTS

None declared.

FUNDING

The study was supported by a grant to GB by the Wellcome Trust Institutional Translational Partnership Awards: Thailand Major Overseas Programme (WT-iTP-2019/004). SMRU is supported by the Wellcome Trust [grant 220111]. For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any author accepted manuscript version arising from this submission. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

DATA AVAILABILITY STATEMENT

De-identified participant data are available from the Mahidol Oxford Tropical Medicine Data Access Committee upon request from this link: https://www.tropmedres.ac/units/moru-bangkok/bioethics-engagement/datasharing.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1
2
3 476 REFERENCES
4
5 477
6

7 478 1. Bhutani VK, Zipursky A, Blencowe H, Khanna R, Sgro M, Ebbesen F, et al. Neonatal
8 479 hyperbilirubinemia and Rhesus disease of the newborn: incidence and impairment estimates
9 480 for 2010 at regional and global levels. Pediatric research. 2013;74 Suppl 1:86-100.
10
11 481 2. Bhutani VK, Vilms RJ, Hamerman-Johnson L. Universal bilirubin screening for severe
12 482 neonatal hyperbilirubinemia. J Perinatol. 2010;30 Suppl:S6-15.
13 483 3. Olusanya BO, Kaplan M, Hansen TWR. Neonatal hyperbilirubinaemia: a global
14 484 perspective. Lancet Child Adolesc Health. 2018;2(8):610-20.
15 485 4. Lauer BJ, Spector ND. Hyperbilirubinemia in the newborn. Pediatr Rev.
16 486 2011;32(8):341-9.
17
18 487 5. Bancone G, Gornsawun G, Peerawaranun P, Penpitchaporn P, Paw MK, Poe DD, et al.
19 488 Contribution of genetic factors to high rates of neonatal hyperbilirubinaemia on the
20 489 Thailand-Myanmar border. Plos Global Public Health. 2022.
21
22 490 6. Cappellini MD, Fiorelli G. Glucose-6-phosphate dehydrogenase deficiency. Lancet.
23 491 2008;371(9606):64-74.
24 492 7. Kaplan M, Hammerman C, Vreman HJ, Stevenson DK, Beutler E. Acute hemolysis and
25 493 severe neonatal hyperbilirubinemia in glucose-6-phosphate dehydrogenase-deficient
26 494 heterozygotes. The Journal of pediatrics. 2001;139(1):137-40.
27
28 495 8. Meloni T, Forteleoni G, Dore A, Cutillo S. Neonatal hyperbilirubinaemia in
29 496 heterozygous glucose-6-phosphate dehydrogenase deficient females. British journal of
30 497 haematology. 1983;53(2):241-6.
31 498 9. Molou E, Schulpis KH, Thodi G, Georgiou V, Dotsikas Y, Papadopoulos K, et al.
32 499 Glucose-6-Phosphate Dehydrogenase (G6PD) deficiency in Greek newborns: the
33 500 Mediterranean C563T mutation screening. Scandinavian journal of clinical and laboratory
34 501 investigation. 2014;74(3):259-63.
35
36 502 10. WHO Working Group. Glucose-6-phosphate dehydrogenase deficiency. Bulletin of
37 503 the World Health Organization. 1989;67(6):601-11.
38
39 504 11. Howes RE, Piel FB, Patil AP, Nyangiri OA, Gething PW, Dewi M, et al. G6PD deficiency
40 505 prevalence and estimates of affected populations in malaria endemic countries: a
41 506 geostatistical model-based map. PLoS medicine. 2012;9(11):e1001339.
42
43 507 12. Bancone G, Chu CS, Somsakchaicharoen R, Chowwiwat N, Parker DM,
44 508 Charunwatthana P, et al. Characterization of G6PD Genotypes and Phenotypes on the
45 509 Northwestern Thailand-Myanmar Border. PloS one. 2014;9(12):e116063.
46 510 13. Turner C, Carrara V, Aye Mya Thein N, Chit Mo Mo Win N, Turner P, Bancone G, et al.
47 511 Neonatal intensive care in a Karen refugee camp: a 4 year descriptive study. PloS one.
48 512 2013;8(8):e72721.
49
50 513 14. Thielemans L, Gornsawun G, Hanboonkunupakarn B, Paw MK, Porn P, Moo PK, et al.
51 514 Diagnostic performances of the fluorescent spot test for G6PD deficiency in newborns along
52 515 the Thailand-Myanmar border: A cohort study. Wellcome open research. 2018;3:1.
53 516 15. Health NCCfWsCs. NICE. Neonatal jaundice: clinical guideline2010.
54
55 517 16. Bancone G, Chowwiwat N, Somsakchaicharoen R, Poodpanya L, Moo PK, Gornsawun
56 518 G, et al. Single Low Dose Primaquine (0.25mg/kg) Does Not Cause Clinically Significant
57 519 Haemolysis in G6PD Deficient Subjects. PloS one. 2016;11(3):e0151898.
58 520 17. Asiimwe C, Kyabayinze DJ, Kyalisiima Z, Nabakooza J, Bajabaite M, Counihan H, et al.
59
60 521 Early experiences on the feasibility, acceptability, and use of malaria rapid diagnostic tests at

For peer review only

peripheral health centres in Uganda-insights into some barriers and facilitators. Implement Sci. 2012;7:5.
18. Bancone G, Gilder ME, Chowwiwat N, Gornsawun G, Win E, Cho WW, et al. Prevalences of inherited red blood cell disorders in pregnant women of different ethnicities living along the Thailand-Myanmar border. Wellcome open research. 2017;2:72.
19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-10.
20. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. Indian pediatrics. 2011;48(4):277-87.
21. Thielemans L, Peerawaranun P, Mukaka M, Paw MK, Wiladphaingern J, Landier J, et al. High levels of pathological jaundice in the first 24 hours and neonatal hyperbilirubinaemia in an epidemiological cohort study on the Thailand-Myanmar border. PloS one. 2021;16(10):e0258127.
22. Aynalem YA, Mulu GB, Akalu TY, Shiferaw WS. Prevalence of neonatal hyperbilirubinaemia and its association with glucose-6-phosphate dehydrogenase deficiency and blood-type incompatibility in sub-Saharan Africa: a systematic review and meta-analysis. BMJ Paediatr Open. 2020;4(1):e000750.
23. Bernardo J, Nock M. Pediatric Provider Insight Into Newborn Screening for Glucose-6-Phosphate Dehydrogenase Deficiency. Clin Pediatr (Phila). 2015;54(6):575-8.
24. DelFavero JJ, Jnah AJ, Newberry D. Glucose-6-Phosphate Dehydrogenase Deficiency and the Benefits of Early Screening. Neonatal Netw. 2020;39(5):270-82.
25. Wong RJ, Montiel C, Kunda M, Stevenson DK, Bhutani VK. A novel point-of-care device for measuring glucose-6-phosphate dehydrogenase enzyme deficiency. Semin Perinatol. 2021;45(1):151356.
26. Alam MS, Kibria MG, Jahan N, Thriemer K, Hossain MS, Douglas NM, et al. Field evaluation of quantitative point of care diagnostics to measure glucose-6-phosphate dehydrogenase activity. PloS one. 2018;13(11):e0206331.
27. Pal S, Bansil P, Bancone G, Hrutkay S, Kahn M, Gornsawun G, et al. Evaluation of a Novel Quantitative Test for Glucose-6-Phosphate Dehydrogenase Deficiency: Bringing Quantitative Testing for Glucose-6-Phosphate Dehydrogenase Deficiency Closer to the Patient. The American journal of tropical medicine and hygiene. 2019;100(1):213-21.
28. Pal S, Myburgh J, Bansil P, Hann A, Robertson L, Gerth-Guyette E, et al. Reference and point-of-care testing for G6PD deficiency: Blood disorder interference, contrived specimens, and fingerstick equivalence and precision. PloS one. 2021;16(9):e0257560.
29. Zobrist S, Brito M, Garbin E, Monteiro WM, Clementino Freitas S, Macedo M, et al. Evaluation of a point-of-care diagnostic to identify glucose-6-phosphate dehydrogenase deficiency in Brazil. PLoS neglected tropical diseases. 2021;15(8):e0009649.
30. Kosaryan M, Mahdavi MR, Jalali H, Roshan P. Why does the Iranian national program of screening newborns for G6PD enzyme deficiency miss a large number of affected infants? Pediatric hematology and oncology. 2014;31(1):95-100.
31. Doherty AN, Kring EA, Posey YF, Maisels MJ. Glucose-6-phosphate dehydrogenase activity levels in white newborn infants. The Journal of pediatrics. 2014;164(6):1416-20.
32. Kaplan M, Abramov A. Neonatal hyperbilirubinemia associated with glucose-6-phosphate dehydrogenase deficiency in Sephardic-Jewish neonates: incidence, severity, and the effect of phototherapy. Pediatrics. 1992;90(3):401-5.
33. Ellis D, Sewell CE, Skinner LG. Reticulocyte enzymes and protein synthesis. Nature. 1956;177(4500):190-1.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

34. Srivastava A, Evans KJ, Sexton AE, Schofield L, Creek DJ. Metabolomics-Based Elucidation of Active Metabolic Pathways in Erythrocytes and HSC-Derived Reticulocytes. Journal of proteome research. 2017;16(4):1492-505.

35. Engel N, Ghergu C, Matin MA, Kibria MG, Thriemer K, Price RN, et al. Implementing radical cure diagnostics for malaria: user perspectives on G6PD testing in Bangladesh. Malaria journal. 2021;20(1):217.

36. Gerth-Guyette E, Adissu W, Brito M, Garbin E, Macedo M, Sharma A, et al. Usability of a point-of-care diagnostic to identify glucose-6-phosphate dehydrogenase deficiency: a multi-country assessment of test label comprehension and results interpretation. Malaria journal. 2021;20(1):307.

37. Chu CS, Bancone G, Kelley M, Advani N, Domingo GJ, Cutiongo-de la Paz EM, et al. Optimizing G6PD testing for Plasmodium vivax case management: why sex, counseling, and community engagement matter. Wellcome open research. 2020;5:21.

38. Vidavalur R, Bhutani VK. Economic evaluation of point of care universal newborn screening for glucose-6-Phosphate dehydrogenase deficiency in United States. J Matern Fetal Neonatal Med. 2021:1-9.

39. Watchko JF, Kaplan M, Stark AR, Stevenson DK, Bhutani VK. Should we screen newborns for glucose-6-phosphate dehydrogenase deficiency in the United States? J Perinatol. 2013;33(7):499-504.

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

Quantitative G6PD point-of-care test can be used reliably on cord blood to identify male and female newborns at increased risk of neonatal hyperbilirubinaemia: a mixed method study

Figure 1. Distribution of G6PD enzymatic activity from cord blood samples detected by gold standard spectrophotometry assay (A) and Biosensor (B) according to sex and genotype

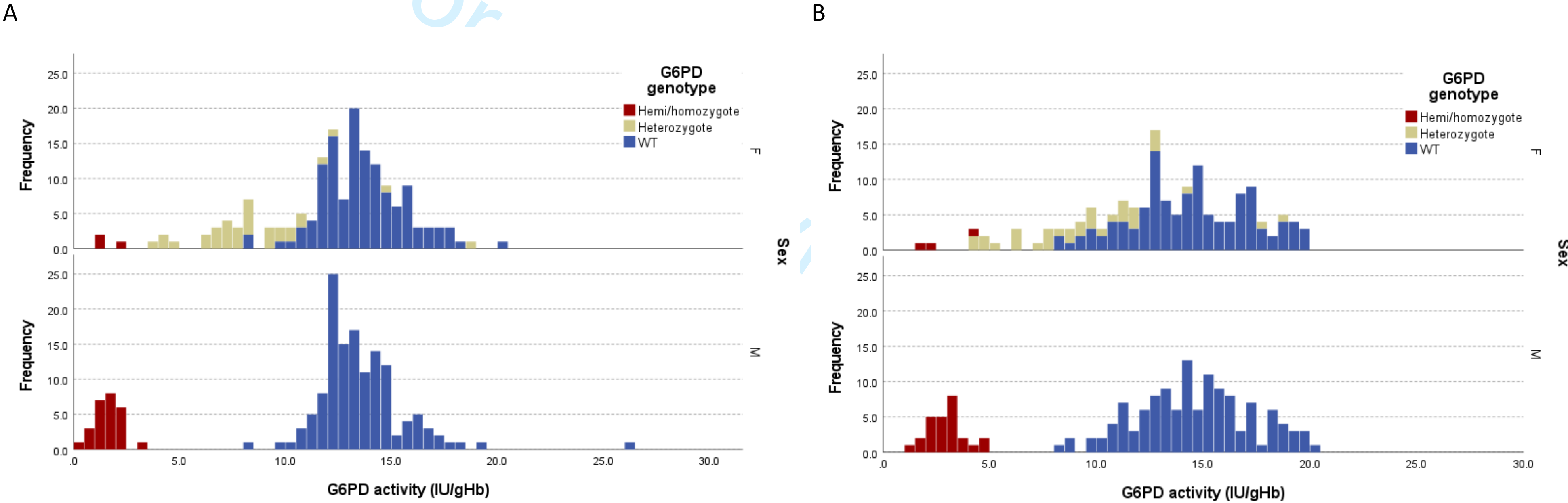
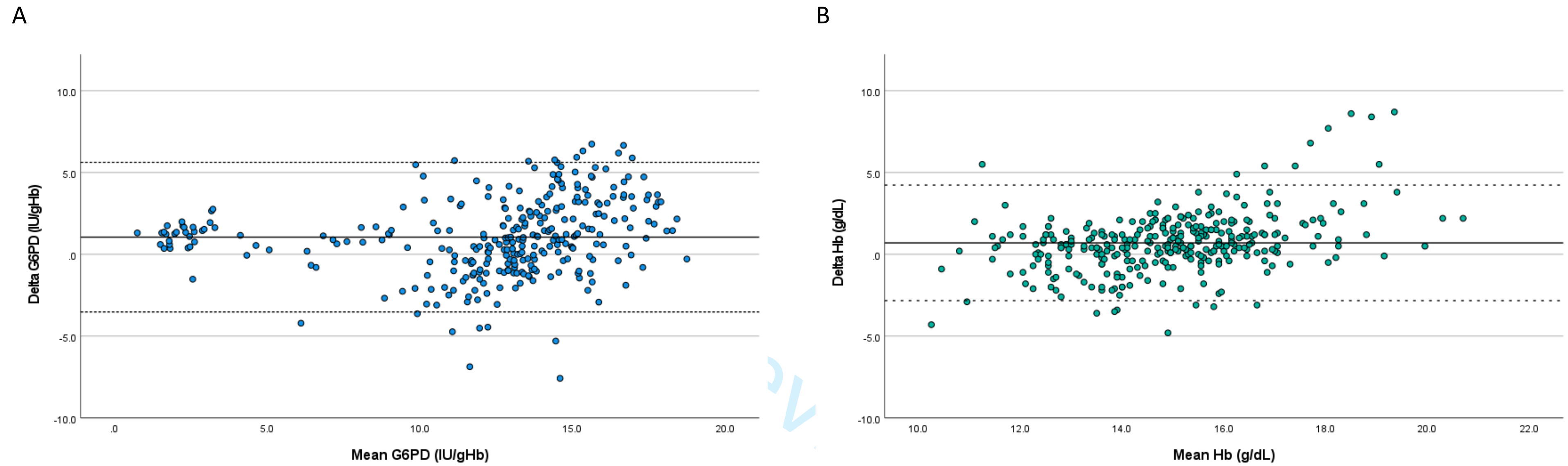


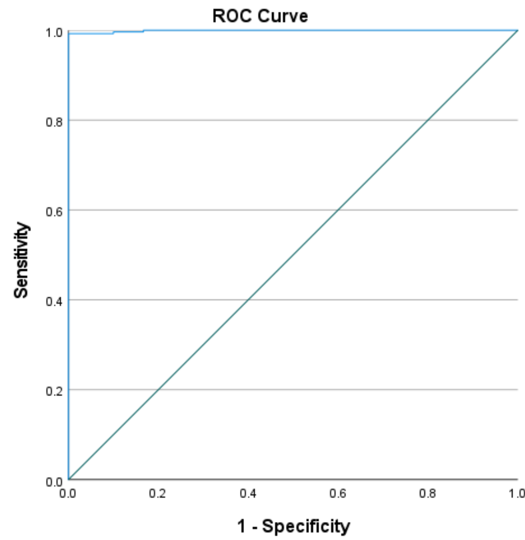
Figure 2. Bland Altman plot of G6PD activity (A) and haemoglobin levels (B) in cord blood comparing gold standard spectrophotometry to Biosensor

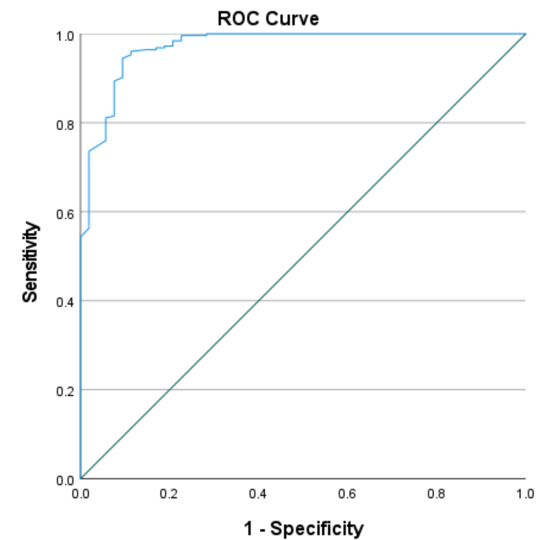


Delta G6PD=G6PD Biosensor- G6PD Spectrophotometry
Full horizontal line indicates mean difference (1.05IU/gHb); dotted horizontal lines indicate limits of agreement (-3.52 to 5.62IU/gHb)

Delta Hb=Hb Biosensor- Hb Spectrophotometry
Full horizontal line indicates mean difference (0.70g/dL); dotted horizontal lines indicate limits of agreement (-2.83 to 4.23g/dL)

1
2
3
4
5
6
7 **Figure 3**. Receiver Operating Characteristic curve of Biosensor for 30% activity (A) and 70% activity (B) thresholds.
8
9 A) 30% activity B) 70% activity





10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Quantitative G6PD point-of-care test can be used reliably on cord blood to identify male and female newborns at increased risk of neonatal hyperbilirubinaemia: a mixed method study

Supplementary Figures and Tables

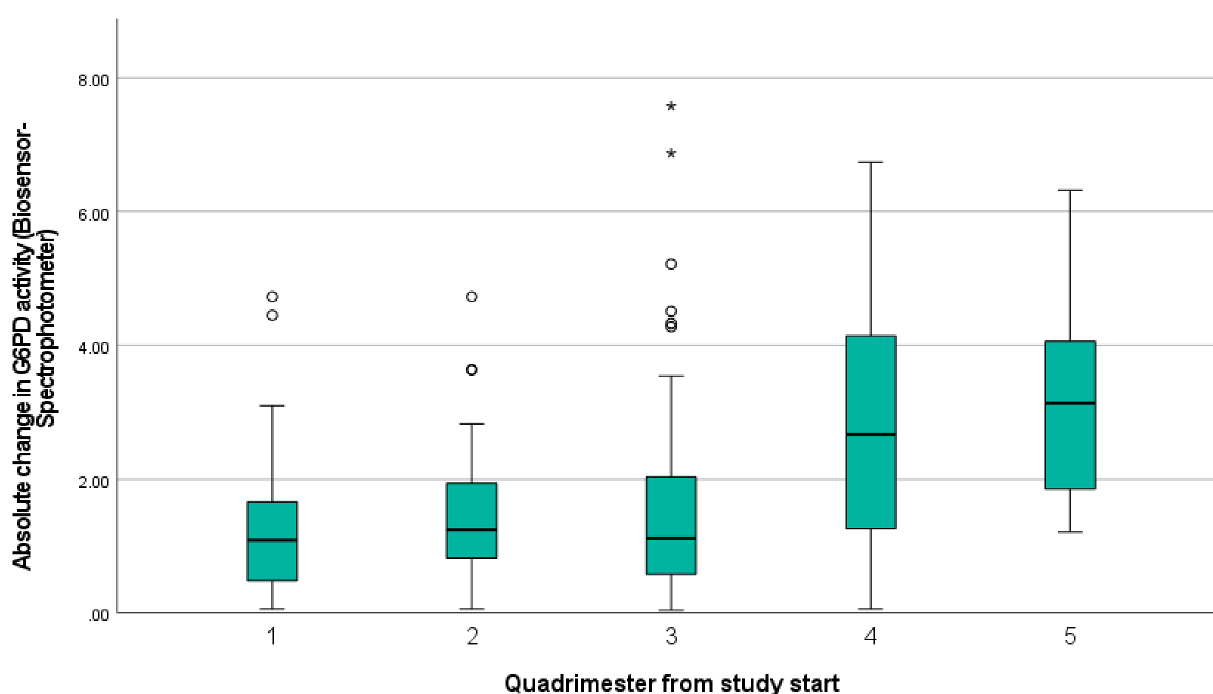
S Table 1. G6PD enzymatic activity (IU/gHb) of cord blood by spectrophotometry according to genotype

G6PD genotype	N	Mean	Std. Deviation	Minimum	Maximum
Hemizygote	26	1.64	0.65	0.09	3.32
Homozygote	3	1.66	0.43	1.38	2.16
Heterozygote	34	8.55	2.97	3.54	18.89
WT	262	13.62	2.02	8.01	26.32
Total	325	12.02	4.14	0.09	26.32

S Table 2. G6PD enzymatic activity (IU/gHb) by Biosensor according to genotype

G6PD genotype	N	Mean	Std. Deviation	Minimum	Maximum
Hemizygote	26	2.87	0.81	1.4	4.6
Homozygote	3	2.70	1.23	1.8	4.1
Heterozygote	34	9.50	3.47	4.0	18.6
WT	244	14.46	2.72	8.1	20.0
Total	307	12.82	4.47	1.4	20.0

S Figure 1. Absolute difference in G6PD activity detected by Biosensor as compared to spectrophotometry over time (only MKT clinic)



For peer review

S Table 3. Characteristics of samples misclassified by Biosensor

Clinic	Year	EGA	Sex	Reference G6PD (IU/gHb)	Reference Hb (g/dL)	Percent activity of reference (%)	Reference phenotype	Biosensor G6PD (IU/gHb)	Biosensor Hb (g/dL)	Percent activity of Biosensor (%)	Percent activity of reference (%)	Biosensor phenotype	G6PD genotype Mahidol	Retics (%)	WBC (10^3/ uL)
MKT	2020	42	F	4.4	15	33	INT	4.3	15.2	30	32	DEF	Heterozygote	1.5	13.6
MKT	2021	40	F	7.1	15.8	54	INT	12.6	16.3	88	95	NOR	Heterozygote	1.3	20.2
MKT	2021	41	F	7.7	14.1	58	INT	12.5	11.5	87	94	NOR	Heterozygote	ND	20.6
MKT	2021	39	F	8.0	15.3	60	INT	10.9	11.7	76	82	NOR	WT	2.3	19.3
MKT	2021	39	F	8.2	14.3	62	INT	4	15.7	28	30	DEF	Heterozygote	2.2	21.1
MKT	2021	39	F	8.5	14.1	64	INT	11.8	14.1	82	89	NOR	Heterozygote	ND	ND
WPA	2021	39	F	9.4	13.3	71	NOR	9.8	13.7	68	74	INT	Heterozygote	1.6	13.8
WPA	2021	38	F	10.2	14.8	77	NOR	7.5	15.3	52	56	INT	Heterozygote	2.1	11.6
WPA	2021	39	F	10.9	15.6	82	NOR	8.8	16.7	61	66	INT	WT	1.6	12.5
MKT	2020	38	F	11.4	16.7	86	NOR	9.3	18.8	65	70	INT	WT	4.8	7.1
WPA	2021	39	F	11.7	14.3	88	NOR	8.1	16.5	56	61	INT	WT	1.8	11.1
WPA	2020	39	F	11.8	15.8	89	NOR	9.8	15.3	68	74	INT	WT	1.9	14.3
MKT	2020	39	F	12.1	12.6	91	NOR	9	13.2	63	68	INT	WT	3.9	14.4
MKT	2021	40	F	14.2	14	107	NOR	9.7	16.3	67	73	INT	WT	1.7	11.3
MKT	2021	37	F	15.1	11.2	113	NOR	8.2	12.1	57	62	INT	WT	3.7	15.5

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

S Table 4. Phototherapy treatment in newborns with EGA≥38 weeks with different G6PD phenotypes

G6PD phenotype by spectrophotometry	PT	No PT	% PT	RR	95%CI	P_{Fisher}
Deficient	8	20	28.6	4.9	2.3-10.5	<0.001
Intermediate	3	17	15.0	2.6	0.8-8.1	0.13
Normal	15	242	5.8			reference
G6PD phenotype by Biosensor						
Deficient	9	21	30.0	5.4	2.5-11.6	<0.001
Intermediate	2	20	9.1	1.7	0.4-6.8	0.49
Normal	13	223	5.5			reference

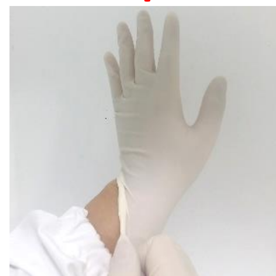
For peer review only

SD G6PD BIOSENSOR (for sample)

SMRU. Shoklo Malaria Research Unit

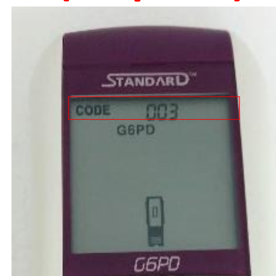
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

Prepare the machine, test device and buffer (step 1-8) BEFORE doing the blood collection (step 9)

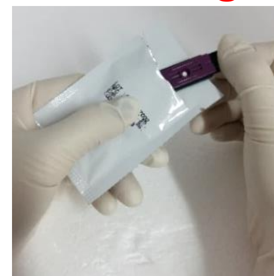

1. Put on gloves

2. Insert codechip (For first time using or open new box of test device)

3. Check the expiry date printed on the foil pouch



4. Check that codechip number on screen correspond to test device

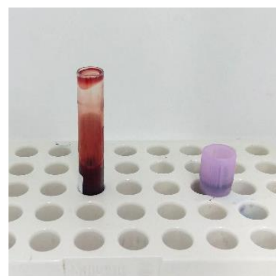

5. Open the foil pouch and take a test device out and hold the test in the right side

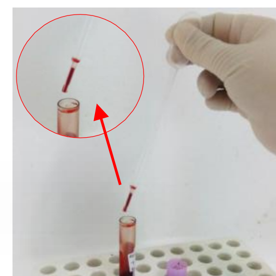
6. Insert the test device

7. Open flap chamber

8. Open buffer tube and place on rack

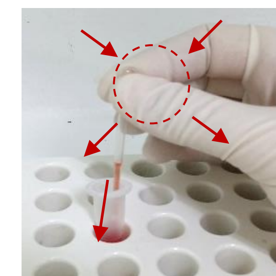
9. Mix sample tube well by inverting* 10 times
*Gently, no bubbles and no shaking


10. Place on rack


11. Collect blood by using Pasteur pipette

12. Drop blood on para film one drop (Avoid to make bubble)

13. Hold the EZI tube horizontally, and touch the tip of the EZI tube to the blood specimen. Do not close hole.


14. Mix the blood specimen with extraction buffer by pressing and releasing the EZI tube 10 times

15. **Discard used EZI tube in the sharp bin**

16. Take new EZI tube

17. Hold the EZI tube horizontally, and touch the tip of the EZI tube to the mixed blood specimen. Do not close hole.

18. Apply mixed specimen to the specimen application hole of the test device

19. Close the flap chamber immediately after applying

20. Wait for 2 min for the test result to appear on the screen (Check date) and report results on the logbook

21. Take the used test device out and discard in sharp bin

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
Participants	**6**	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	6
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	NA
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	8-9
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	8-9
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	5
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	8-9
	15	How indeterminate index test or reference standard results were handled	5
	16	How missing data on the index test and reference standard were handled	NA
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	NA
	18	Intended sample size and how it was determined	8
RESULTS			
Participants	**19**	Flow of participants, using a diagram	
	20	Baseline demographic and clinical characteristics of participants	10
	21a	Distribution of severity of disease in those with the target condition	11-12
	21b	Distribution of alternative diagnoses in those without the target condition	NA
	22	Time interval and any clinical interventions between index test and reference standard	NA
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	table 2
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	11-12
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	16
	27	Implications for practice, including the intended use and clinical role of the index test	17
OTHER INFORMATION			
	28	Registration number and name of registry	NA
	29	Where the full study protocol can be accessed	10
	30	Sources of funding and other support; role of funders	9-10

For peer review only

cc BY SA

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

BMJ Open

Technical evaluation and usability of a quantitative G6PD POC test in cord blood: a mixed methods study in a low-resource setting

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2022-066529.R1
Article Type:	Original research
Date Submitted by the Author:	09-Nov-2022

Complete List of Authors:	Bancone, Germana; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University; University of Oxford Centre for Tropical Medicine and Global Health Gilder, Mary Ellen; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Win, Elsie; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Gornsawun, Gornpan; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Penpitchaporn, Penporn; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Moo, Phaw Khu; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Archasuksan, Laypaw; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Wai, Nan San; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Win, Sylverine; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Aung, Ko Ko; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University Hashmi, Ahmar; Institute for Implementation Science, University of Texas Health Sciences Center (UTHealth), Houston, TX USA; Department of Health Promotion and Behavioral Sciences, School of Public Health, University of Texas Health Sciences Center (UTHealth) Hanboonkunupakarn, Borimas; Mahidol-Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University Nosten, Francois; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University; University of Oxford Centre for Tropical Medicine and Global Health Carrara, Verena Ilona; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University; Institute of Global Health, Faculty of Medicine, University of Geneva, Switzerland

	McGready, Rose; Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University; University of Oxford Centre for Tropical Medicine and Global Health
Primary Subject Heading:	Diagnostics
Secondary Subject Heading:	Global health, Haematology (incl blood transfusion), Paediatrics, Patient-centred medicine
Keywords:	PAEDIATRICS, NEONATOLOGY, MOLECULAR BIOLOGY

SCHOLARONE™
Manuscripts

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

For peer review only

TITLE

Technical evaluation and usability of a quantitative G6PD POC test in cord blood: a mixed methods study in a low-resource setting

Germana Bancone[1,2,*], Mary Ellen Gilder[1], Elsie Win[1], Gornpan Gornsawun[1], Penporn Penpitchaporn[1], Paw Khu Moo[1], Laypaw Archasuksan[1], Nan San Wai[1], Sylverine Win[1], Ko Ko Aung[1], Ahmar Hashmi[4,5], Borimas Hanboonkunupakarn[3], Francois Nosten[1,2], Verena I Carrara[1,6], Rose McGready[1,2]

1. Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Mae Sot, Thailand.

2. Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom.

3. Mahidol-Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand.

4. Institute for Implementation Science, University of Texas Health Sciences Center (UTHealth), Houston, TX USA

5. Department of Health Promotion and Behavioral Sciences, School of Public Health, University of Texas Health Sciences Center (UTHealth) Houston, TX USA

6. Institute of Global Health, Faculty of Medicine, University of Geneva, Switzerland

* corresponding author: germana@tropmedres.ac

ABSTRACT

Objectives: New point-of-care (POC) quantitative G6PD testing devices developed to provide safe radical cure for P. vivax malaria may be used to diagnose G6PD deficiency in newborns at risk of severe neonatal hyperbilirubinaemia, improving clinical care, and preventing related morbidity and mortality.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Methods: We conducted a mixed-methods study analyzing technical performance and usability of the "STANDARD G6PD" Biosensor when used by trained midwives on cord blood samples at two rural clinics on the Thailand-Myanmar border.

Results: In 307 cord blood samples, the Biosensor had a sensitivity of 1.000 (95%CI 0.859-1.000) and a specificity of 0.993 (95% CI 0.971-0.999) as compared to gold standard spectrophotometry to diagnose G6PD deficient newborns using a receiving operator characteristic (ROC) analysis-derived threshold of ≤4.8IU/gHb. The Biosensor had a sensitivity of 0.727 (95%CI: 0.498-0.893) and specificity of 0.933 (95%CI: 0.876-0.969) for 30-70% activity range in females using ROC analysis-derived range of 4.9 to 9.9IU/gHb. These thresholds allowed identification of all G6PD deficient neonates and 80% of female neonates with intermediate phenotypes.

Need of phototherapy treatment for neonatal hyperbilirubinaemia was higher in neonates with deficient and intermediate phenotypes as diagnosed by either reference spectrophotometry or Biosensor.

Focus group discussions found high levels of learnability, willingness, satisfaction, and suitability for the Biosensor in this setting. The staff valued the capacity of the Biosensor to identify newborns with G6PD deficiency early ("We can know that early, we can counsel the parents about the chances of their children getting jaundice") and at the POC, including in more rural settings ("Because we can know the right result of the G6PD deficiency in a short time. Especially for the clinic which does not have a lab").

Conclusions: The Biosensor is a suitable tool in this resource-constrained setting to identify newborns with abnormal G6PD phenotypes at increased risk of neonatal hyperbilirubinaemia.

Strengths and limitations of this study

- The technical performance of the G6PD quantitative point-of-care diagnostic device was assessed against the current gold-standard spectrophotometric assay.
- Receiving operator characteristic analysis was used to identify the best diagnostic thresholds.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

- Usability among clinical personnel from a resource-constrained setting was analysed using a conceptual framework developed for similar settings.
- Fewer than planned focus group discussions were conducted and they occurred in a single clinical site providing a possibly narrower point of view on the usability topics explored.

INTRODUCTION

Pathologically increased levels of bilirubin during the first week of life, i.e. neonatal hyperbilirubinaemia (NH), are common and dangerous for the developing brain. The most severe form of NH, kernicterus, causes neurological sequelae in >80% of neonates (56/100,000 live births globally, [1]). Every year, an estimated twenty-four million newborns are at risk of NH-related adverse outcomes with three-quarters of mortality occurring in sub-Saharan Africa and South Asia [1,2]. These preventable deaths and disabilities disproportionally affect neonates where universal health care and treatment options are scarce, if not absent [3].

Several genetic and clinical factors influence the timing and evolution of NH, including G6PD deficiency, ABO blood group incompatibility, prematurity/low birth weight and sepsis [4]. Early identification of these risk factors can dramatically improve neonatal clinical management during the first days of life [5].

The enzymatic defect of G6PD deficiency, caused by mutations on the X-linked G6PD gene, is a known risk factor for increased levels of bilirubin after birth and it is associated with susceptibility to drug-induced haemolysis [6]. Risk of severe NH is increased in both deficient and heterozygous newborns with abnormal phenotypes [7-9] and universal neonatal screening of G6PD deficiency is supported by WHO in populations where more than 3-5% of males are affected [10].

G6PD deficiency is particularly prevalent among neonates from tropical regions [11], where clinical care is often provided in a non-tertiary hospital or clinic context. Knowledge of G6PD status by medical

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

staff and parents can aid in avoiding potentially haemolytic antibiotics or other agents (such as naphthalene), improved follow-up, and heightened awareness of signs and symptoms of severe NH.

G6PD deficiency is very common among the Karen and Burman population along the Thailand-Myanmar border (9-18% in males, [12]) where it is associated with an increased risk to develop NH requiring phototherapy both in G6PD deficient (over 4-fold [13]) and in heterozygous females (over 2-fold [5]) as compared to wild type genotype neonates. In a recent study, screening of G6PD by qualitative Fluorescent Spot Test (FST) on cord blood failed to identify almost 10% of G6PD deficient neonates [14].

Demonstrating usability of new quantitative Point-Of-Care (POC) G6PD diagnostic tests by locally trained clinical staff can inform clinical deployment in this setting and in other rural settings. This study assessed the technical performance and usability of the "G6PD STANDARD" (SD Biosensor, Korea) test when used by trained midwives in two clinics along the Thailand-Myanmar border.

MATERIALS AND METHODS

Study design

A mixed-methods study was conducted to evaluate both the technical performance of the "G6PD STANDARD" (SD Biosensor, Korea) test (henceforth "Biosensor") and its usability by midwives in a non-tertiary setting. G6PD enzymatic activity and haemoglobin concentration measured by the device were compared to the gold standard reference spectrophotometric assay and haematology analyser, respectively. Performance of the G6PD fluorescent spot test (FST) currently used routinely at the point-of-care, was also compared to the reference and new test.

Following local staff training, user proficiency was assessed before study start; usability was explored using focus group discussions (FGD) at the end of the study.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Study setting and population

The study was conducted in SMRU clinics situated along the Thailand-Myanmar border in Tak province (Thailand) where free antenatal care and birthing services are provided for migrant women of predominantly Karen and Burman ethnicity.

SMRU midwives come from the same population as the pregnant women and patients seeking care at SMRU clinics. The majority of midwives have primary or secondary education and receive clinical training on-site. Pregnant women attending SMRU clinics at Wang Pha (WPA) and Maw Ker Thai (MKT) were informed about the study at regular antenatal care visits in the 3rd trimester. Informed consent procedures and eligibility assessments for mothers were completed before labour commenced. Eligibility of neonates was assessed immediately after delivery, and those born at an estimated gestational age (EGA) by ultrasound ≥35 weeks with no severe maternal complications at delivery and no severe neonatal illness were included. In order to allow laboratory analyses to be performed within 30 hours from collection, only neonates born during week days were included. For all neonates, indication for starting phototherapy treatment followed the recommendations of the UK NICE guidelines [15].

Blood analyses for technical evaluation of Biosensor

Two milliliters of cord blood were collected into EDTA from the umbilical cord using an established SMRU SOP. An aliquot of anticoagulated blood was used by the midwives in the delivery room for the Biosensor following manufacturer's instructions within one hour of collection (Supplementary file 1). Tests were repeated if the test result was an error or "HI" (a result obtained when G6PD activity is very high, outside the instrument analytic range). High-level and low-level Biosensor controls were run weekly or monthly (depending on availability) from April 2020 until May 2021.

An aliquot of anticoagulated blood was analysed by G6PD fluorescent spot test (FST) at the clinical laboratory. The remaining blood was stored at 4°C until shipment to the central SMRU laboratory on the same day.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1
2
3 125 Gold standard reference testing for G6PD and haemoglobin were performed by spectrophotometric
4
5 126 assay and haematology analyser (with complete blood and reticulocyte counts), respectively, at the
6
7 127 SMRU central laboratory.
8
9
10 128 G6PD spectrophotometric assay was performed using Pointe Scientific kits (assay kit # G7583-180, Lysis
11
12 129 Buffer # G7583-LysSB). Kinetic determination of G6PD activity at 340 nm was performed using a
13
14 130 SHIMAZU UV-1800 spectrophotometer with temperature controlled cuvette compartment (30°C).
15
16
17 131 Samples were analysed in double and mean activity was expressed in IU/gHb using the Hb
18
19 132 concentration obtained by complete blood count analysis. The final result was calculated using
20
21 133 manufacturer's Temperature Control Factor of 1.37. Two controls (Normal, Intermediate or Deficient;
22
23 134 Analytic Control Systems, Inc. USA) were analysed at every run and results compared to expected
24
25
26 135 ranges provided by manufacturer. Complete blood count was performed using a CeltacF MEK-8222K
27
28 136 haematology analyser (Nihon Kohden, Japan). Three-levels quality controls were run every day and
29
30 137 device maintenance and calibration were performed regularly. Reticulocytes were analysed by
31
32 138 microscopy after staining with supervital staining Crystal Violet.
33
34
35 139 Buffy coat recovered from whole blood after centrifugation was stored at -20°C for later DNA
36
37 140 extraction using standard columns kit (Favorgen Biotech, Taiwan). Genotyping for G6PD common
38
39 141 mutations was performed through established SOPs [16]. Mahidol mutation was analysed in all
40
41 142 samples. Other mutations were only analysed in phenotypically deficient or intermediate samples
42
43 143 (G6PD < 9.31IU/gHb by reference test) with wild type or heterozygote Mahidol genotypes. Viangchan,
44
45
46 144 Chinese-4, Kaiping, Canton, Union and Mediterranean were analysed first and full gene sequence was
47
48 145 performed if none of these mutations were found.
49
50 146
51
52
53 147 ## Biosensor training, user proficiency and usability assessment
54
55 148 Midwives of WPA and MKT SMRU clinics were trained for use of Biosensor and were eligible to
56
57 149 participate in the usability component of the study following informed consent. Two to four training
58
59 150 sessions were provided at each clinic in the local language by an experienced laboratory technician
60

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

(author LA). The sessions lasted from 1 to 2 hours and included a short introduction about the test, a practical demonstration using imitation blood, and supervised use of the biosensor by each midwife. Midwives were allowed to practice the procedure the week following the training prior to taking a user proficiency test. The proficiency test was administered by author LA in the local language and it consisted of a questionnaire (modified from a questionnaire developed by PATH (https://www.finddx.org/wp-content/uploads/2020/09/PATH_STANDARD-G6PD-User-Competency-Assessment-quiz_08oct19.pdf) and direct observation of two consecutive tests. Midwives were asked to explain out-loud their actions while performing the first test. The proficiency test was analysed by authors GB and GG and midwives who scored <85% were re-trained before study start. A visual aid with all critical steps of the procedure was printed and available in the delivery room during the study.

The usability component of the study followed the conceptual framework for acceptance and use of a rapid diagnostic test for malaria proposed by Asiimwe et al. [17] that evaluates 6 components: learnability, willingness, suitability, satisfaction, efficacy, and effectiveness. The focus group discussions (FGD) specifically focused on 4 main themes of learnability, willingness, satisfaction, and suitability. Due to COVID, only two of the planned six total FGD were conducted. The midwives were grouped by their seniority, with senior and junior midwives together, and midwife assistants in a separate group in order to encourage honest and open conversation. One researcher (KKA) facilitated the FGD while an experienced assistant took notes; both were fluent in Burmese and Karen languages used in the FGD. Immediately following the FGD, research staff debriefed and noted main themes of the discussion. FGDs were audio-recorded and subsequently translated and transcribed in English. Two researchers (MG and GB) independently analysed the transcript using thematic analysis based on the pre-set framework [17] using Taguette (a free and open access qualitative data analysis software, https://joss.theoj.org/papers/10.21105/joss.03522) and confirmed findings with KKA. Face-to-face meeting and exchange of notes allowed for triangulation between the researchers.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Blood analysis for assessment of neonatal hyperbilirubinaemia

Routine clinical care for newborns included at least one total serum bilirubin (TSB) test before discharge (around 48h of life) using capillary blood measured on-site by the rapid quantitative bilirubinometer BR-501 (Apel Co. Ldt, Japan).

Sample size and statistical analyses

The expected prevalence of G6PD deficiency in the population living at the border is 9-18% in males and 2-4% in female [12, 16] corresponding to approximately 20-30% heterozygous females, 60% of whom have intermediate activity [18]. Assuming that the proportion of females and males in the neonate population is 50%, 9% were expected to be G6PD deficient and 7% to be G6PD intermediate. In order to obtain 95% CI of the limits of agreement within 0.5 SD of the difference, about 31 neonates with deficiency and 25 with intermediate phenotypes were needed, with a minimum total sample size of 350 samples.

Clinical data were double entered in MACRO and collated with laboratory data; data were analysed using SPSSv27.

Male median (MM) was calculated in all males with wild type genotypes in both the references spectrophotometric assay and the Biosensor. Deficiency was defined as enzymatic activity below 30% of MM by reference spectrophotometry and receiving operator characteristic (ROC)-derived 30% threshold by Biosensor; intermediate phenotypes were defined as enzymatic activity between 30% and 70% of the MM or ROC-derived threshold.

Mean and standard deviation (SD) were reported for continuous variables. Categorical variables were compared by Chi-squared test and ANOVA. Bland-Altman plot was used to inspect correspondence between G6PD activity detected by Biosensor compared to the spectrophotometry assay [19]. Correlation was assessed using Pearson's coefficient of correlation and Interclass Correlation Coefficient (ICC). Area under the curve (AUC) of the ROC curve [20] was calculated at different activity thresholds to analyse clinical performances (i.e. sensitivity and specificity) of the Biosensor. Cohen's

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Kappa coefficient was calculated for categories of phenotypes identified by Biosensor and spectrophotometry.

For analysis of haematologic features and risk of neonatal hyperbilirubinaemia, neonates gestational ages assessed by ultrasound were categorized as ≤38 and >38 weeks according to epidemiologic studies conducted previously in the same population [21].

Statistical significance was assessed at the 5% level.

Patient and Public Involvement statement

At the outset of the study, the research team engaged the local population through a local ethics and research advisory committee, the Tak Province Community Advisory Board, Thailand. This group is comprised of community leaders who were asked to advise on study design, process, and outcomes of interest, and subsequently approved the study (TCAB201904).

RESULTS

A total of 331 cord blood samples were collected between April 2020 and November 2021; six were clotted and excluded from all analysis. Of the remaining 325 samples, 257 (79%) were collected in MKT clinic and 68 in WPA clinic, in 166 (51%) female and 159 male neonates. Mean (SD) of estimated gestational age of newborns was 39.1 (1.0) weeks.

General haematologic characteristics

As expected for this specimen, haematological characteristics of cord blood (Table 1) showed higher white blood cell count, haemoglobin concentrations, reticulocyte counts and larger cellular volumes compared to adult blood. Reticulocyte counts and red cell distribution width were higher in neonates <38 weeks gestational age (P=0.02 and P=0.01 respectively) while the other indexes did not differ by gestational age groups.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

225 **Table 1**. Haematologic characteristics of cord blood samples according to newborn gestational age. Results are shown as mean (SD)

EGA (weeks)	N*	WBC (10^3/ uL)	NEU (10^3/ uL)	LYM (10^3/ uL)	RBC (10^6/ uL)	HGB (g/dL)	HCT (%)	MCV (fL)	MCH (pg)	MCHC (g/dL)	RDW (%)	PLT (10^3/ uL)	Reticulocyte (%)
<38	19	13.1 (3.6)	9.6 (3.3)	2.7 (1.7)	4.3 (0.4)	14.4 (1.7)	48.0 (5.4)	110.9 (6.6)	33.2 (2.7)	29.9 (1.5)	16.8 (1.5)	259.2 (66.2)	2.8 (1.8)
≥38	298	14.3 (3.8)	10.8 (3.6)	2.8 (1.6)	4.5 (0.5)	14.5 (1.7)	49.0 (5.2)	109.0 (7.9)	32.3 (3.0)	29.6 (1.4)	16.0 (1.2)	261.4 (47.7)	2.1 (1.1)
P_{ANOVA}		*0.17*	*0.16*	*0.88*	*0.14*	*0.68*	*0.43*	*0.30*	*0.21*	*0.41*	*0.01*	*0.85*	*0.02*

226

227 * Number of samples analysed by haematology analyser was 317 out of 325; 7 samples were analysed by Hemocue and result used to calculate G6PD
228 enzymatic activity.

229

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

G6PD genotypes

A total of 26 hemizygous mutated males (21 Mahidol, 2 Kaiping, 1 Viangchan, 1 Coimbra, 1 Orissa), 3 homozygous mutated females (Mahidol), 34 heterozygous females (32 Mahidol, 1 Canton, 1 Viangchan) and 262 wild type (129 females and 133 males) were found. Overall allelic frequency of all mutated alleles was 13.4%. The distribution of G6PD activity by spectrophotometry and biosensor associated with different genotypes are shown in Figures 1 and Supplementary Tables 1 and 2.

Fluorescent spot test

The poor performances of the FST in cord blood were confirmed here, with the FST failing to identify 23% (7/30) of deficient neonates and 100% of the intermediate females (22/22; Table 2).

Technical evaluation of Biosensor

Male medians by reference spectrophotometric assay and Biosensor

MM G6PD activity by spectrophotometer was 13.3 IU/gHb giving a 30% threshold of 4.0 IU/gHb for diagnosis of deficiency; intermediate activity (30-70%) in females ranged between 4.1 and 9.3 IU/gHb. The cord blood-specific 30% spectrophotometric threshold identified all the hemizygous male and homozygous female newborns (Figure 1A).

MM of G6PD activity by Biosensor calculated on 307 samples was 14.4 IU/gHb giving a 30% threshold of 4.3 IU/gHb for diagnosis of deficiency. Intermediate activity (30-70%) in females ranged between 4.4 and 10.1 IU/gHb (Figure 1B).

In 7% of cases (23/325), the Biosensor provided an initial result of “HI” activity without a numeric value. Of the 19 samples retested, 14 had “HI” results again and 5 samples had an activity ranging from 17.3 to 20.0 IU/gHb; all samples with initial or confirmed “HI” results were normal by spectrophotometry and had a wild type genotype. Overall, 18 samples (5.5% of the total) did not have a final numeric result by Biosensor but would have been considered “normal”, according to the spectrophotometric assay.

For peer review only

Biosensor performance

Biosensor performance was assessed for 307/325 samples that yielded numeric results. The mean (±1.96SD) difference in enzymatic activity between Biosensor and spectrophotometry was 1.05 IU/gHb (LoA: -3.52 to 5.62 IU/gHb) as represented in the Bland-Altman plot in Figure 2A. A very strong correlation between enzymatic activity by Biosensor and reference spectrophotometry was observed (Pearson's r=0.855, p<0.001; ICC=0.905, p<0.001).

The mean (±1.96SD) difference in Hb between the Biosensor and haematology analyser was 0.70 g/dL (LoA: -2.83 to 4.23 g/dL) (Figure 2B). A moderate correlation between Hb levels by Biosensor and haematology analyser was observed (Pearson's r=0.637, p<0.001; ICC=0.728, p<0.001).

Area under the curve (AUC) of the ROC analysis (Figure 3A) of the 30% threshold was 0.999 (95%CI: 0.997-1.000); ROC analysis showed that 30% of Biosensor MM (4.3IU/gHb) was associated with sensitivity of 0.931 (95%CI: 0.758-0.988) and specificity of 0.989 (95%CI: 0.966-0.997) while a threshold of 4.8IU/gHb had a sensitivity of 1.000 (95%CI: 0.859-1.000) and a specificity of 0.993 (95% CI: 0.971-0.999). This second threshold was therefore used for the subsequent analyses.

AUC of the ROC analysis (Figure 3B) for the 70% threshold was 0.972 (95%CI: 0.949-0.994) and ROC analysis showed that a threshold of 9.9IU/gHb had a better sensitivity and specificity as compared to the 70% of Biosensor MM (10.1 IU/gHb). The ROC-derived threshold had a sensitivity of 0.842 (95%CI: 0.716-0.921) and specificity of 0.984 (95%CI: 0.957-0.995) to identify samples with ≤70% activity and was used for subsequent analyses.

AUC of the ROC analysis for the range 30-70% activity was 0.935 (95%CI: 0.887-0.983); sensitivity and specificity for intermediate phenotypes in females were 0.727 (95%CI 0.498-0.893) and 0.933 (95%CI: 0.876-0.969) respectively based on ROC-derived thresholds as compared to 0.592 (95%CI: 0.390-0.770) and 0.953 (95%CI: 0.897-0.980) using Biosensor MM thresholds.

When comparing phenotypes defined according to the 30% and 70% thresholds of spectrophotometry and ROC-derived threshold for Biosensor (Table 2), the Biosensor correctly identified all deficient and

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

normal males and all deficient females. In females, the Biosensor incorrectly identified 9% (2/22) of intermediate females (activity by spectrophotometry 33% and 62%) as deficient, and 7% (9/130) of phenotypically normal female neonates as intermediate (activity by spectrophotometer ranging from 71% to 113%). It also misdiagnosed 18% (4/22) of intermediate samples as normal. Of these 4 samples, 3 were Mahidol heterozygotes and 1 was a wild type and their enzymatic activity by spectrophotometry ranged from 54% to 64%. Cohen's kappa coefficient was 0.841, p<0.001. Overall, the majority of samples with discordant results (11/15) were identified by the Biosensor as having a "worse" phenotype. Characteristics of the 15 samples with discordant results are reported in Supplementary Table 3.

Table 2. Diagnostic performance of FST and Biosensor as compared to gold standard spectrophotometry.

		Spectrophotometry				
		Male		Female		
		Deficient	Normal	Deficient	Intermediate	Normal
FST	Deficient	20	0	2	0	0
	Normal	6*	133	2*	22	137
	Total	**26**	**133**	**4**	**22**	**137**
Biosensor	Deficient	26	0	4	2[#]	0
	Intermediate	NA	NA	0	16	9[$]
	Normal	0	125	0	4[&]	121
	Total	**26**	**125**	**4**	**22**	**130**

Phenotypes are based on 30% and 70% thresholds for spectrophotometry. For Biosensor, threshold for deficiency is ≤4.8IU/gHb and 4.9 to 9.9IU/gHb for intermediate, both obtained by ROC analysis. Total sample for Biosensor was 307; total sample for FST was 322 (3 samples were not analysed by FST at the clinic)

*Enzymatic activities ranging from 12% to 27% of spectrophotometry MM.

[#] Two Mahidol heterozygotes with activity by spectrophotometry of 33% and 62% of MM.

[$] Two Mahidol heterozygotes and 7 wild type samples with enzymatic activity by spectrophotometry ranging from 71% to 113%.

[&] Three Mahidol heterozygotes and 1 wild type samples with enzymatic activity by spectrophotometry ranging from 54% to 64%.

Characteristics of discordant samples are reported in Supplementary Table 1.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

No difference in results were observed by clinic (ICC=0.899, p<0.001 in MKT and ICC=0.930, p<0.001 in WPA) or user. In MKT clinic where the test was used over 20 months, a trend of larger absolute mean differences in activity (Biosensor - Spectrophotometry) were observed in the last 4-8 months of use as compared to the first 12 months (Supplementary Figure 1).

Risk of neonatal hyperbilirubinaemia

Risk of neonatal hyperbilirubinemia by phenotype (determined by spectrophotometry) was assessed in term neonates (EGA≥38weeks). A significantly larger proportion of G6PD deficient neonates (29%) underwent phototherapy for treatment of NH as compared to G6PD normal (6%, RR[95%CI] =4.9 [2.3-10.5]; P<0.001). A larger proportion of female neonates with intermediate phenotypes (90% of whom were heterozygotes) required phototherapy (15%), although in this small cohort the difference did not reach statistical significance (RR[95%CI] =2.6 [0.8-8.1]; P=0.13; supplementary Table 4. Relative risk by quantitative phenotypes were similar to those already established by genotypes in the same population [5].

Biosensor training, user proficiency and usability assessment

A total of 22 midwives in two clinics were initially trained and completed the users' proficiency test, including 7 senior, 10 junior and 5 assistant midwives. Median (min-max) observed score from the questionnaire (max 7 points) and observed tests (max 18 points) was 22.1 (18-24.5). The median score did not differ by seniority: assistant 21.4 (18.0-23.5), junior 22.0 (19.3-24.5), senior 22.8 (21.0-24.5); most midwives (72%) had a score >21 points (>85% of maximum score). The most common mistakes in the questionnaire were on how to mix the blood and the buffer (pipetting 10 times *vs* shaking the buffer tube) and on volume of blood mixture to transfer into the device. On observation, the most common mistakes were failure to check the date on Biosensor screen and failure to check test expiry date (rated as minor mistakes since expired test strips are automatically recognized by the Biosensor and rejected).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1
2
3 341 Two focus group discussions were held in December 2021 in MKT clinic, four weeks after completion
4
5 342 of the sample collection at that site; one FGD included 6 senior and junior midwives, and one included
6
7 343 6 assistant midwives. Discussions on satisfaction, learnability, willingness, and suitability and future
8
9 344 use are summarized in Table 3. Overall satisfaction was high, although staff were concerned with
10
11 345 invalid results, and found it challenging to dedicate one member of the team to perform the biosensor
12
13 346 test in the delivery room in the busy postpartum period. In terms of learnability, the midwife assistants
14
15 347 reported learning the device more easily, though some were anxious about missing steps. The senior
16
17 348 staff were anxious about mistakes and clotted blood, and reported the need to refer to the instructions
18
19 349 as a problem. Contrary to the positive expressions to keep using the device at the clinic, the midwives'
20
21 350 willingness to use the device was not high and they requested a dedicated staff to perform the test or
22
23 351 the test to be done in the laboratory. In terms of suitability and future use, the midwives found the
24
25 352 results clinically useful and a valuable diagnostic tool in both their setting and field clinics. However,
26
27 353 they were concerned about neglecting clinical care while doing a laboratory test, the cost of the device,
28
29 354 and emphasized the need for good training.
30
31
32
33
34
35 355 **Table 3.** Selected quotes by theme from focus group discussions.

Theme	Quotes
A. Satisfaction	"It is very good for the children. It is good to know if the child has G6PD deficiency or not from birth. The advantage of the device is that it can detect the children without having to do a heel stick on the baby. On the other hand, there is an increase in work…. But now that we are good at using it, it's fine." [FGD1] "Sometimes if someone is doing the test by using the device it means there are fewer staffs to be with mothers and babies which is not good." [FGD1]
B. Learnability	"After the one-time training, we had 1 or 2 times experiences practically. Then we can do it." [FGD2] "I am really scared I will forget the steps." [FGD2] "We have to look at the book very often, if not we forget the process of what to put and how to put it." [FGD1]
C. Willingness	"Facilitator: Yes. What do you think about keeping on using this device in the future? Participant: Of course. It is good. Participant: Yes, it is good. But if we can have a specific staff to do it then it will be better." [FGD2] "To make changes, take out the blood and send it to the lab. Then only lab staff have to

36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	do that." [FGD1]
D. Suitability & Future Use	"Because we can know that early, we can have counseling with the parents about the chances of their children getting yellow skin. We can take time to counsel." [FGD1] "Because we can know the right result of the G6PD deficiency in a short time. Especially for the clinic which doesn't have a lab then it is difficult to know the G6PD status. But with this device, they will only need to take a little blood from the baby and they can know the result of G6PD." [FGD2]

DISCUSSION

This is the first study to assess clinical performance and usability by locally trained health workers of the "STANDARD G6PD" Biosensor test for identification of G6PD deficient and intermediate phenotypes in cord blood. Current data, together with previously collected evidence from clinical trials in the same population [5], clearly indicate that newborn heterozygous girls with G6PD intermediate phenotypes, who are not identified by the FST, are at increased risk of NH and require phototherapy [7, 8]. The availability of a validated POC quantitative test such as the Biosensor and its inclusion in diagnostics guidelines for neonatal care at birth will allow identification of this group of neonates and better clinical care in several settings [22-25]. Together with other easy-to-use non-invasive tools for diagnosis of NH (e.g. Transcutaneous bilirubinometers), this study provides evidence that Biosensor could be used in non-tertiary rural settings for identification of neonates who need referral to higher levels of care. In settings where phototherapy is available, this study indicates that the Biosensor is a better option than FST to support clinical management of neonates. Technical performance of the Biosensor using ROC-derived threshold was comparable to that observed in adult blood in laboratory and field studies [26-29].

The phenotypic classification provided by the Biosensor was superior to the currently available qualitative test (FST) both for deficient and for intermediate phenotypes. Among intermediate phenotypes, 80% were identified as either deficient or intermediate, allowing a better identification of neonates at potential jaundice risk as compared to the currently used FST-based diagnosis [14, 30]. Poor performance of FST can be explained by the higher G6PD enzymatic activity at birth as compared

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1
2
3 377 to adulthood [31, 32]; this is probably the result of several haematological factors including younger
4
5 378 red cell age, increased number of reticulocytes with higher G6PD activity [33, 34] and higher WBC
6
7 379 count [28] as observed here. Importantly, because of higher enzymatic activity in cord blood,
8
9 380 thresholds established in adult blood cannot be used to identify deficient or intermediate phenotypes
10
11 381 by either spectrophotometry or Biosensor at birth and would have missed identification of 10% (3/29)
12
13 382 deficient neonates (2/26 deficient males and 1/4 deficient females) and 86% (19/22) intermediate
14
15 383 females.
16
17 384 Biosensor haemoglobin values had a moderate correlation with those assessed by automatic
18
19 385 haematology analyser. Although cord (and neonatal) blood samples have higher haemoglobin levels
20
21 386 and increased viscosity, Biosensor's performance in measuring G6PD activity was not worse at higher
22
23 387 haemoglobin levels.
24
25 388 While the Biosensor provided a numeric result in 94.5% of cases, in few cases an "error" message or a
26
27 389 "HI" result was obtained which, according to the protocol, required re-analysis of the sample. Samples
28
29 390 that tested "HI" were confirmed to be normal, both phenotypically by spectrophotometry and by
30
31 391 genotype (all wild type). In routine practice it will not be needed to repeat the test in samples showing
32
33 392 "HI" result should the manufacturers include this information in the instructions for use.
34
35 393 The usability component of the study highlighted important themes to be taken into consideration for
36
37 394 future use of the Biosensor at birth. The midwives have been involved in previous research regarding
38
39 395 neonatal jaundice and appreciated the importance of early G6PD diagnosis to identify newborns most
40
41 396 at risk of neonatal hyperbilirubinaemia and to facilitate optimal clinical care and parental counselling.
42
43 397 The non-invasive nature of cord blood analysis was considered an advantage. In this setting, the SMRU
44
45 398 midwives recommended that the test be performed by dedicated staff or by the available laboratory
46
47 399 to assure appropriate clinical care is provided to the newborns and mothers; nevertheless, they
48
49 400 estimated that in more rural contexts it may be appropriate for trained birth attendants to perform
50
51 401 the test. Of note, midwives considered their reliance on reading the visual aid while performing the
52
53 402 test (which is standard practice in laboratories) a weakness and this aspect might need to be taken
54
55
56
57
58
59
60

For peer review only

into account when training clinic field staff. Usability results obtained here might not be generalizable to every other context but there are data being collected in several rural and community-based settings that corroborate ease of use of this device to guide malaria treatment after appropriate training [26, 35, 36] .

Although midwives felt uncertain about properly conducting the test at the beginning of the study, the laboratory data showed highly accurate results in the first 12 months of use and very good results in the latter 8 months, supporting suitability of the test among health care workers without prior experience in diagnostics. Follow up studies should explore the causes of this slight decrease in quality over time which could be attributed to environmental or users' factors as well as device durability over >1 year of use in tropical conditions.

Limitations

A practical limitation of Biosensor testing on cord blood is the extra step needed to collect the blood with a syringe from the cord. A sampling device that collects a fixed volume of blood directly from the cord would streamline the process.

It is very likely that performance and reference ranges observed here in cord blood could apply to neonatal capillary or venous blood collected within the first 24 hours of life but this was not evaluated during the study.

The study was conducted in a period critically influenced by the COVID-19 pandemic. Travel restrictions resulted in a delayed study start, reduced enrolment in one clinic (WPA), and a protracted enrolment duration of the study overall. Fewer than planned FGD were conducted—including planned discussions at key time points during the study—and they occurred in a single clinical site providing a possibly narrower point of view on the usability topics explored. Additional staff stressors and human resource limitations due to COVID-19 and the political unrest in Myanmar in 2021 were not assessed but may have influenced the results of both the technical and usability components of the study.

CONCLUSIONS

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The “STANDARD G6PD” Biosensor is a reliable POC tool to support the perinatal care of newborns at higher risk of neonatal hyperbilirubinemia by demonstrating very high sensitivity in identification of deficient newborns and high sensitivity in identification of female newborns with intermediate activity. Its use by trained personnel in rural clinics and birthing centers with a high prevalence of G6PD deficiency, together with assessment of bilirubin levels before discharge, has the potential to avert disability and death from hyperbilirubinaemia.

Extending use of the Biosensor for newborn testing in countries where it is already deployed for malaria case management in resource-constrained settings [37], would provide a higher return on this investment. Use of Biosensor in populations with prevalent G6PD deficiency outside malaria endemic regions might increase the benefit-cost ratio of universal screening [38] in all settings [39].

Figures Legends and Captions

Figure 1. Distribution of G6PD enzymatic activity from cord blood samples detected by gold standard spectrophotometry assay (A) and Biosensor (B) according to sex and genotype

Figure 2. Bland Altman plot of G6PD activity (A) and haemoglobin levels (B) in cord blood comparing gold standard spectrophotometry to Biosensor

(A) Delta G6PD=G6PD Biosensor- G6PD Spectrophotometry
Full horizontal line indicates mean difference (1.05IU/gHb); dotted horizontal lines indicate limits of agreement (-3.52 to 5.62IU/gHb)
(B) Delta Hb=Hb Biosensor- Hb Spectrophotometry
Full horizontal line indicates mean difference (0.70g/dL); dotted horizontal lines indicate limits of agreement (-2.83 to 4.23g/dL)

Figure 3. Receiver Operating Characteristic curve of Biosensor for 30% activity (A) and 70% activity (B) thresholds.

ACKNOWLEDGMENTS

The authors wish to thank all the mothers for their collaboration and understanding; the study would not have been possible without the hard work and dedication of all SMRU staff involved, especially

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

during such a difficult time of political unrest and COVID-19 pandemic. Acknowledgments are also extended to SD Biosensor for donating the devices and the tests for the study.

ETHICAL APPROVAL

The study was approved by Oxford Tropical Research Ethics Committee, UK (OxTREC 532-19), the Mahidol University Faculty of Tropical Medicine Ethics Committee, Thailand (TMEC 19-048, MUTM 2019-080-02) and the Tak Province Border Community Ethics Advisory Board (TCAB201904). Written informed consent was obtained from literate mothers and midwives; a thumbprint was obtained in the presence of a literate witness for illiterate mothers.

AUTHORS' CONTRIBUTION

Substantial contributions to the conception or design of the work: GB, AH, FN, VIC and RM. Acquisition, analysis or interpretation of data for the work: GB, MEG, EW, GG, PP, PKM, LA, NSW, SW, KKA, AH, BH, FN, VIC and RM. Drafting the work or revising it critically for important intellectual content: GB, MEG, EW, GG, PP, PKM, LA, NSW, SW, KKA, AH, BH, FN, VIC and RM. Final approval of the version to be published: all authors. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved: all authors.

COMPETING INTERESTS

None declared.

FUNDING

The study was supported by a grant to GB by the Wellcome Trust Institutional Translational Partnership Awards: Thailand Major Overseas Programme (WT-iTP-2019/004). SMRU is supported by the Wellcome Trust [grant 220211]. For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any author accepted manuscript version arising from this submission. The

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1
2
3 483 funders had no role in study design, data collection and analysis, decision to publish, or preparation of
4
5 484 the manuscript.
6
7
8
9 485 DATA AVAILABILITY STATEMENT
10
11 486 De-identified participant data are available from the Mahidol Oxford Tropical Medicine Data Access
12
13 487 Committee upon request from this link: https://www.tropmedres.ac/units/moru-bangkok/bioethics-
14
15 488 engagement/datasharing.
16
17
18 489
19
20
21
22 490 REFERENCES
23
24 491
25

26 492 1. Bhutani VK, Zipursky A, Blencowe H, Khanna R, Sgro M, Ebbesen F, et al. Neonatal
27 493 hyperbilirubinemia and Rhesus disease of the newborn: incidence and impairment estimates
28 494 for 2010 at regional and global levels. Pediatric research. 2013;74 Suppl 1:86-100.
29 495 2. Bhutani VK, Vilms RJ, Hamerman-Johnson L. Universal bilirubin screening for severe
30 496 neonatal hyperbilirubinemia. J Perinatol. 2010;30 Suppl:S6-15.
31 497 3. Olusanya BO, Kaplan M, Hansen TWR. Neonatal hyperbilirubinaemia: a global
32 498 perspective. Lancet Child Adolesc Health. 2018;2(8):610-20.
33 499 4. Lauer BJ, Spector ND. Hyperbilirubinemia in the newborn. Pediatr Rev.
34 500 2011;32(8):341-9.
35 501 5. Bancone G, Gornsawun G, Peerawaranun P, Penpitchaporn P, Paw MK, Poe DD, et al.
36 502 Contribution of genetic factors to high rates of neonatal hyperbilirubinaemia on the
37 503 Thailand-Myanmar border. Plos Global Public Health. 2022.
38 504 6. Cappellini MD, Fiorelli G. Glucose-6-phosphate dehydrogenase deficiency. Lancet.
39 505 2008;371(9606):64-74.
40 506 7. Kaplan M, Hammerman C, Vreman HJ, Stevenson DK, Beutler E. Acute hemolysis and
41 507 severe neonatal hyperbilirubinemia in glucose-6-phosphate dehydrogenase-deficient
42 508 heterozygotes. The Journal of pediatrics. 2001;139(1):137-40.
43 509 8. Meloni T, Forteleoni G, Dore A, Cutillo S. Neonatal hyperbilirubinaemia in
44 510 heterozygous glucose-6-phosphate dehydrogenase deficient females. British journal of
45 511 haematology. 1983;53(2):241-6.
46 512 9. Molou E, Schulpis KH, Thodi G, Georgiou V, Dotsikas Y, Papadopoulos K, et al.
47 513 Glucose-6-Phosphate Dehydrogenase (G6PD) deficiency in Greek newborns: the
48 514 Mediterranean C563T mutation screening. Scandinavian journal of clinical and laboratory
49 515 investigation. 2014;74(3):259-63.
50 516 10. WHO Working Group. Glucose-6-phosphate dehydrogenase deficiency. Bulletin of
51 517 the World Health Organization. 1989;67(6):601-11.
52 518 11. Howes RE, Piel FB, Patil AP, Nyangiri OA, Gething PW, Dewi M, et al. G6PD deficiency
53 519 prevalence and estimates of affected populations in malaria endemic countries: a
54 520 geostatistical model-based map. PLoS medicine. 2012;9(11):e1001339.

55
56
57
58
59
60

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

12. Bancone G, Chu CS, Somsakchaicharoen R, Chowwiwat N, Parker DM, Charunwatthana P, et al. Characterization of G6PD Genotypes and Phenotypes on the Northwestern Thailand-Myanmar Border. PloS one. 2014;9(12):e116063.
13. Turner C, Carrara V, Aye Mya Thein N, Chit Mo Mo Win N, Turner P, Bancone G, et al. Neonatal intensive care in a Karen refugee camp: a 4 year descriptive study. PloS one. 2013;8(8):e72721.
14. Thielemans L, Gornsawun G, Hanboonkunupakarn B, Paw MK, Porn P, Moo PK, et al. Diagnostic performances of the fluorescent spot test for G6PD deficiency in newborns along the Thailand-Myanmar border: A cohort study. Wellcome open research. 2018;3:1.
15. Health NCCfWsCs. NICE. Neonatal jaundice: clinical guideline2010.
16. Bancone G, Chowwiwat N, Somsakchaicharoen R, Poodpanya L, Moo PK, Gornsawun G, et al. Single Low Dose Primaquine (0.25mg/kg) Does Not Cause Clinically Significant Haemolysis in G6PD Deficient Subjects. PloS one. 2016;11(3):e0151898.
17. Asiimwe C, Kyabayinze DJ, Kyalisiima Z, Nabakooza J, Bajabaite M, Counihan H, et al. Early experiences on the feasibility, acceptability, and use of malaria rapid diagnostic tests at peripheral health centres in Uganda-insights into some barriers and facilitators. Implement Sci. 2012;7:5.
18. Bancone G, Gilder ME, Chowwiwat N, Gornsawun G, Win E, Cho WW, et al. Prevalences of inherited red blood cell disorders in pregnant women of different ethnicities living along the Thailand-Myanmar border. Wellcome open research. 2017;2:72.
19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-10.
20. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. Indian pediatrics. 2011;48(4):277-87.
21. Thielemans L, Peerawaranun P, Mukaka M, Paw MK, Wiladphaingern J, Landier J, et al. High levels of pathological jaundice in the first 24 hours and neonatal hyperbilirubinaemia in an epidemiological cohort study on the Thailand-Myanmar border. PloS one. 2021;16(10):e0258127.
22. Aynalem YA, Mulu GB, Akalu TY, Shiferaw WS. Prevalence of neonatal hyperbilirubinaemia and its association with glucose-6-phosphate dehydrogenase deficiency and blood-type incompatibility in sub-Saharan Africa: a systematic review and meta-analysis. BMJ Paediatr Open. 2020;4(1):e000750.
23. Bernardo J, Nock M. Pediatric Provider Insight Into Newborn Screening for Glucose-6-Phosphate Dehydrogenase Deficiency. Clin Pediatr (Phila). 2015;54(6):575-8.
24. DelFavero JJ, Jnah AJ, Newberry D. Glucose-6-Phosphate Dehydrogenase Deficiency and the Benefits of Early Screening. Neonatal Netw. 2020;39(5):270-82.
25. Wong RJ, Montiel C, Kunda M, Stevenson DK, Bhutani VK. A novel point-of-care device for measuring glucose-6-phosphate dehydrogenase enzyme deficiency. Semin Perinatol. 2021;45(1):151356.
26. Alam MS, Kibria MG, Jahan N, Thriemer K, Hossain MS, Douglas NM, et al. Field evaluation of quantitative point of care diagnostics to measure glucose-6-phosphate dehydrogenase activity. PloS one. 2018;13(11):e0206331.
27. Pal S, Bansil P, Bancone G, Hrutkay S, Kahn M, Gornsawun G, et al. Evaluation of a Novel Quantitative Test for Glucose-6-Phosphate Dehydrogenase Deficiency: Bringing Quantitative Testing for Glucose-6-Phosphate Dehydrogenase Deficiency Closer to the Patient. The American journal of tropical medicine and hygiene. 2019;100(1):213-21.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

567 28. Pal S, Myburgh J, Bansil P, Hann A, Robertson L, Gerth-Guyette E, et al. Reference and
568 point-of-care testing for G6PD deficiency: Blood disorder interference, contrived specimens,
569 and fingerstick equivalence and precision. PloS one. 2021;16(9):e0257560.
570 29. Zobrist S, Brito M, Garbin E, Monteiro WM, Clementino Freitas S, Macedo M, et al.
571 Evaluation of a point-of-care diagnostic to identify glucose-6-phosphate dehydrogenase
572 deficiency in Brazil. PLoS neglected tropical diseases. 2021;15(8):e0009649.
573 30. Kosaryan M, Mahdavi MR, Jalali H, Roshan P. Why does the Iranian national program
574 of screening newborns for G6PD enzyme deficiency miss a large number of affected infants?
575 Pediatric hematology and oncology. 2014;31(1):95-100.
576 31. Doherty AN, Kring EA, Posey YF, Maisels MJ. Glucose-6-phosphate dehydrogenase
577 activity levels in white newborn infants. The Journal of pediatrics. 2014;164(6):1416-20.
578 32. Kaplan M, Abramov A. Neonatal hyperbilirubinemia associated with glucose-6-
579 phosphate dehydrogenase deficiency in Sephardic-Jewish neonates: incidence, severity, and
580 the effect of phototherapy. Pediatrics. 1992;90(3):401-5.
581 33. Ellis D, Sewell CE, Skinner LG. Reticulocyte enzymes and protein synthesis. Nature.
582 1956;177(4500):190-1.
583 34. Srivastava A, Evans KJ, Sexton AE, Schofield L, Creek DJ. Metabolomics-Based
584 Elucidation of Active Metabolic Pathways in Erythrocytes and HSC-Derived Reticulocytes.
585 Journal of proteome research. 2017;16(4):1492-505.
586 35. Engel N, Ghergu C, Matin MA, Kibria MG, Thriemer K, Price RN, et al. Implementing
587 radical cure diagnostics for malaria: user perspectives on G6PD testing in Bangladesh.
588 Malaria journal. 2021;20(1):217.
589 36. Gerth-Guyette E, Adissu W, Brito M, Garbin E, Macedo M, Sharma A, et al. Usability
590 of a point-of-care diagnostic to identify glucose-6-phosphate dehydrogenase deficiency: a
591 multi-country assessment of test label comprehension and results interpretation. Malaria
592 journal. 2021;20(1):307.
593 37. Chu CS, Bancone G, Kelley M, Advani N, Domingo GJ, Cutiongo-de la Paz EM, et al.
594 Optimizing G6PD testing for Plasmodium vivax case management: why sex, counseling, and
595 community engagement matter. Wellcome open research. 2020;5:21.
596 38. Vidavalur R, Bhutani VK. Economic evaluation of point of care universal newborn
597 screening for glucose-6-Phosphate dehydrogenase deficiency in United States. J Matern
598 Fetal Neonatal Med. 2021:1-9.
599 39. Watchko JF, Kaplan M, Stark AR, Stevenson DK, Bhutani VK. Should we screen
600 newborns for glucose-6-phosphate dehydrogenase deficiency in the United States? J
601 Perinatol. 2013;33(7):499-504.

602

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

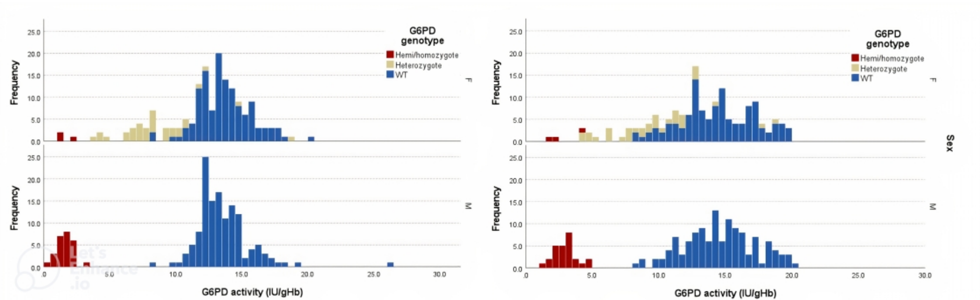


Figure 1. Distribution of G6PD enzymatic activity from cord blood samples detected by gold standard spectrophotometry assay (A) and Biosensor (B) according to sex and genotype

2217x675mm (72 x 72 DPI)

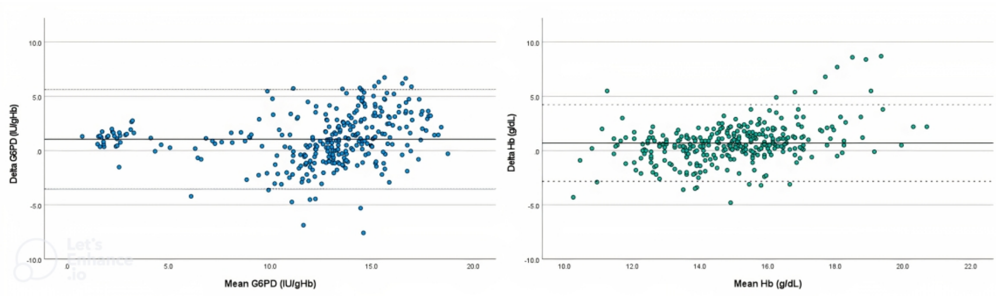


Figure 2. Bland Altman plot of G6PD activity (A) and haemoglobin levels (B) in cord blood comparing gold standard spectrophotometry to Biosensor

2142x633mm (72 x 72 DPI)

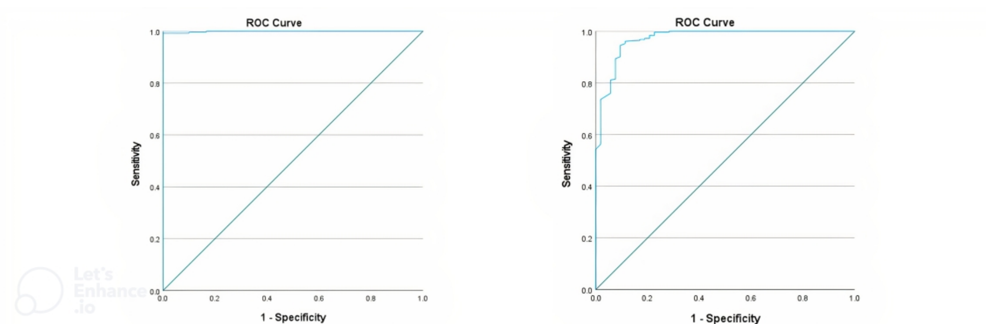


Figure 3. Receiver Operating Characteristic curve of Biosensor for 30% activity (A) and 70% activity (B) thresholds.

1972x642mm (72 x 72 DPI)

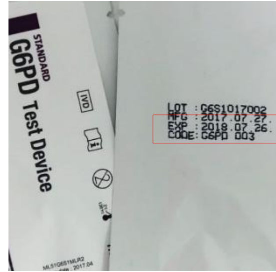
SD G6PD BIOSENSOR (for sample)

SMRU Shoklo Malaria Research Unit

Prepare the machine, test device and buffer (step 1-8) BEFORE doing the blood collection (step 9)

1. Put on gloves

2. Insert codechip (For first time using or open new box of test device)



3. Check the expiry date printed on the foil pouch

4. Check that codechip number on screen correspond to test device

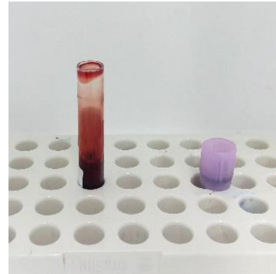
5. Open the foil pouch and take a test device out and hold the test in the right side

6. Insert the test device

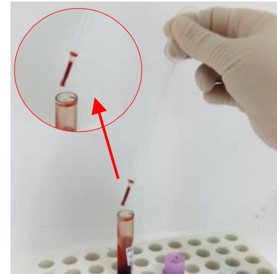
7. Open flap chamber

8. Open buffer tube and place on rack

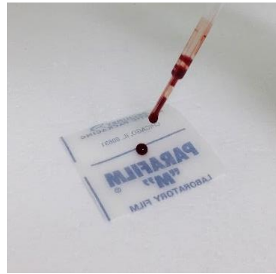
9. Mix sample tube well by inverting* 10 times *Gently, no bubbles and no shaking



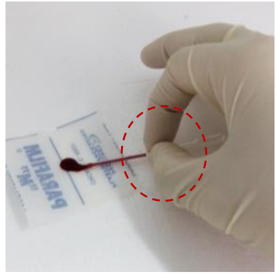
10. Place on rack



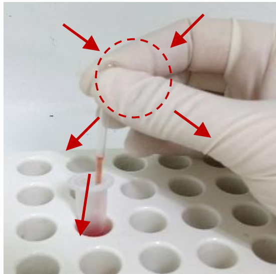
11. Collect blood by using Pasteur pipette



12. Drop blood on para film one drop (Avoid to make bubble)



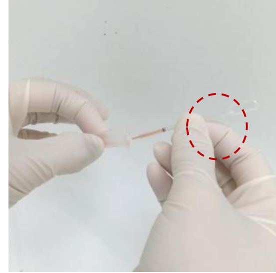
13. Hold the EZI tube horizontally, and touch the tip of the EZI tube to the blood specimen. Do not close hole.



14. Mix the blood specimen with extraction buffer by pressing and releasing the EZI tube 10 times

15. **Discard used EZI tube in the sharp bin**

16. Take new EZI tube



17. Hold the EZI tube horizontally, and touch the tip of the EZI tube to the mixed blood specimen. Do not close hole.

18. Apply mixed specimen to the specimen application hole of the test device

19. Close the flap chamber immediately after applying

20. Wait for 2 min for the test result to appear on the screen (Check date) and report results on the logbook

21. Take the used test device out and discard in sharp bin

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Quantitative G6PD point-of-care test can be used reliably on cord blood to identify male and female newborns at increased risk of neonatal hyperbilirubinaemia: a mixed method study

Supplementary Figures and Tables

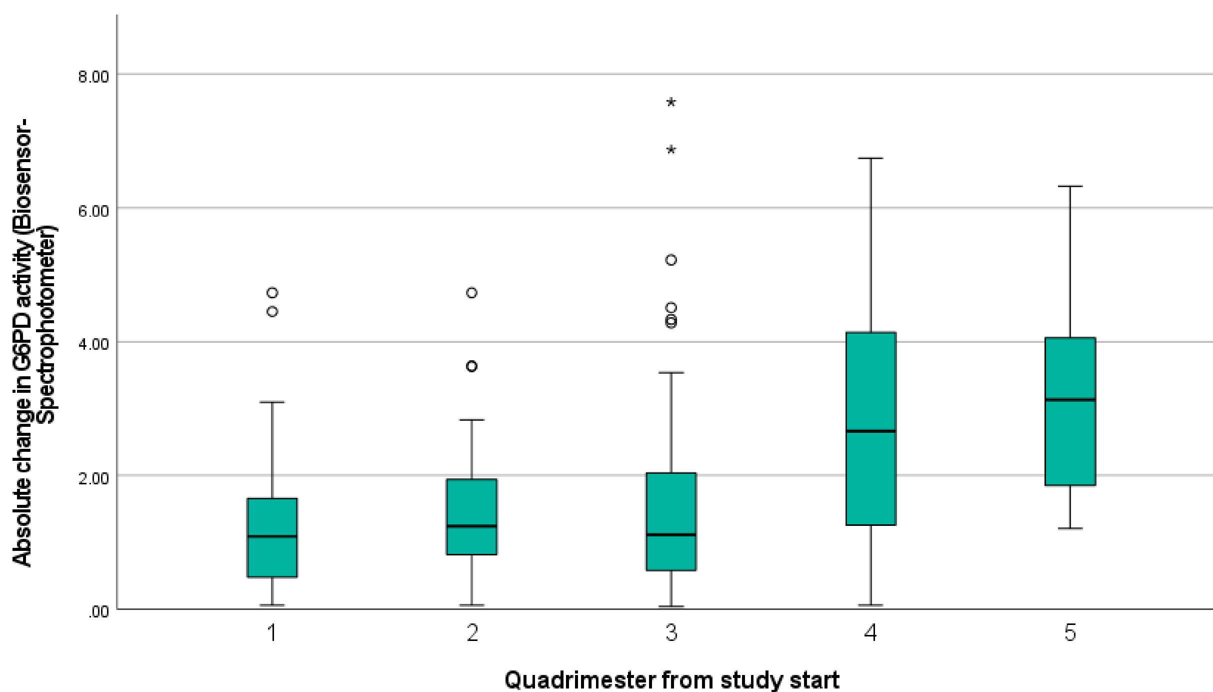
S Table 1. G6PD enzymatic activity (IU/gHb) of cord blood by spectrophotometry according to genotype

G6PD genotype	N	Mean	Std. Deviation	Minimum	Maximum
Hemizygote	26	1.64	0.65	0.09	3.32
Homozygote	3	1.66	0.43	1.38	2.16
Heterozygote	34	8.55	2.97	3.54	18.89
WT	262	13.62	2.02	8.01	26.32
Total	325	12.02	4.14	0.09	26.32

S Table 2. G6PD enzymatic activity (IU/gHb) by Biosensor according to genotype

G6PD genotype	N	Mean	Std. Deviation	Minimum	Maximum
Hemizygote	26	2.87	0.81	1.4	4.6
Homozygote	3	2.70	1.23	1.8	4.1
Heterozygote	34	9.50	3.47	4.0	18.6
WT	244	14.46	2.72	8.1	20.0
Total	307	12.82	4.47	1.4	20.0

S Figure 1. Absolute difference in G6PD activity detected by Biosensor as compared to spectrophotometry over time (only MKT clinic)



For peer review

S Table 3. Characteristics of samples misclassified by Biosensor

Clinic	Year	EGA	Sex	Reference G6PD (IU/gHb)	Reference Hb (g/dL)	Percent activity of reference (%)	Reference phenotype	Biosensor G6PD (IU/gHb)	Biosensor Hb (g/dL)	Percent activity of Biosensor (%)	Percent activity of reference (%)	Biosensor phenotype	G6PD genotype Mahidol	Retics (%)	WBC (10^3/ uL)
MKT	2020	42	F	4.4	15	33	INT	4.3	15.2	30	32	DEF	Heterozygote	1.5	13.6
MKT	2021	40	F	7.1	15.8	54	INT	12.6	16.3	88	95	NOR	Heterozygote	1.3	20.2
MKT	2021	41	F	7.7	14.1	58	INT	12.5	11.5	87	94	NOR	Heterozygote	ND	20.6
MKT	2021	39	F	8.0	15.3	60	INT	10.9	11.7	76	82	NOR	WT	2.3	19.3
MKT	2021	39	F	8.2	14.3	62	INT	4	15.7	28	30	DEF	Heterozygote	2.2	21.1
MKT	2021	39	F	8.5	14.1	64	INT	11.8	14.1	82	89	NOR	Heterozygote	ND	ND
WPA	2021	39	F	9.4	13.3	71	NOR	9.8	13.7	68	74	INT	Heterozygote	1.6	13.8
WPA	2021	38	F	10.2	14.8	77	NOR	7.5	15.3	52	56	INT	Heterozygote	2.1	11.6
WPA	2021	39	F	10.9	15.6	82	NOR	8.8	16.7	61	66	INT	WT	1.6	12.5
MKT	2020	38	F	11.4	16.7	86	NOR	9.3	18.8	65	70	INT	WT	4.8	7.1
WPA	2021	39	F	11.7	14.3	88	NOR	8.1	16.5	56	61	INT	WT	1.8	11.1
WPA	2020	39	F	11.8	15.8	89	NOR	9.8	15.3	68	74	INT	WT	1.9	14.3
MKT	2020	39	F	12.1	12.6	91	NOR	9	13.2	63	68	INT	WT	3.9	14.4
MKT	2021	40	F	14.2	14	107	NOR	9.7	16.3	67	73	INT	WT	1.7	11.3
MKT	2021	37	F	15.1	11.2	113	NOR	8.2	12.1	57	62	INT	WT	3.7	15.5

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

S Table 4. Phototherapy treatment in newborns with EGA≥38 weeks with different G6PD phenotypes

G6PD phenotype by spectrophotometry	PT	No PT	% PT	RR	95%CI	P_{Fisher}
Deficient	8	20	28.6	4.9	2.3-10.5	<0.001
Intermediate	3	17	15.0	2.6	0.8-8.1	0.13
Normal	15	242	5.8			reference
G6PD phenotype by Biosensor						
Deficient	9	21	30.0	5.4	2.5-11.6	<0.001
Intermediate	2	20	9.1	1.7	0.4-6.8	0.49
Normal	13	223	5.5			reference

19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
Participants	**6**	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	6
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	NA
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	8-9
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	8-9
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	5
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	5
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	8-9
	15	How indeterminate index test or reference standard results were handled	5
	16	How missing data on the index test and reference standard were handled	NA
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	NA
	18	Intended sample size and how it was determined	8
RESULTS			
Participants	**19**	Flow of participants, using a diagram	
	20	Baseline demographic and clinical characteristics of participants	10
	21a	Distribution of severity of disease in those with the target condition	11-12
	21b	Distribution of alternative diagnoses in those without the target condition	NA
	22	Time interval and any clinical interventions between index test and reference standard	NA
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	table 2
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	11-12
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	16
	27	Implications for practice, including the intended use and clinical role of the index test	17
OTHER INFORMATION			
	28	Registration number and name of registry	NA
	29	Where the full study protocol can be accessed	10
	30	Sources of funding and other support; role of funders	9-10

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

CC BY SA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

For peer review only