

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

Diagnostic accuracy of self-administered urine glucose test strips as a diabetes screening tool in a low-resource setting in Cambodia

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-019924
Article Type:	Research
Date Submitted by the Author:	09-Oct-2017
Complete List of Authors:	Storey, Helen; PATH, Diagnostics Kahn, Ansley; PATH, Noncommunicable Diseases; van Pelt, Maurits; MoPoTsyo Bun, Socheath; MoPoTsyo Daily, Frances Neogi, Tina; PATH Thompson, Matthew; University of Washington, Department of Family Medicine McGuire, Helen; PATH, Noncommunicable Diseases Weigl, Bernhard; PATH
Primary Subject Heading:	Diabetes and endocrinology
Secondary Subject Heading:	Diagnostics, Global health
Keywords:	Low-resource settings, Diabetes, Diagnostics, Urine glucose test strip, Screening

SCHOLARONE™
Manuscripts

1 Diagnostic accuracy of self-administered urine glucose test strips as 2 a diabetes screening tool in a low-resource setting in Cambodia

3
4 Helen L. Storey; PATH, Seattle WA 98109, USA; hstorey@path.org

5 Maurits H. van Pelt; MoPoTsyo, Phnom Penh, Cambodia; mopotsyo@gmail.com

6 Socheath Bun; MoPoTsyo, Phnom Penh, Cambodia; mopo.medtrainer@gmail.com

7 Frances Daily; Phnom Penh, Cambodia; fdaily@bigpond.com

8 Tina Neogi¹; PATH, Seattle WA 98109, USA; tinaneogi@gmail.com

9 Matthew Thompson; Department of Family Medicine, University of Washington, Seattle WA
10 98195, USA; mjt@uw.edu

11 Helen McGuire; PATH, Washington DC 20001, USA; hmcguire@path.org

12 Bernhard H Weigl²; PATH, Seattle WA 98109, USA; bweigl@intven.com

13
14 * Corresponding author: Helen McGuire, PATH, PO Box 900922 Seattle WA 98109,
15 hmcguire@path.org, 202-540-4360

16 ¹ Present Address: UW Neighborhood Ravenna Clinic, Seattle WA 98105

17 ² Present Address: Intellectual Ventures/Global Good, Bellevue WA 98005, USA

18
19
20 **Abstract** (word count: 248)

21 **Objective:** Screening for diabetes in low resource countries is a growing challenge, necessitating
22 tests that are resource and context appropriate. The aim of this study was to determine the

23 diagnostic accuracy of a self-administered urine glucose test strip compared to alternative
24 diabetes screening tools in a low resource setting of Cambodia.

25 **Design:** Prospective cross-sectional study

26 **Setting:** Members of the Borey Santhepheap community in Cambodia (Phnom Penh
27 Municipality, District Dangkao, Commune Chom Chao).

28 **Participants:** All households on randomly selected streets were invited to participate, and adults
29 at least 18 years of age living in the study area were eligible for inclusion.

30 **Outcomes:** The accuracy of self-administered UGTS positivity, HbA1c >6.5%, and cFBG \geq 126
31 mg/dL were assessed against a composite reference standard of capillary FBG \geq 200 mg/dL or
32 venous blood glucose 2 hours after OGTT \geq 200 mg/dL.

33 **Results:** Of the 1289 participants, 234 (18%) had diabetes based on either cFBG (74, 32%) or the
34 OGTT (160, 68%). The UGTS was 14% sensitive and 99% specific, and failed to identify 201
35 individuals with diabetes, while falsely identifying 7 without diabetes. Those missed by the
36 UGTS had lower venous FBG, lower 2-hour OGTT, and lower HbA1c compared with those
37 correctly diagnosed.

38 **Conclusions:** Low cost, easy to use diabetes tools are essential for low-resource communities
39 with minimal infrastructure. While the UGTS may identify persons with diabetes that might
40 otherwise go undiagnosed in these settings, its poor sensitivity cannot be ignored. The massive
41 burden of diabetes in low-resource settings demands improvements in test technologies.

42 **Keywords:** Diabetes, Low-resource settings, Diagnostics, Urine glucose test strip, Screening,

43

44 **Article Summary (word count: 2261)**

45 **Strengths and limitations of the study**

- 46 • This is one of the first studies to determine the prevalence of diabetes and report on the
47 screening accuracy of urine glucose test strips in Cambodia, which are commonly used as
48 screening tests in this setting.
- 49 • We used a prospective community-based design and had a large sample size with high
50 participation rate, though participation bias towards those able to miss a day of work to
51 attend a clinic visit may still have been an issue.
- 52 • Use of a composite reference test and not evaluating those with cFBG > 200 mg/dL by the
53 OGTT, could have affected our study results, though the use of OGTT allows comparison
54 of our results to those in a number of other studies.
- 55 • The urine glucose test was self-administered and self reported, which is pragmatic and
56 aligns with the practices at MoPoTyso and other clinical settings in Cambodia, however
57 errors in interpreting the test result could influence accuracy.

59 Background

60 According to the International Diabetes Federation (IDF), 415 million adults are living with
61 diabetes globally, almost half of which are undiagnosed, and this number is expected to increase
62 to 642 million by 2040.[1] As is the case for most non-communicable diseases (NCDs), three
63 quarters of those affected live in low- and middle-income countries. In Cambodia for example,
64 there are an estimated 230,000 people with diabetes, who are at risk for the associated micro- and
65 macrovascular complications of this disease, including cardiovascular disease (CVD).[1,2]
66 Strategies to reduce CVD risk may also prevent and control diabetes, which would further reduce
67 rates of eye, kidney, and neural damage due to diabetes complications.[3] To facilitate screening

1
2
3 68 and monitoring for diabetes in these low- and middle-income countries, a low-cost, point-of-care
4
5 69 diagnostic test that is resource and context appropriate is needed.
6
7
8 70

9
10 71 In low-resource settings, urine glucose test strips have been used as diabetes screening tools
11
12 72 because they are inexpensive, noninvasive, and easy to use.[4,5] While these tests do not require
13
14 73 fasting and are user friendly, they can only detect glucose after it has exceeded the threshold for
15
16 74 reabsorption by the kidneys and appears in the urine. The reported threshold varies and is
17
18 75 affected by kidney function.[6] Although their low sensitivity makes them inadequate for use as
19
20 76 a screening tool,[7-9] the World Health Organisation (WHO) acknowledges that they may have a
21
22 77 place in low resource settings where other tests are not possible and the prevalence of
23
24 78 undiagnosed diabetes may be high.[9] Currently many people are not diagnosed until severe
25
26 79 complications develop. Although the sensitivity of the urine test delays diagnosis relative to
27
28 80 other methods, it may provide an opportunity to reduce further advancement of complications.
29
30
31
32

33 81
34
35 82 MoPoTsyo, a nongovernmental organization, provides screening and care services to people with
36
37 83 diabetes and hypertension in Cambodia through an innovative, community-based peer educator
38
39 84 model.[10-12] MoPoTsyo uses urine glucose test strips issued in the community and self-
40
41 85 administered by patients as the initial method of diabetes screening, which has allowed them to
42
43 86 screen over 700,000 adults, followed by confirmation with blood glucose testing for those who
44
45 87 have a positive urine test. The aim of this study was to determine the diagnostic accuracy of a
46
47 88 self-administered urine glucose test strip compared to alternative diabetes screening tools in a
48
49 89 low resource setting of Cambodia. We also explored whether individuals with diabetes who were
50
51 90 detected by urine glucose test strips differed in health status compared to those who were missed
52
53
54
55
56
57
58
59
60

91 by this test but detected by blood glucose measurement. Greater understanding of the
92 performance of this test by the MoPoTsyo program will help to inform its optimal use.

93

94 **Methods**

95 *Study design and procedures*

96 A prospective cross-sectional study was performed among members of the Borey Santhepheap
97 community in Cambodia (Phnom Penh Municipality, District Dangkao, Commune Chom Chao)
98 from November 2013 to October 2014. All households on randomly selected streets were invited
99 to participate by a local peer educator, who described the study to all potential household
100 members. Adults at least 18 years of age living in the study area were eligible for inclusion.
101 Individuals were excluded if they had diabetes or hypertension or had taken medications for
102 diabetes and/or high blood pressure in the last 30 days, had kidney disease, or had received
103 dialysis. Informed consent was obtained from all participants. The protocol was approved by the
104 PATH Research Ethics Committee and the National Ethics Committee for Health Research
105 (Cambodia Institutional Review Board). Study methods and results are reported in alignment
106 with the 2015 STARD recommendations.[13]

107

108 After enrollment, all participants were screened for diabetes using a self-administered and self-
109 reported urine glucose test strip (Sichuan Medicines and Health Products, Chengdu, China).
110 Participants were taught how to use the test strip and read the results with assistance of a color
111 chart, and were given several ways to report results to their peer educator. All participants were
112 then invited to attend the clinic following an 8-hour fast for laboratory confirmed tests for
113 diabetes and associated co-morbid risk factors. Upon arriving at the clinic all participants

1
2
3 114 provided a urine sample, a venous blood sample, and a finger stick blood sample for capillary
4
5 115 fasting blood glucose measurement (cFBG) (On Call Plus glucometer, Acon Laboratories, San
6
7
8 116 Diego, USA). If the cFBG was less than 200 mg/dL they were asked to consume a 75g oral
9
10 117 glucose load for the oral glucose tolerance test (OGTT). The oral glucose load was ingested
11
12 118 within 5 minutes of starting consumption, and two hours after ingestion, further venous blood
13
14 119 and finger stick blood samples were obtained for glucose measurements. During the visit, a
15
16 120 health history was completed based on the WHO STEPS surveillance questionnaire [14] and
17
18 121 blood pressure measured by trained clinical staff using an electronic device (Omron Corporation,
19
20 122 Tokyo, Japan). All devices used in the study were owned and used previously by MoPoTsyo
21
22 123 within the guidelines of the Cambodian Ministry of Health; none of the devices were
23
24 124 investigational. Additional laboratory tests performed included HbA1c (DCA Vantage Analyzer,
25
26 125 Siemens AG, Germany), serum creatinine, glucose, total cholesterol, high-density lipoprotein
27
28 126 cholesterol, and triglycerides (Humalyzer 3000 Chemistry Analyzer, Human Diagnostics,
29
30 127 Germany), spot urine creatinine, protein, and albumin tests (Combilyzer dipstick reader, Human
31
32 128 Diagnostics, Germany).

33
34
35
36
37
38 129
39
40 130 A sample size of 1315 participants was calculated for a desired precision range of 10% and an
41
42 131 estimated sensitivity and specificity of the urine glucose test strip of 21% and 90%, respectively,
43
44 132 which is also sufficient for analysis of HbA1c, OGTT, and FBG as the test strip has the lowest
45
46 133 performance. The sample size for the study was calculated based on Buderer's formula [15],
47
48 134 accounting for a 3% drop-out rate and a 5% national prevalence of diabetes [16].
49
50

51
52 135
53
54
55
56
57
58
59
60

136 **Data Analysis**

137 The index tests of interest were a positive self-administered urine glucose test strip, HbA1c
138 >6.5%, and cFBG \geq 126 mg/dL. Diagnostic accuracy was assessed against a composite reference
139 standard, which was cFBG \geq 200 mg/dL, or venous blood glucose, 2 hours after OGTT \geq 200
140 mg/dL.[17,18] If the participant's cFBG was >200 mg/dL, the patient was considered to have
141 diabetes and an OGTT was not performed. Other measures were defined as follows: Overweight
142 (BMI \geq 25 or waist circumference >90cm for men or >80cm for women[19]), elevated blood
143 pressure (systolic pressure \geq 140mmHg or diastolic pressure \geq 90mmHg), albuminuria (\geq 20
144 mg/L), and elevated albumin/creatinine ratio (\geq 30mg/g). We calculated sensitivity, specificity,
145 positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio
146 (LR+), negative LR (LR-),with 95% confidence intervals (CI).

147
148 Subgroup analyses were used to explore the performance of the urine glucose test strip in
149 participants at increased risk for diabetes mellitus (DM), including age (\geq 50 years), BMI
150 (\geq 25), gender, and waist circumference (>90cm for men or >80cm for women). Prevalence of
151 diabetes by subgroup was compared by chi-squared test. We also explored whether the
152 individuals correctly classified by the urine glucose test strip had better or worse controlled
153 diabetes than those misclassified by the test, as defined by various clinical and laboratory
154 measures. Continuous values were compared using Student's t-test and dichotomous values were
155 compared using the chi-squared test. Data were analyzed using Stata/SE 13.1 (StataCorp LP,
156 Texas, USA).

158 **Results**

1
2
3 159 Of 1328 eligible study subjects, 1316 participated in the study and 1289 were included in the
4
5 160 analysis (Figure 1). Participants were excluded from the analysis if they did not complete the
6
7 161 OGTT due to vomiting or other reasons (16), were not fasting prior to the clinic visit (5), or
8
9 162 reported taking medication for diabetes that day (6). Of the analyzed participants, 75%
10
11 163 (972/1289) were female, mean age was 51 years, 31% had high BMI, and 13% had elevated
12
13 164 blood pressure, although only 8% were taking antihypertensive medications. Characteristics of
14
15 165 the participants included in the analysis are presented in Table 1.
16
17
18
19 166

20
21 167 A total of 234 individuals had diabetes based on the composite reference standard of either
22
23 168 cFBG(74, 32%) or the OGTT (160, 68%), corresponding to a prevalence of 18%. Of the index
24
25 169 tests evaluated, the urine glucose test strip had lower sensitivity (14.1% sensitive), than cFBG
26
27 170 (73.9%), and HbA1c (75.2% sensitive). All three tests offered high specificity (99.3%, 96.8%
28
29 171 and 98.5% respectively) (Table 2). The urine glucose test strip failed to identify 201 individuals
30
31 172 with diabetes (false negatives) and falsely identified seven participants without diabetes (false
32
33 173 positives). The 201 patients with diabetes who were not identified by the urine test had
34
35 174 significantly lower venous FBG, lower 2 hr OGTT, and lower HbA1c compared to those
36
37 175 correctly diagnosed, but were similar in other characteristics (Table 3). The seven false positive
38
39 176 individuals had higher HbA1c, higher systolic BP, and higher proportion receiving treatment for
40
41 177 hypertension than those with true negative results (Table 3).
42
43
44
45
46
47 178

48
49 179 The prevalence of diabetes (diagnosed by the composite reference standard) was significantly
50
51 180 higher in participants who were 50 years of age or older compared to those under 50 years (24%
52
53 181 vs. 9.6%, $p < 0.001$); those with high BMI compared to those with normal BMI (22% vs. 17%,
54
55
56
57
58
59
60

182 p=0.03); and those with greater waist circumference compared to those with normal waist (24%
183 vs. 13%, p<0.001), but was the same in males and females (Table 4). The diagnostic accuracy of
184 the urine glucose test strip was similar among subgroups of patients with various cofactors, with
185 overlapping confidence intervals (Table 4).

187 Discussion

188 Urine glucose test strips had much lower sensitivity than either cFBG or HbA1c, but all three
189 tests offered high specificity. Patients who tested positive with the urine glucose test who were
190 confirmed to have diabetes by the reference standard (true positives) had higher FBG, higher
191 OGTT and higher HbA1c levels compared to the false negative group (urine test negative in
192 patients with diabetes), suggesting that the urine glucose test may identify individuals with poor
193 glycemic control. This suggests a subset of diabetes patients is being identified that is potentially
194 at higher risk of advancing complications or comorbidities, and who may benefit the most from
195 further care [20]. In addition, testing for urine glucose was highly specific (99%), with positive
196 LR_s in the 20s, indicating that when positive, this test is highly indicative of diabetes.

198 The prevalence of diabetes in the MoPoTsyo population in Cambodia was 18%. This is much
199 higher than the national prevalence for Cambodia, which is reported at 3.0%.[1] This may be due
200 to the high proportion of individuals over 50 years of age in our study population, which could
201 be explained by a participation bias towards those who were able to miss a day of work to attend
202 a clinic visit. Additionally, our study took place in a rapidly changing urban population, which
203 had a 2.4 times higher diabetes prevalence in the STEP survey, country report from 2010.[21]

204

1
2
3 205 A wide range of sensitivities for the urine glucose test strip has been reported, and its use
4
5 206 remains controversial. A review in 2000 found six adequately designed studies that reported
6
7 207 performance of urine test strips for glucose.[8] Among these, sensitivities in two reports of
8
9 208 fasting patients were 16% and 35%; two using random samples found sensitivities of 18% and
10
11 209 64%; and three using postprandial and post-load measurements reported sensitivities between
12
13 210 39% and 48%. This review concluded that blood glucose measurements were preferred over
14
15 211 urinary glucose or HbA1c, and particularly, postprandial over fasting measures. Another review
16
17 212 found five studies reporting a range of sensitivity from 18% to 74% for urine glucose test
18
19 213 strips.[7] The review concluded that urine glucose test strips are not sufficient for screening for
20
21 214 diabetes.
22
23
24
25
26
27

28 216 This is one of the first studies to determine the prevalence of diabetes in Cambodia, and report on
29
30 217 the screening accuracy of urine glucose test strips which are commonly used as screening tests in
31
32 218 this setting. We used a prospective community-based design and had a large sample size with
33
34 219 high participation rate. The study had several limitations. Firstly, we used a composite reference
35
36 220 test and those with cFBG > 200 mg/dL were not evaluated by the OGTT. While OGTT is
37
38 221 considered the gold standard reference test for assessing diagnostic accuracy, there has been
39
40 222 some question of its performance. Two studies in China, each on more than 200 participants,
41
42 223 found that the reproducibility of the OGTT was 56% [22] and 66% [23]. Though our choice of
43
44 224 the reference standards, particularly OGTT, could have affected our study results, its use allows
45
46 225 comparison of our results to those in a number of other studies. Second, the urine glucose test
47
48 226 was self-administered and self reported. While this was pragmatic, and aligns with the practices
49
50 227 at MoPoTyso and other clinical settings in Cambodia, errors in interpreting the test result could
51
52
53
54
55
56
57
58
59
60

1
2
3 228 influence accuracy. We were not able to repeat this test when patients attended their clinic visit
4
5 229 as they were fasting at the clinic visit, and thus their urine would not have been the random non-
6
7
8 230 fasting urine test obtained at home. Third, we were not able to obtain hemoglobin levels (or test
9
10 231 for hemoglobin variants) as these tests are not available in this setting, and hence cannot assess
11
12 232 the impact of anemia or hemoglobinopathy on test performance. Fourth, glucose test strip
13
14 233 accuracy may be subject to effects of heat and humidity, we were not able to explore their
15
16
17 234 possible impact on our results.
18

19 235
20
21 236 For clinicians working in settings similar to ours, the question is how useful is the urine glucose
22
23 237 test as a screening or diagnostic test, and is it “better than nothing”? The low sensitivity certainly
24
25
26 238 reduces the value of this test as a screening tool, but the high specificity means that positive tests
27
28 239 can be used to rule in patients with diabetes, suggesting that urine glucose may have some
29
30 240 diagnostic value in this setting. The false positive rate was extremely low, and only 7 patients
31
32 241 without disease were identified as positive by urine glucose test strip. From a population
33
34 242 perspective, the value of a low cost, poorly sensitive yet highly specific test for diabetes is
35
36
37 243 unclear in terms of balancing the opportunity to identify a subset of patients with less well
38
39 244 controlled diabetes who would not have been identified otherwise, with the downside of a high
40
41
42 245 false negative rate.[24]
43

44 246
45
46
47 247 Not surprisingly, usability parameters and cost make urine glucose test strips a highly desirable
48
49 248 test in this and other low-resource settings.[9] Product attributes such as low complexity and
50
51 249 infrastructure requirements, short time to results, and low participant burden greatly contribute to
52
53 250 the acceptability and desirability of the screening tool. The large patient burden and the frequent
54
55
56
57
58
59

1
2
3 251 inability to comply with fasting requirements reduce the feasibility of using OGTT or FBG tests.
4

5 252 While HbA1c testing does not require fasting, current tests are too expensive for use in most
6
7 253 low-income countries. The role of a poorly sensitive test like urine glucose in resource poor
8
9 254 settings such as Cambodia is debatable, on the one hand the test will identify some patients
10
11 255 previously undiagnosed, and assuming treatment can be initiated, reduce severity of
12
13 256 complications from this disease. On the other hand, the test will miss the majority of patients
14
15 257 with diabetes, thus risking a false reassurance, further postponement of diagnosis, and risking
16
17 258 patient's respect for the health care system.
18
19
20
21
22 259

23
24 260 There may be strategies to improve the performance (particularly sensitivity) of the urine glucose
25
26 261 test strip. First, using presence of risk factors such as high waist circumference or BMI, may
27
28 262 increase the pretest probability of diabetes and lead to improved performance. Second, using
29
30 263 random, postprandial, or glucose-loaded measurements may be superior than fasting because the
31
32 264 renal threshold for glucose is more often reached in non-fasting states.[8] Third, improving the
33
34 265 limit of detection may be possible by modifications in the test strip itself, or improvement in the
35
36 266 way it is read either manually (with trained users) or automatically (with electronic reading
37
38 267 devices). Finally, increasing screening frequency may be feasible in low resource settings, if the
39
40 268 urine glucose test strip truly does identify a smaller but more advanced fraction of diabetes
41
42 269 patients.
43
44
45
46
47 270

48 49 271 **Conclusion**

50
51 272 Low cost, easy to use diabetes screening, diagnosis, and monitoring tools are essential for low-
52
53 273 resource communities with minimal infrastructure. While the urine glucose test strip has some
54
55
56
57
58
59
60

1
2
3 274 value as a screening test in these settings, its performance is far from optimal. Progress is
4
5 275 urgently needed to improve the performance, availability, and access of essential testing
6
7
8 276 technologies for diabetes.
9

10 277

11
12 278

13
14
15 279 **List of abbreviations**

16
17 280 urine glucose test strip (UGTS)

18
19 281 International Diabetes Federation (IDF)

20
21 282 non-communicable diseases (NCDs)

22
23 283 cardiovascular disease (CVD)

24
25 284 World Health Organisation (WHO)

26
27 285 capillary fasting blood glucose measurement (cFBG)

28
29 286 oral glucose tolerance test (OGTT)

30
31 287 positive predictive value (PPV)

32
33 288 negative predictive value (NPV)

34
35 289 positive likelihood ratio (LR+)

36
37 290 negative likelihood ratio (LR-)

38
39 291 confidence intervals (CI)

40
41 292 diabetes mellitus (DM)

42
43 293

44
45
46
47
48
49 294 **Declarations**

50
51 295 *Ethical approval and consent to participate*
52
53
54
55
56
57
58
59
60

1
2
3 296 The protocol was approved by the PATH Research Ethics Committee and the National Ethics
4
5 297 Committee for Health Research (Cambodia Institutional Review Board). Informed consent was
6
7
8 298 obtained from all participants.

9
10 299 ***Consent for publication***

11
12 300 Not applicable.

13
14 301 ***Availability of data and material***

15
16
17 302 The datasets used during the current study are available from the corresponding author on
18
19 303 reasonable request.

20
21 304 ***Competing Interests***

22
23
24 305 The authors declare that they have no competing interests.

25
26 306 ***Funding***

27
28 307 This work was supported by a grant from Medtronic Foundation, and received additional support
29
30
31 308 from PATH and the University of Washington Department of Family Medicine. The funding
32
33 309 source had no involvement in study design, data collection, data analysis, data interpretation,
34
35 310 writing of the manuscript, or the decision to publish the results.

36
37 311 ***Authors contributions***

38
39
40 312 MHP, SB, TN, HM and BW designed the study; MHP, SB, TN, and BW implemented the study;
41
42 313 HLS, MT, HM, and BW analysed and interpreted the data; HLS, MHP, FD, MT, HM, and BW
43
44 314 contributed to writing. All authors read and approved the final manuscript.

45
46 315 ***Acknowledgements***

47
48
49 316 We would like to acknowledge the Borey Santhepheap community in Cambodia (Phnom Penh
50
51 317 Municipality, District Dangkao, Commune Chom Chao) for participating in this study. We also
52
53
54
55
56
57
58
59
60

318 acknowledge the input of Dr Annette Fitzpatrick and Dr Jim LoGerfo from the University of
 319 Washington.

320 *Authors' information*

321 Not applicable.

322

323

324 **Tables**

325

326 **Table 1.** Characteristics of included participants.

	Mean (SD) or % n=1289
Age, years	51.4 (14.9)
Female (%)	75.4
BMI ¹	23.2 (4.1)
High BMI (%)	30.5
Waist circumference above cutoff ² (%)	46.1
Systolic blood pressure, mmHg	123.5 (20.6)
Diastolic blood pressure, mmHg	80.8 (12.1)
Elevated blood pressure (%)	12.9
Take treatment for high blood pressure (%)	8.2

327 ¹ n=1288

328 ² >90cm for men, >80cm for women. [19]

329

330

331 **Table 2.** Diagnostic accuracy of urine glucose test strip, capillary fasting glucose, and HbA1c determined
 332 by comparison with the composite reference standard (n=1289)¹.

333

	Urine glucose test strip positive	cFBG ≥126 mg/dL	HbA1c >6.5%
True positive (n)	33	173	176
False positive (n)	7	34	16
False negative (n)	201	61	58
True negative (n)	1048	1021	1039
True diabetes prevalence¹ (95%CI)	18%, 234/1289 (16, 20.4)		
Sensitivity (95% CI)	14.1 (9.90, 19.2)	73.9 (67.8, 79.4)	75.2 (69.2, 80.6)
Specificity (95% CI)	99.3 (98.6, 99.7)	96.8 (95.5, 97.8)	98.5 (97.5, 99.1)
Positive PV (95% CI)	82.5 (67.2, 92.7)	83.6 (77.8, 88.3)	91.7 (86.8, 95.2)
Negative PV (95% CI)	83.9 (81.7, 85.9)	94.4 (92.8, 95.7)	94.7 (93.2, 96.0)
Positive LR (95% CI)	21.3 (9.50, 47.5)	22.9 (16.3, 32.2)	49.6 (30.3, 81.1)
Negative LR (95% CI)	0.90 (0.80, 0.90)	0.30 (0.20, 0.30)	0.30 (0.20, 0.30)

334 ¹ Excludes individuals taking diabetes treatment that day (n=6), did not fast before OGTT as instructed (n=5), or did
 335 not complete the OGTT (n=16). 74 patients with cFBG ≥200 were not tested by OGTT.

336 ² Composite reference standard: OGTT ≥200 mg/dL or cFBG ≥200 mg/dL.

337

338

339
340
341
342**Table 3.** Diagnostic accuracy of the urine glucose test strip by patient characteristics.

Patient characteristic	Diabetic ¹		Non-diabetic ¹	
	True Positive n=33 Mean (SD) or %	False Negative n=201 Mean (SD) or %	False Positive n=7 Mean (SD) or %	True Negative n=1048 Mean (SD) or %
Age	57 (9.3)	58 (10.5)	56 (11.9)	50 (15.5)
Female (%)	81.8	74.6	85.7	75.3
Venous fasting blood glucose	207 (75.3)	166 (73.2)	95 (16.9)	90 (13.1)
Venous blood glucose 2 hrs after OGTT	310 (60.8)	275 (62.2)	115 (43.2)	120 (31.0)
Change in venous blood glucose during OGTT	160 (50.8)	146 (49.8)	20 (47.7)	30 (30.0)
HbA1c	10 (2.3)	8 (2.4)	6 (0.7)	5 (0.5)
BMI	24 (3.9)	24 (3.9)	26 (3.2)	23 (4.1)
High BMI (%)	33.3	36.8	57.1	29.0
Waist circumference above cutoff (%)	60.6	61.7	71.4	42.8
Systolic blood pressure	132 (24.9)	130 (20.6)	146 (14.0)	122 (20.2)
Diastolic blood pressure	85 (9.6)	84 (11.7)	87 (6.5)	80 (12.1)
Elevated blood pressure (%)	15.2	20.9	14.3	11.3
Take treatment for high blood pressure (%)	18.2	11.4	28.6	7.1
Total Cholesterol	242 (62.3)	227 (69.8)	240 (63.1)	213 (56.3)
Proteinuria (n=1116) ² (%)	20.0	17.2	0	3.0
Albuminuria (%)	51.5	47.8	14.3	21.7
Abnormal albumin/creatinine ratio (%)	39.3	39.3	14.3	17.3

¹ Diagnosis by the composite reference standard: venous OGTT ≥ 200 mg/dL or cFBG ≥ 200 mg/dL.² 4 missing values, 169 indeterminate measurements not included in analysis.**Bold = significantly different ($p \leq 0.05$) by Student's t-test or chi-squared test.**343
344
345
346
347
348
349
350**Table 4.** Diagnostic accuracy of urine glucose test strip by participant cofactors (n=1289)¹.

Results	Cofactors							
	Age		BMI ¹		Gender		Waist circumference ³	
	<50	≥ 50	<25	≥ 25	Male	Female	Normal	High
Number of participants	531	758	895	393	317	972	691	598
True positive (n)	8	25	22	11	6	27	13	20
False positive (n)	3	4	3	4	1	6	2	5
False negative (n)	43	158	127	74	51	150	77	124
True negative (n)	477	571	743	304	259	789	599	449
True diabetes prevalence ²	9.6% (7.2, 12.4)	24% (21.0, 27.4)	17% (14.0, 19.3)	22% (18.0, 26.0)	18% (14.0, 22.7)	18% (16.0, 20.8)	13% (11.0, 15.8)	24% (21.0, 27.7)
Sensitivity (95% CI)	15.7 (7.0, 28.6)	13.7 (9.0, 19.5)	14.8 (9.5, 21.5)	12.9 (6.6, 22.0)	10.5 (4.0, 21.5)	15.3 (10.3, 21.4)	14.4 (7.9, 23.4)	13.9 (8.7, 20.6)

Specificity (95% CI)	99.4 (98.2, 99.9)	99.3 (98.2, 99.8)	99.6 (98.8, 99.9)	98.7 (96.7, 99.6)	99.6 (97.9, 100)	99.2 (98.4, 99.7)	99.7 (98.8, 100)	98.9 (97.4, 99.6)
Positive PV (95% CI)	72.7 (39, 94.0)	86.2 (68.3, 96.1)	88 (68.8, 97.5)	73.3 (44.96, 92.2)	85.7 (42.1, 99.6)	81.8 (64.5, 93.0)	86.7 (59.5, 98.3)	80 (59.3, 93.2)
Negative PV (95% CI)	91.7 (89, 94)	78.3 (75.2, 81.3)	85.4 (82.9, 87.7)	80.4 (76.1, 84.3)	83.5 (78.9, 87.5)	84 (81.5, 86.3)	88.6 (86, 90.9)	78.4 (74.8, 81.7)
Positive LR (95% CI)	25.1 (6.9, 91.6)	19.6 (6.9, 55.7)	36.7 (11.1, 121)	10.0 (3.3, 30.5)	27.4 (3.4, 223)	20.2 (8.5, 48.2)	43.4 (10.0, 189)	12.6 (4.8, 33)
Negative LR (95% CI)	0.8 (0.8, 1.0)	0.9 (0.8, 0.9)	0.9 (0.8, 0.9)	0.9 (0.8, 1.0)	0.9 (0.8, 1.0)	0.85 (0.80, 0.91)	0.86 (0.79, 0.94)	0.87 (0.82, 0.93)

¹ Excluded individuals taking diabetes treatment that day (n=6), did not fast before OGTT as instructed (n=5), or did not complete the OGTT (n=16). 74 patients with cFBG \geq 200 were not tested by OGTT; 1 patient had cFBG \geq 200 and also tested OGTT positive.

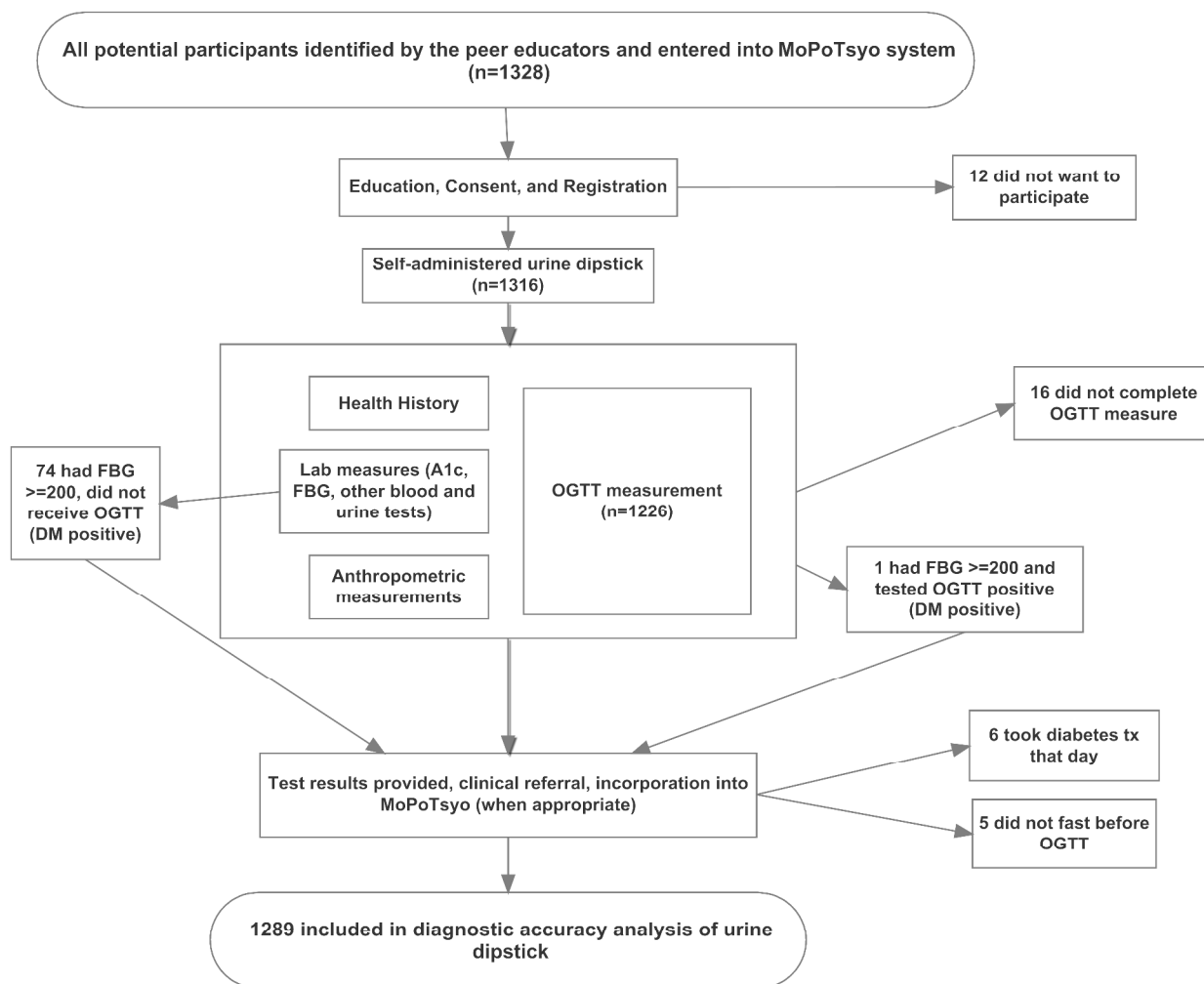
² True prevalence as determined by the composite reference standard. Total number of diabetes diagnoses: 234 (18% prevalence).

³ High Waist circumference = >90cm for men, >80cm for women.[19]

Bold = significantly different ($p \leq 0.05$), chi-squared test.

Figures

Figure 1: Study flow diagram.



Reference List

1. IDF Diabetes Atlas, 7th Edition. International Diabetes Federation . 2015. International Diabetes Federation.
Ref Type: Electronic Citation
2. Forbes JM, Cooper ME: **Mechanisms of diabetic complications.** *Physiol Rev* 2013, **93(1)**: 137-188.
3. Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, Ilanne-Parikka P *et al.*: **Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance.** *N Engl J Med* 2001, **344**: 1343-1350.
4. Dyerberg J, Pedersen L, Aagaard O: **Evaluation of a dipstick test for glucose in urine.** *Clin Chem* 1976, **22**: 205-210.

- 1
2
3 381 5. van der Sande MA, Walraven GE, Bailey R, Rowley JT, Banya WA, Nyan OA *et al.*: **Is there a**
4 382 **role for glycosuria testing in sub-Saharan Africa?** *Trop Med Int Health* 1999, **4**: 506-513.
5
6 383 6. Taal MW, Chertow GM, Marsden PA, Skorecki K, Yu ASL, Brenner BM: **Laboratory**
7 384 **Assessment of Kidney Disease: Glomerular Filtration Rate, Urinalysis, and Proteinuria.** In
8 385 *Brenner & Rector's The Kidney*. Philadelphia: Elsevier; 2012.
9
10 386 7. Wei OY, Teece S: **Best evidence topic report. Urine dipsticks in screening for diabetes**
11 387 **mellitus.** *Emerg Med J* 2006, **23**: 138.
12
13 388 8. Engelgau MM, Narayan KM, Herman WH: **Screening for type 2 diabetes.** *Diabetes Care* 2000,
14 389 **23**: 1563-1580.
15
16 390 9. World Health Organization. Screening for Type 2 Diabetes. Report of a World Health Organization
17 391 and International Diabetes Federation meeting. 2003. Geneva, Switzerland, World Health
18 392 Organization.
19 393 Ref Type: Report
20
21 394 10. van Pelt M: **Improving access to education and care in Cambodia.** *Diabetes Voice* 2009, **54**.
22
23 395 11. Taniguchi D, LoGerfo J, van PM, Mielcarek B, Huster K, Haider M *et al.*: **Evaluation of a multi-**
24 396 **faceted diabetes care program including community-based peer educators in Takeo province,**
25 397 **Cambodia, 2007-2013.** *PLoS One* 2017, **12**: e0181582.
26
27 398 12. van Olmen J, Eggermont N, van Pelt M, Hen H, de Man J, Schellevis F *et al.*: **Patient-centred**
28 399 **innovation to ensure access to diabetes care in Cambodia: the case of MoPoTsyo.** *J Pharm*
30 400 *Policy Pract* 2016, **9**: 1.
31
32 401 13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L *et al.*: **STARD 2015: an**
33 402 **updated list of essential items for reporting diagnostic accuracy studies.** *BMJ* 2015, **351**:
34 403 h5527.
35
36 404 14. World Health Organization. WHO STEPS Surveillance Manual. World Health Organization . 6-13-
37 405 2008. World Health Organization. 4-30-2015.
38 406 Ref Type: Electronic Citation
39
40 407 15. Obuchowski NA: **Sample size calculations in studies of test accuracy.** *Stat Methods Med Res*
41 408 1998, **7**: 371-392.
42
43 409 16. King H, Keuky L, Seng S, Khun T, Roglic G, Pinget M: **Diabetes and associated disorders in**
44 410 **Cambodia: two epidemiological surveys.** *Lancet* 2005, **366**: 1633-1639.
45
46 411 17. **Standards of Medical Care in Diabetes-2016 Abridged for Primary Care Providers.** *Clin*
47 412 *Diabetes* 2016, **34**: 3-21.
48
49 413 18. Bartoli E, Fra GP, Carnevale Schianca GP: **The oral glucose tolerance test (OGTT) revisited.**
50 414 *Eur J Intern Med* 2011, **22**: 8-12.
51
52 415 19. World Health Organization. Waist circumference and waist-hip ratio: report of a WHO expert
53 416 consultation, Geneva, 8-11 December 2008. 2011. Geneva, Switzerland, World Health
54 417 Organization.

1
2
3 418 Ref Type: Report
4

5 419 20. **The effect of intensive treatment of diabetes on the development and progression of long-term**
6 **420 complications in insulin-dependent diabetes mellitus. The Diabetes Control and**
7 **421 Complications Trial Research Group. *N Engl J Med* 1993, **329**: 977-986.**

9 422 21. University of Health Sciences Cambodia, Ministry of Health Cambodia. Prevalence of non-
10 423 communicable disease risk factors in Cambodia; STEPS Survey Country Report. 2010.

11 424 Ref Type: Report

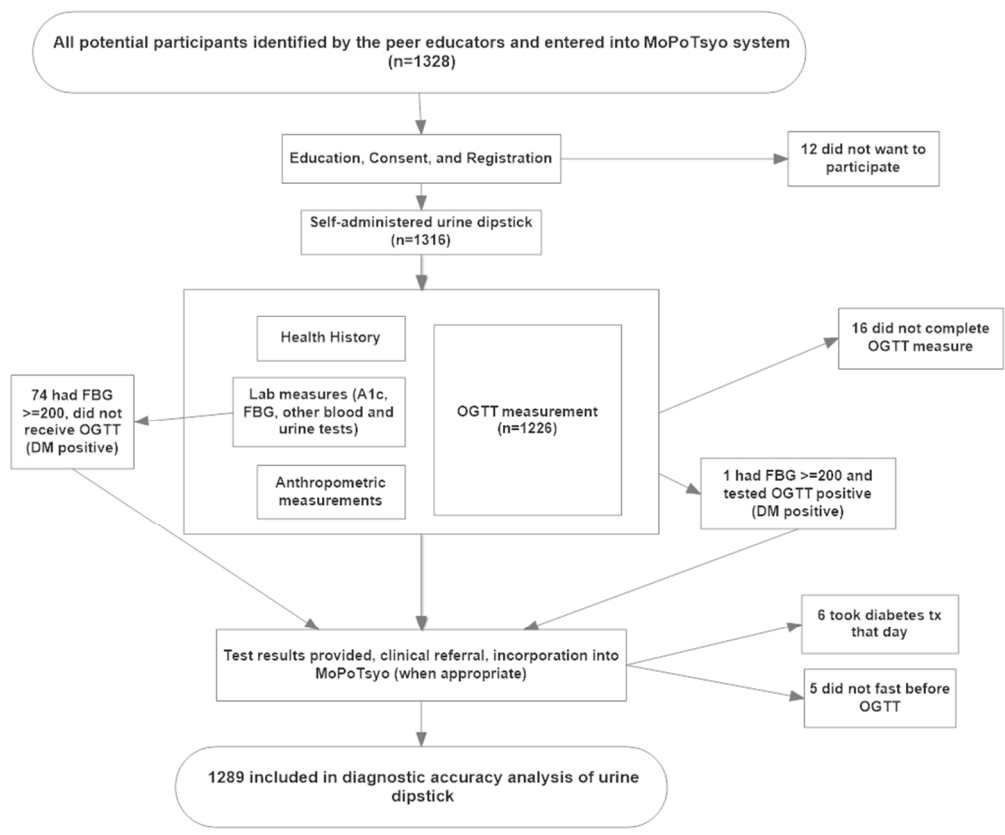
13 425 22. Liu M, Pan CY, Jin MM, Su HY, Lu JM: **[The reproducibility and clinical significance of oral**
14 **426 glucose tolerance test for abnormal glucose metabolism]. *Zhonghua Nei Ke Za Zhi* 2007, **46**:**
15 **427 1007-1010.**

17 428 23. Ko GT, Chan JC, Woo J, Lau E, Yeung VT, Chow CC *et al.*: **The reproducibility and usefulness**
18 **429 of the oral glucose tolerance test in screening for diabetes and other cardiovascular risk**
20 **430 factors. *Ann Clin Biochem* 1998, **35** (Pt 1): 62-67.**

21 431 24. Flessa S, Zembok A: **Costing of diabetes mellitus type II in Cambodia. *Health Econ Rev* 2014,**
23 **432 4: 24.**

24 433
25 434
26 435
27 436

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



86x71mm (300 x 300 DPI)

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)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	1
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3
	4	Study objectives and hypotheses	4
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
<i>Participants</i>	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
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	7
	15	How indeterminate index test or reference standard results were handled	7
	16	How missing data on the index test and reference standard were handled	7
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	7
	18	Intended sample size and how it was determined	6
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	17
	20	Baseline demographic and clinical characteristics of participants	15
	21a	Distribution of severity of disease in those with the target condition	15
	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	5
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	15
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	15
	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	10
	27	Implications for practice, including the intended use and clinical role of the index test	11
OTHER INFORMATION			
	28	Registration number and name of registry	NA
	29	Where the full study protocol can be accessed	NA
	30	Sources of funding and other support; role of funders	14

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

Diagnostic accuracy of self-administered urine glucose test strips as a diabetes screening tool in a low-resource setting in Cambodia

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-019924.R1
Article Type:	Research
Date Submitted by the Author:	09-Jan-2018
Complete List of Authors:	Storey, Helen; PATH, Diagnostics Kahn, Ansley; PATH, Noncommunicable Diseases; van Pelt, Maurits; MoPoTsyo Bun, Socheath; MoPoTsyo Daily, Frances Neogi, Tina; PATH Thompson, Matthew; University of Washington, Department of Family Medicine McGuire, Helen; PATH, Noncommunicable Diseases Weigl, Bernhard; PATH
Primary Subject Heading:	Diabetes and endocrinology
Secondary Subject Heading:	Diagnostics, Global health
Keywords:	Low-resource settings, Diabetes, Diagnostics, Urine glucose test strip, Screening

SCHOLARONE™
Manuscripts

1 **Diagnostic accuracy of self-administered urine glucose test strips as** 2 **a diabetes screening tool in a low-resource setting in Cambodia**

3
4 Helen L. Storey; PATH, Seattle WA 98109, USA; hstorey@path.org

5 Maurits H. van Pelt; MoPoTsyo, Phnom Penh, Cambodia; mopotsyo@gmail.com

6 Socheath Bun; MoPoTsyo, Phnom Penh, Cambodia; mopo.medtrainer@gmail.com

7 Frances Daily; Phnom Penh, Cambodia; fdaily@bigpond.com

8 Tina Neogi¹; PATH, Seattle WA 98109, USA; tinaneogi@gmail.com

9 Matthew Thompson; Department of Family Medicine, University of Washington, Seattle WA
10 98195, USA; mjt@uw.edu

11 Helen McGuire; PATH, Washington DC 20001, USA; hmcguire@path.org

12 Bernhard H Weigl²; PATH, Seattle WA 98109, USA; bweigl@intven.com

13
14 * Corresponding author: Helen McGuire, PATH, PO Box 900922 Seattle WA 98109,
15 hmcguire@path.org, 202-540-4360

16 ¹ Present Address: UW Neighborhood Ravenna Clinic, Seattle WA 98105

17 ² Present Address: Intellectual Ventures/Global Good, Bellevue WA 98005, USA

18
19
20 **Abstract** (word count: 287)

21 **Objective:** Screening for diabetes in low resource countries is a growing challenge, necessitating
22 tests that are resource and context appropriate. The aim of this study was to determine the

23 diagnostic accuracy of a self-administered urine glucose test strip compared to alternative
24 diabetes screening tools in a low resource setting of Cambodia.

25 **Design:** Prospective cross-sectional study

26 **Setting:** Members of the Borey Santhepheap community in Cambodia (Phnom Penh
27 Municipality, District Dangkao, Commune Chom Chao).

28 **Participants:** All households on randomly selected streets were invited to participate, and adults
29 at least 18 years of age living in the study area were eligible for inclusion.

30 **Outcomes:** The accuracy of self-administered urine glucose test strip positivity, HbA1c >6.5%,
31 and capillary fasting blood glucose measurement ≥ 126 mg/dL were assessed against a composite
32 reference standard of capillary fasting blood glucose measurement ≥ 200 mg/dL or venous blood
33 glucose 2 hours after oral glucose tolerance test ≥ 200 mg/dL.

34 **Results:** Of the 1289 participants, 234 (18%) had diabetes based on either capillary fasting blood
35 glucose measurement (74, 32%) or the oral glucose tolerance test (160, 68%). The urine glucose
36 test strip was 14% sensitive and 99% specific, and failed to identify 201 individuals with
37 diabetes, while falsely identifying 7 without diabetes. Those missed by the urine glucose test
38 strip had lower venous fasting blood glucose, lower venous blood glucose 2 hours after oral
39 glucose tolerance test, and lower HbA1c compared with those correctly diagnosed.

40 **Conclusions:** Low cost, easy to use diabetes tools are essential for low-resource communities
41 with minimal infrastructure. While the urine glucose test strip may identify persons with diabetes
42 that might otherwise go undiagnosed in these settings, its poor sensitivity cannot be ignored. The
43 massive burden of diabetes in low-resource settings demands improvements in test technologies.

44 **Keywords:** Diabetes, Low-resource settings, Diagnostics, Urine glucose test strip, Screening,
45

46 Article Summary (word count: 2261)

47 *Strengths and limitations of the study*

- 48 • This is one of the first studies to determine the prevalence of diabetes and report on the
49 screening accuracy of urine glucose test strips in Cambodia, which are commonly used as
50 screening tests in this setting.
- 51 • We used a prospective community-based design and had a large sample size with high
52 participation rate, though participation bias towards those able to miss a day of work to
53 attend a clinic visit may still have been an issue.
- 54 • Use of a composite reference test and not evaluating those with cFBG > 200 mg/dL by the
55 OGTT, could have affected our study results, though the use of OGTT allows comparison
56 of our results to those in a number of other studies.
- 57 • The urine glucose test was self-administered and self reported, which is pragmatic and
58 aligns with the practices at MoPoTyso and other clinical settings in Cambodia, however
59 errors in interpreting the test result could influence accuracy.

61 **Background**

62 According to the International Diabetes Federation (IDF), 415 million adults are living with
63 diabetes globally, almost half of which are undiagnosed, and this number is expected to increase
64 to 642 million by 2040.[1] As is the case for most non-communicable diseases (NCDs), three
65 quarters of those affected live in low- and middle-income countries. In Cambodia for example,
66 there are an estimated 230,000 people with diabetes, who are at risk for the associated micro- and
67 macrovascular complications of this disease, including cardiovascular disease (CVD).[1,2]
68 Strategies to reduce CVD risk may also prevent and control diabetes, which would further reduce

1
2
3 69 rates of eye, kidney, and neural damage due to diabetes complications.[3] To facilitate screening
4
5 70 and monitoring for diabetes in these low- and middle-income countries, a low-cost, point-of-care
6
7 71 diagnostic test that is resource and context appropriate is needed.
8
9
10 72

11
12 73 In low-resource settings, urine glucose test strips have been used as diabetes screening tools
13
14 74 because they are inexpensive, noninvasive, and easy to use.[4,5] While these tests do not require
15
16 75 fasting and are user friendly, they can only detect glucose after it has exceeded the threshold for
17
18 76 reabsorption by the kidneys and appears in the urine. The reported threshold varies and is
19
20 77 affected by kidney function.[6] Although their low sensitivity makes them inadequate for use as
21
22 78 a screening tool,[7-9] the World Health Organisation (WHO) acknowledges that they may have a
23
24 79 place in low resource settings where other tests are not possible and the prevalence of
25
26 80 undiagnosed diabetes may be high.[9] Currently many people are not diagnosed until severe
27
28 81 complications develop. Although the sensitivity of the urine test delays diagnosis relative to
29
30 82 other methods, it may provide an opportunity to reduce further advancement of complications.
31
32 83

33
34
35 84 MoPoTsyo, a nongovernmental organization, provides screening and care services to people with
36
37 85 diabetes and hypertension in Cambodia through an innovative, community-based peer educator
38
39 86 model.[10-12] MoPoTsyo uses urine glucose test strips issued in the community and self-
40
41 87 administered by patients as the initial method of diabetes screening, which has allowed them to
42
43 88 screen over 700,000 adults, followed by confirmation with blood glucose testing for those who
44
45 89 have a positive urine test. The aim of this study was to determine the diagnostic accuracy of a
46
47 90 self-administered urine glucose test strip compared to alternative diabetes screening tools in a
48
49 91 low resource setting of Cambodia. We also explored whether individuals with diabetes who were
50
51
52
53
54
55
56
57
58
59
60

1
2
3 92 detected by urine glucose test strips differed in health status compared to those who were missed
4
5 93 by this test but detected by blood glucose measurement. Greater understanding of the
6
7
8 94 performance of this test by the MoPoTsyo program will help to inform its optimal use.
9
10
11

12 96 **Methods**

13 14 97 *Study design and procedures*

15
16
17 98 A prospective cross-sectional study was performed among members of the Borey Santhepheap
18
19 99 community in Cambodia (Phnom Penh Municipality, District Dangkao, Commune Chom Chao)
20
21 100 from November 2013 to October 2014. All households on randomly selected streets were invited
22
23
24 101 to participate by a local peer educator, who described the study to all potential household
25
26 102 members. Adults at least 18 years of age living in the study area were eligible for inclusion.
27
28 103 Individuals were excluded if they had diabetes or hypertension or had taken medications for
29
30 104 diabetes and/or high blood pressure in the last 30 days, had kidney disease, or had received
31
32
33 105 dialysis. Written informed consent was obtained from all participants. The protocol was approved
34
35 106 by the PATH Research Ethics Committee and the National Ethics Committee for Health
36
37 107 Research (Cambodia Institutional Review Board). Study methods and results are reported in
38
39
40 108 alignment with the 2015 STARD recommendations.[13]
41
42
43

44 110 After enrollment, all participants were screened for diabetes using a self-administered and self-
45
46 111 reported urine glucose test strip (Sichuan Medicines and Health Products, Chengdu, China).
47
48 112 Participants were taught how to use the test strip and read the results with assistance of a color
49
50 113 chart, and were given several ways to report results to their peer educator. All participants were
51
52
53 114 then invited to attend the clinic following an 8-hour fast for laboratory confirmed tests for
54
55
56
57
58
59
60

1
2
3 115 diabetes and associated co-morbid risk factors. Upon arriving at the clinic all participants
4
5 116 provided a urine sample, a venous blood sample, and a finger stick blood sample for capillary
6
7
8 117 fasting blood glucose measurement (cFBG) (On Call Plus glucometer, Acon Laboratories, San
9
10 118 Diego, USA, <https://www.aconlabs.com/us/glucose/on-call/plus-bgms/>). If the cFBG was less
11
12 119 than 200 mg/dL they were asked to consume a 75g oral glucose load for the oral glucose
13
14 120 tolerance test (OGTT). The oral glucose load was ingested within 5 minutes of starting
15
16 121 consumption, and two hours after ingestion, further venous blood and finger stick blood samples
17
18 122 were obtained for glucose measurements. During the visit, a health history was completed based
19
20 123 on the WHO STEPS surveillance questionnaire [14] and blood pressure measured by trained
21
22 124 clinical staff using an electronic device (Omron Corporation, Tokyo, Japan). All devices used in
23
24 125 the study were owned and used previously by MoPoTsyo within the guidelines of the Cambodian
25
26 126 Ministry of Health; none of the devices were investigational. Additional laboratory tests
27
28 127 performed included HbA1c (DCA Vantage Analyzer, Siemens AG, Germany), serum creatinine,
29
30 128 glucose, total cholesterol, high-density lipoprotein cholesterol, and triglycerides (Humalyzer
31
32 129 3000 Chemistry Analyzer, Human Diagnostics, Germany), spot urine creatinine, protein, and
33
34 130 albumin tests (Combilyzer dipstick reader, Human Diagnostics, Germany).
35
36
37
38
39
40
41

42 132 A sample size of 1315 participants was calculated for a desired precision range of 10% and an
43
44 133 estimated sensitivity and specificity of the urine glucose test strip of 21% and 90%, respectively,
45
46 134 which is also sufficient for analysis of HbA1c, OGTT, and FBG as the test strip has the lowest
47
48 135 performance. The sample size for the study was calculated based on Buderer's formula [15],
49
50 136 accounting for a 3% drop-out rate and a 5% national prevalence of diabetes [16].
51
52
53

54 137
55
56
57
58
59
60

138 ***Data Analysis***

139 The index tests of interest were a positive self-administered urine glucose test strip, HbA1c
140 >6.5%, and cFBG \geq 126 mg/dL. Diagnostic accuracy was assessed against a composite reference
141 standard, which was cFBG \geq 200 mg/dL, or venous blood glucose, 2 hours after OGTT \geq 200
142 mg/dL.[17,18] If the participant's cFBG was >200 mg/dL, the patient was considered to have
143 diabetes and an OGTT was not performed. Other measures were defined as follows: Overweight
144 (BMI \geq 25 or waist circumference >90cm for men or >80cm for women[19]), elevated blood
145 pressure (systolic pressure \geq 140mmHg or diastolic pressure \geq 90mmHg), albuminuria (\geq 20
146 mg/L), and elevated albumin/creatinine ratio (\geq 30mg/g). We calculated sensitivity, specificity,
147 positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio
148 (LR+), negative LR (LR-),with 95% confidence intervals (CI).

149
150 Subgroup analyses were not prespecified, and therefore used to explore the performance of the
151 urine glucose test strip in participants at increased risk for diabetes mellitus (DM), including age
152 (\geq 50 years), BMI (\geq 25), gender, and waist circumference (>90cm for men or >80cm for
153 women). Logistic regression analyses were also used to determine if the diagnostic accuracy of
154 the index test was impacted by these clinical features. Prevalence of diabetes by subgroup was
155 compared by chi-squared test. We also explored whether the individuals correctly classified by
156 the urine glucose test strip had better or worse controlled diabetes than those misclassified by the
157 test, as defined by various clinical and laboratory measures. Mean values of continuous variables
158 were compared using Student's t-test while proportions of dichotomous values were compared
159 using the chi-squared test. Data were analyzed using Stata/SE 13.1 (StataCorp LP, Texas, USA).

160

161 Results

162 Of 1328 eligible study subjects, 1316 participated in the study and 1289 were included in the
163 analysis (Figure 1). Participants were excluded from the analysis if they did not complete the
164 OGTT due to vomiting or other reasons (16), were not fasting prior to the clinic visit (5), or
165 reported taking medication for diabetes that day (6). Of the analyzed participants, 75%
166 (972/1289) were female, mean age was 51 years, 31% had high BMI, and 13% had elevated
167 blood pressure, although only 8% were taking antihypertensive medications. Characteristics of
168 the participants included in the analysis are presented in Table 1.

169
170 A total of 234 individuals had diabetes based on the composite reference standard of either cFBG
171 (70, 30%) or OGTT (164, 70%), corresponding to a prevalence of 18%. The 70 individuals with
172 cFBG ≥ 200 mg/dL, also all had HbA1c measurements $>6.5\%$. Of the index tests evaluated, the
173 urine glucose test strip had lower sensitivity (14.1%, 95% CI: 9.90-19.2) than cFBG (73.9%,
174 95% CI: 67.8-79.4), and HbA1c (75.2%, 95% CI: 69.2-80.6). All three tests offered high
175 specificity (99.3%, 95% CI: 98.6-99.7; 96.8%, 95% CI: 95.5-97.8; and 98.5%, 95% CI: 97.5-
176 99.1; respectively) (Table 2). The urine glucose test strip failed to identify 201 individuals with
177 diabetes (false negatives) and falsely identified seven participants without diabetes (false
178 positives). The 201 patients with diabetes who were not identified by the urine test had
179 significantly lower venous FBG, lower 2 hr OGTT, and lower HbA1c compared to those
180 correctly diagnosed, but were similar in other characteristics (Table 3). The seven false positive
181 individuals had higher HbA1c, higher systolic BP, and higher proportion receiving treatment for
182 hypertension than those with true negative results (Table 3).

183

184 The prevalence of diabetes (diagnosed by the composite reference standard) was significantly
185 higher in participants who were 50 years of age or older compared to those under 50 years (24%
186 vs. 9.6%, $p<0.001$); those with high BMI compared to those with normal BMI (22% vs. 17%,
187 $p=0.03$); and those with greater waist circumference compared to those with normal waist (24%
188 vs. 13%, $p<0.001$), but was the same in males and females (Table 4). The diagnostic accuracy of
189 the urine glucose test strip was similar among subgroups of patients with various cofactors, with
190 overlapping confidence intervals (Table 4). Additionally, multivariate and univariate logistic
191 regression analyses also indicated that the diagnostic accuracy of the index test was not
192 significantly impacted by these cofactors.

194 Discussion

195 Urine glucose test strips had much lower sensitivity than either cFBG or HbA1c, but all three
196 tests offered high specificity. Patients who tested positive with the urine glucose test who were
197 confirmed to have diabetes by the reference standard (true positives) had higher FBG, higher
198 OGTT and higher HbA1c levels compared to the false negative group (urine test negative in
199 patients with diabetes), suggesting that the urine glucose test may identify individuals with poor
200 glycemic control. This suggests a subset of diabetes patients is being identified that may
201 potentially be at higher risk of advancing complications or comorbidities, and who may benefit
202 the most from further care [20]. In addition, testing for urine glucose was highly specific (99%),
203 with positive LR_s in the 20s, indicating that when positive, this test is highly indicative of
204 diabetes.

205

1
2
3 206 The prevalence of diabetes in the MoPoTsyo population in Cambodia was 18%. This is much
4
5 207 higher than the national prevalence for Cambodia, which is reported at 3.0%.[1] This may be due
6
7 208 to the high proportion of individuals over 50 years of age in our study population, which could
8
9
10 209 be explained by a participation bias towards those who were able to miss a day of work to attend
11
12 210 a clinic visit. Additionally, our study took place in a rapidly changing urban population, which
13
14 211 had a 2.4 times higher diabetes prevalence in the STEP survey, country report from 2010.[21]
15
16
17 212

18
19 213 A wide range of sensitivities for the urine glucose test strip has been reported, and its use
20
21 214 remains controversial. A review in 2000 found six adequately designed studies that reported
22
23 215 performance of urine test strips for glucose.[8] Among these, sensitivities in two reports of
24
25 216 fasting patients were 16% and 35%; two using random samples found sensitivities of 18% and
26
27 217 64%; and three using postprandial and post-load measurements reported sensitivities between
28
29 218 39% and 48%. This review concluded that blood glucose measurements were preferred over
30
31 219 urinary glucose or HbA1c, and particularly, postprandial over fasting measures. Another review
32
33 220 found five studies reporting a range of sensitivity from 18% to 74% for urine glucose test
34
35 221 strips.[7] The review concluded that urine glucose test strips are not sufficient for screening for
36
37 222 diabetes.
38
39
40
41
42 223

43
44 224 This is one of the first studies to determine the prevalence of diabetes in Cambodia, and report on
45
46 225 the screening accuracy of urine glucose test strips which are commonly used as screening tests in
47
48 226 this setting. We used a prospective community-based design and had a large sample size with
49
50 227 high participation rate. The study had several limitations. Firstly, we used a composite reference
51
52 228 test and those with cFBG > 200 mg/dL were not evaluated by the OGTT. When evaluating the
53
54
55
56
57
58
59
60

1
2
3 229 index test of cFBG, the index test is included in the reference test, though at a different threshold,
4
5 230 which can cause incorporation bias resulting in an inflated test accuracy. While OGTT is
6
7
8 231 considered the gold standard reference test for assessing diagnostic accuracy, there has been
9
10 232 some question of its performance. Two studies in China, each on more than 200 participants,
11
12 233 found that the reproducibility of the OGTT was 56% [22] and 66% [23]. Though our choice of
13
14 234 the reference standards, particularly OGTT, could have affected our study results, its use allows
15
16
17 235 comparison of our results to those in a number of other studies. Second, the urine glucose test
18
19 236 was self-administered and self reported. While this was pragmatic, and aligns with the practices
20
21 237 at MoPoTyso and other clinical settings in Cambodia, errors in interpreting the test result could
22
23
24 238 influence accuracy. We were not able to repeat this test when patients attended their clinic visit
25
26 239 as they were fasting at the clinic visit, and thus their urine would not have been the random non-
27
28 240 fasting urine test obtained at home. Third, we were not able to obtain hemoglobin levels (or test
29
30
31 241 for hemoglobin variants) as these tests are not available in this setting, and hence cannot assess
32
33 242 the impact of anemia or hemoglobinopathy on test performance.[24] Fourth, glucose test strip
34
35 243 accuracy may be subject to effects of heat and humidity, we were not able to explore their
36
37
38 244 possible impact on our results.

39
40 245
41
42 246 For clinicians working in settings similar to ours, the question is how useful is the urine glucose
43
44 247 test as a screening or diagnostic test, and is it “better than nothing”? The low sensitivity certainly
45
46
47 248 reduces the value of this test as a screening tool, but the high specificity means that positive tests
48
49 249 can be used to rule in patients with diabetes, suggesting that urine glucose may have some
50
51 250 diagnostic value in this setting. The false positive rate was extremely low, and only 7 patients
52
53
54 251 without disease were identified as positive by urine glucose test strip. From a population

1
2
3 252 perspective, the value of a low cost, poorly sensitive yet highly specific test for diabetes is
4
5 253 unclear in terms of balancing the opportunity to identify a subset of patients with less well
6
7 254 controlled diabetes who would not have been identified otherwise, with the downside of a high
8
9
10 255 false negative rate.[25]
11

12 256

13
14 257 Not surprisingly, usability parameters and cost make urine glucose test strips a highly desirable
15
16 258 test in this and other low-resource settings.[9] Product attributes such as low complexity and
17
18 259 infrastructure requirements, short time to results, and low participant burden greatly contribute to
19
20 260 the acceptability and desirability of the screening tool. The large patient burden and the frequent
21
22 261 inability to comply with fasting requirements reduce the feasibility of using OGTT or FBG tests.
23

24 262 While HbA1c testing does not require fasting, current tests are too expensive for use in most
25
26 263 low-income countries. The role of a poorly sensitive test like urine glucose in resource poor
27
28 264 settings such as Cambodia is debatable, on the one hand the test will identify some patients
29
30 265 previously undiagnosed, and assuming treatment can be initiated, reduce severity of
31
32 266 complications from this disease. On the other hand, the test will miss the majority of patients
33
34 267 with diabetes, thus risking a false reassurance, further postponement of diagnosis, and risking
35
36 268 patient's respect for the health care system.
37

38 269

39
40
41
42 270 There may be strategies to improve the performance (particularly sensitivity) of the urine glucose
43
44 271 test strip. First, using presence of risk factors such as high waist circumference or BMI, may
45
46 272 increase the pretest probability of diabetes and lead to improved performance. In our study, the
47
48 273 sensitivity of the UGTS among overweight men with high waist circumference was twice the
49
50 274 overall sensitivity (29% vs. 14% respectively). Second, using random, postprandial, or glucose-
51
52
53
54
55
56
57
58
59

1
2
3 275 loaded measurements may be superior than fasting because the renal threshold for glucose is
4
5 276 more often reached in non-fasting states.[8] Third, improving the limit of detection may be
6
7 277 possible by modifications in the test strip itself, or improvement in the way it is read either
8
9 278 manually (with trained users) or automatically (with electronic reading devices). Finally,
10
11 279 increasing screening frequency may be feasible in low resource settings, if the urine glucose test
12
13 280 strip truly does identify a smaller but more advanced fraction of diabetes patients.
14
15
16
17
18

19 282 **Conclusion**

20
21 283 Low cost, easy to use diabetes screening, diagnosis, and monitoring tools are essential for low-
22
23 284 resource communities with minimal infrastructure. While the urine glucose test strip has some
24
25 285 value as a screening test in these settings, its performance is far from optimal. Progress is
26
27 286 urgently needed to improve the performance, availability, and access of essential testing
28
29 287 technologies for diabetes.
30
31
32

33 288

34 289

35 290 **List of abbreviations**

36 291 urine glucose test strip (UGTS)

37 292 International Diabetes Federation (IDF)

38 293 non-communicable diseases (NCDs)

39 294 cardiovascular disease (CVD)

40 295 World Health Organisation (WHO)

41 296 capillary fasting blood glucose measurement (cFBG)

42 297 oral glucose tolerance test (OGTT)

43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 298 positive predictive value (PPV)
4

5 299 negative predictive value (NPV)
6

7
8 300 positive likelihood ratio (LR+)
9

10 301 negative likelihood ratio (LR-)
11

12 302 confidence intervals (CI)
13

14 303 diabetes mellitus (DM)
15

16
17 304

18
19 305 **Declarations**
20

21 306 *Ethical approval and consent to participate*
22

23
24 307 The protocol was approved by the PATH Research Ethics Committee and the National Ethics
25
26 308 Committee for Health Research (Cambodia Institutional Review Board). Informed consent was
27
28 309 obtained from all participants.
29

30
31 310 *Consent for publication*
32

33 311 Not applicable.
34

35 312 *Availability of data and material*
36

37 313 The datasets used during the current study are available from the corresponding author on
38
39
40 314 reasonable request.
41

42 315 *Competing Interests*
43

44 316 The authors declare that they have no competing interests.
45
46

47 317 *Funding*
48

49 318 This work was supported by a grant from Medtronic Foundation, and received additional support
50
51 319 from PATH and the University of Washington Department of Family Medicine. The funding
52
53
54
55
56
57
58
59
60

320 source had no involvement in study design, data collection, data analysis, data interpretation,
 321 writing of the manuscript, or the decision to publish the results.

322 *Authors contributions*

323 MHP, SB, TN, HM and BW designed the study; MHP, SB, TN, and BW implemented the study;
 324 HLS, MT, HM, and BW analysed and interpreted the data; HLS, MHP, FD, MT, HM, and BW
 325 contributed to writing. All authors read and approved the final manuscript.

326 *Acknowledgements*

327 We would like to acknowledge the Borey Santhepheap community in Cambodia (Phnom Penh
 328 Municipality, District Dangkao, Commune Chom Chao) for participating in this study. We also
 329 acknowledge the input of Dr Annette Fitzpatrick and Dr Jim LoGerfo from the University of
 330 Washington.

331 *Authors' information*

332 Not applicable.

333

334

335 **Tables**

336

337 **Table 1.** Characteristics of included participants.

	Mean (SD) or % n=1289
Age, years	51.4 (14.9)
Female (%)	75.4
BMI ¹	23.2 (4.1)
High BMI (%)	30.5
Waist circumference above cutoff ² (%)	46.1
Systolic blood pressure, mmHg	123.5 (20.6)
Diastolic blood pressure, mmHg	80.8 (12.1)
Elevated blood pressure (%)	12.9
Take treatment for high blood pressure (%)	8.2

338 ¹ n=1288

339 ² >90cm for men, >80cm for women. [19]

340

341

Table 2. Diagnostic accuracy of urine glucose test strip, capillary fasting glucose, and HbA1c determined by comparison with the composite reference standard (n=1289)¹.

	Urine glucose test strip positive	cFBG \geq 126 mg/dL	HbA1c $>$ 6.5%
True positive (n)	33	173	176
False positive (n)	7	34	16
False negative (n)	201	61	58
True negative (n)	1048	1021	1039
True diabetes prevalence ² (95%CI)	18%, 234/1289 (16, 20.4)		
Sensitivity (95% CI)	14.1 (9.90, 19.2)	73.9 (67.8, 79.4)	75.2 (69.2, 80.6)
Specificity (95% CI)	99.3 (98.6, 99.7)	96.8 (95.5, 97.8)	98.5 (97.5, 99.1)
Positive PV (95% CI)	82.5 (67.2, 92.7)	83.6 (77.8, 88.3)	91.7 (86.8, 95.2)
Negative PV (95% CI)	83.9 (81.7, 85.9)	94.4 (92.8, 95.7)	94.7 (93.2, 96.0)
Positive LR (95% CI)	21.3 (9.50, 47.5)	22.9 (16.3, 32.2)	49.6 (30.3, 81.1)
Negative LR (95% CI)	0.90 (0.80, 0.90)	0.30 (0.20, 0.30)	0.30 (0.20, 0.30)

¹ Excludes individuals taking diabetes treatment that day (n=6), did not fast before OGTT as instructed (n=5), or did not complete the OGTT (n=16).

² Composite reference standard: OGTT \geq 200 mg/dL or cFBG \geq 200 mg/dL. 70 patients with cFBG \geq 200 were not tested by OGTT.

Table 3. Diagnostic accuracy of the urine glucose test strip by patient characteristics.

Patient characteristic: Mean (SD) or %	Diabetic ¹		Non-diabetic ¹	
	True Positive n=33	False Negative n=201	False Positive n=7	True Negative n=1048
Age	57 (9.3)	58 (10.5)	56 (11.9)	50 (15.5)
Female (%)	81.8	74.6	85.7	75.3
Venous fasting blood glucose	207 (75.3)	166 (73.2)	95 (16.9)	90 (13.1)
Venous blood glucose 2 hrs after OGTT	310 (60.8)	275 (62.2)	115 (43.2)	120 (31.0)
Change in venous blood glucose during OGTT	160 (50.8)	146 (49.8)	20 (47.7)	30 (30.0)
HbA1c	10 (2.3)	8 (2.4)	6 (0.7)	5 (0.5)
BMI	24 (3.9)	24 (3.9)	26 (3.2)	23 (4.1)
High BMI (%)	33.3	36.8	57.1	29.0
Waist circumference above cutoff (%)	60.6	61.7	71.4	42.8
Systolic blood pressure	132 (24.9)	130 (20.6)	146 (14.0)	122 (20.2)
Diastolic blood pressure	85 (9.6)	84 (11.7)	87 (6.5)	80 (12.1)
Elevated blood pressure (%)	15.2	20.9	14.3	11.3
Take treatment for high blood pressure (%)	18.2	11.4	28.6	7.1
Total Cholesterol	242 (62.3)	227 (69.8)	240 (63.1)	213 (56.3)
Proteinuria (n=1116) ² (%)	20.0	17.2	0	3.0
Albuminuria (%)	51.5	47.8	14.3	21.7
Abnormal albumin/creatinine ratio (%)	39.3	39.3	14.3	17.3

¹ Diagnosis by the composite reference standard: venous OGTT \geq 200 mg/dL or cFBG \geq 200 mg/dL. 70 patients with cFBG \geq 200 were not tested by OGTT.

² 4 missing values, 169 indeterminate measurements not included in analysis.

359 **Bold = significantly different ($p \leq 0.05$) by Student's t-test or chi-squared test.**

360

361

362 **Table 4.** Diagnostic accuracy of urine glucose test strip by participant cofactors (n=1289) ¹.

363

Results	Cofactors							
	Age		BMI ³		Gender		Waist circumference ⁴	
	<50	≥50	<25	≥25	Male	Female	Normal	High
Number of participants	531	758	895	393	317	972	691	598
True positive (n)	8	25	22	11	6	27	13	20
False positive (n)	3	4	3	4	1	6	2	5
False negative (n)	43	158	127	74	51	150	77	124
True negative (n)	477	571	743	304	259	789	599	449
True diabetes prevalence ²	9.6% (7.2, 12.4)	24% (21.0, 27.4)	17% (14.0, 19.3)	22% (18.0, 26.0)	18% (14.0, 22.7)	18% (16.0, 20.8)	13% (11.0, 15.8)	24% (21.0, 27.7)
Sensitivity (95% CI)	15.7 (7.0, 28.6)	13.7 (9.0, 19.5)	14.8 (9.5, 21.5)	12.9 (6.6, 22.0)	10.5 (4.0, 21.5)	15.3 (10.3, 21.4)	14.4 (7.9, 23.4)	13.9 (8.7, 20.6)
Specificity (95% CI)	99.4 (98.2, 99.9)	99.3 (98.2, 99.8)	99.6 (98.8, 99.9)	98.7 (96.7, 99.6)	99.6 (97.9, 100)	99.2 (98.4, 99.7)	99.7 (98.8, 100)	98.9 (97.4, 99.6)
Positive PV (95% CI)	72.7 (39, 94.0)	86.2 (68.3, 96.1)	88 (68.8, 97.5)	73.3 (44.96, 92.2)	85.7 (42.1, 99.6)	81.8 (64.5, 93.0)	86.7 (59.5, 98.3)	80 (59.3, 93.2)
Negative PV (95% CI)	91.7 (89, 94)	78.3 (75.2, 81.3)	85.4 (82.9, 87.7)	80.4 (76.1, 84.3)	83.5 (78.9, 87.5)	84 (81.5, 86.3)	88.6 (86, 90.9)	78.4 (74.8, 81.7)
Positive LR (95% CI)	25.1 (6.9, 91.6)	19.6 (6.9, 55.7)	36.7 (11.1, 121)	10.0 (3.3, 30.5)	27.4 (3.4, 223)	20.2 (8.5, 48.2)	43.4 (10.0, 189)	12.6 (4.8, 33)
Negative LR (95% CI)	0.8 (0.8, 1.0)	0.9 (0.8, 0.9)	0.9 (0.8, 0.9)	0.9 (0.8, 1.0)	0.9 (0.8, 1.0)	0.85 (0.80, 0.91)	0.86 (0.79, 0.94)	0.87 (0.82, 0.93)

364

365 ¹ Excludes individuals taking diabetes treatment that day (n=6), did not fast before OGTT as instructed (n=5), or did not complete the OGTT (n=16).

366 ² Composite reference standard: OGTT ≥ 200 mg/dL or cFBG ≥ 200 mg/dL. 70 patients with cFBG ≥ 200 were not tested by OGTT.

367 ³ n=1288.

368 ⁴ High Waist circumference = >90 cm for men, >80 cm for women.[19]

369 **Bold = significantly different ($p \leq 0.05$), chi-squared test.**

370

371

372

373 Figure legend

374

375

376 **Figure 1:** Study flow diagram.

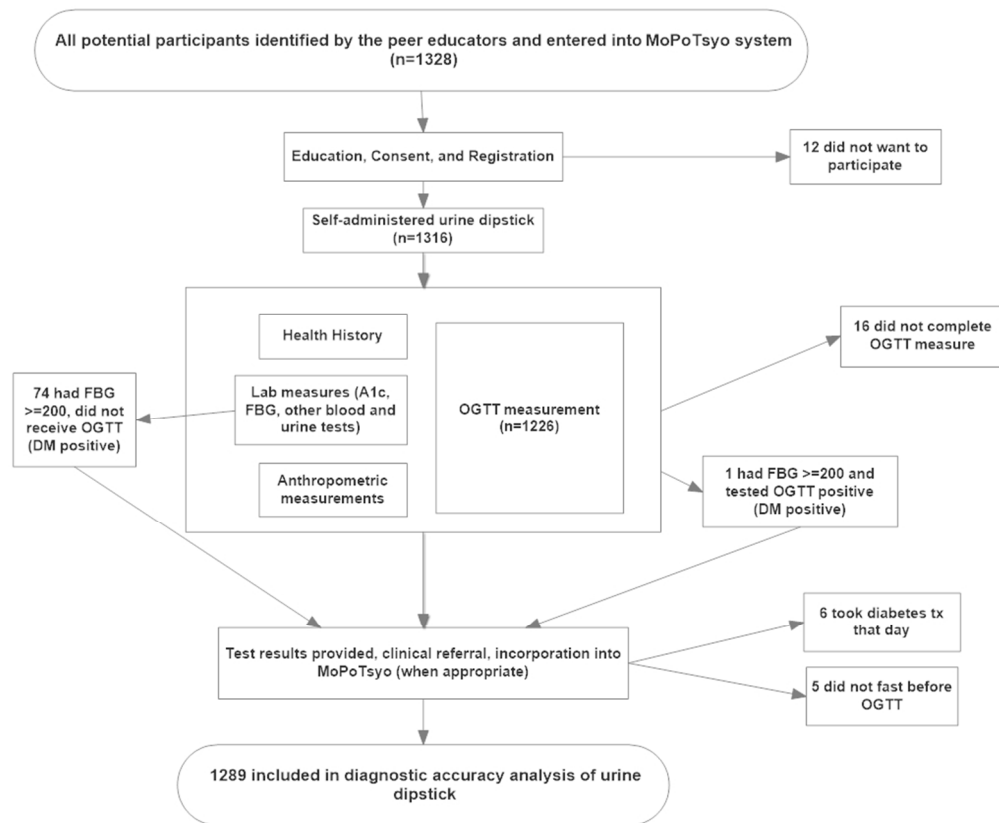
377

378

Reference List

1. IDF Diabetes Atlas, 7th Edition. International Diabetes Federation . 2015. International Diabetes Federation.
Ref Type: Electronic Citation
2. Forbes JM, Cooper ME: **Mechanisms of diabetic complications.** *Physiol Rev* 2013, **93(1)**: 137-188.
3. Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, Ilanne-Parikka P *et al.*: **Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance.** *N Engl J Med* 2001, **344**: 1343-1350.
4. Dyerberg J, Pedersen L, Aagaard O: **Evaluation of a dipstick test for glucose in urine.** *Clin Chem* 1976, **22**: 205-210.
5. van der Sande MA, Walraven GE, Bailey R, Rowley JT, Banya WA, Nyan OA *et al.*: **Is there a role for glycosuria testing in sub-Saharan Africa?** *Trop Med Int Health* 1999, **4**: 506-513.
6. Taal MW, Chertow GM, Marsden PA, Skorecki K, Yu ASL, Brenner BM: **Laboratory Assessment of Kidney Disease: Glomerular Filtration Rate, Urinalysis, and Proteinuria.** In *Brenner & Rector's The Kidney*. Philadelphia: Elsevier; 2012.
7. Wei OY, Teece S: **Best evidence topic report. Urine dipsticks in screening for diabetes mellitus.** *Emerg Med J* 2006, **23**: 138.
8. Engelgau MM, Narayan KM, Herman WH: **Screening for type 2 diabetes.** *Diabetes Care* 2000, **23**: 1563-1580.
9. World Health Organization. Screening for Type 2 Diabetes. Report of a World Health Organization and International Diabetes Federation meeting. 2003. Geneva, Switzerland, World Health Organization.
Ref Type: Report
10. van Pelt M: **Improving access to education and care in Cambodia.** *Diabetes Voice* 2009, **54**.
11. Taniguchi D, LoGerfo J, van PM, Mielcarek B, Huster K, Haider M *et al.*: **Evaluation of a multi-faceted diabetes care program including community-based peer educators in Takeo province, Cambodia, 2007-2013.** *PLoS One* 2017, **12**: e0181582.
12. van Olmen J, Eggermont N, van Pelt M, Hen H, de Man J, Schellevis F *et al.*: **Patient-centred innovation to ensure access to diabetes care in Cambodia: the case of MoPoTsyo.** *J Pharm Policy Pract* 2016, **9**: 1.
13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L *et al.*: **STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies.** *BMJ* 2015, **351**: h5527.

- 1
2
3 417 14. World Health Organization. WHO STEPS Surveillance Manual. World Health Organization . 6-13-
4 418 2008. World Health Organization. 4-30-2015.
5 419 Ref Type: Electronic Citation
6
- 7 420 15. Obuchowski NA: **Sample size calculations in studies of test accuracy.** *Stat Methods Med Res*
8 421 1998, **7**: 371-392.
9
- 10 422 16. King H, Keuky L, Seng S, Khun T, Roglic G, Pinget M: **Diabetes and associated disorders in**
11 423 **Cambodia: two epidemiological surveys.** *Lancet* 2005, **366**: 1633-1639.
12
- 13 424 17. **Standards of Medical Care in Diabetes-2016 Abridged for Primary Care Providers.** *Clin*
14 425 *Diabetes* 2016, **34**: 3-21.
15
- 16 426 18. Bartoli E, Fra GP, Carnevale Schianca GP: **The oral glucose tolerance test (OGTT) revisited.**
17 427 *Eur J Intern Med* 2011, **22**: 8-12.
18
- 19 428 19. World Health Organization. Waist circumference and waist-hip ratio: report of a WHO expert
20 429 consultation, Geneva, 8-11 December 2008. 2011. Geneva, Switzerland, World Health
21 430 Organization.
22 431 Ref Type: Report
23
- 24 432 20. **The effect of intensive treatment of diabetes on the development and progression of long-term**
25 433 **complications in insulin-dependent diabetes mellitus. The Diabetes Control and**
26 434 **Complications Trial Research Group.** *N Engl J Med* 1993, **329**: 977-986.
27
- 28 435 21. University of Health Sciences Cambodia, Ministry of Health Cambodia. Prevalence of non-
29 436 communicable disease risk factors in Cambodia; STEPS Survey Country Report. 2010.
30 437 Ref Type: Report
31
- 32 438 22. Liu M, Pan CY, Jin MM, Su HY, Lu JM: **[The reproducibility and clinical significance of oral**
33 439 **glucose tolerance test for abnormal glucose metabolism].** *Zhonghua Nei Ke Za Zhi* 2007, **46**:
34 440 1007-1010.
35
- 36 441 23. Ko GT, Chan JC, Woo J, Lau E, Yeung VT, Chow CC *et al.*: **The reproducibility and usefulness**
37 442 **of the oral glucose tolerance test in screening for diabetes and other cardiovascular risk**
38 443 **factors.** *Ann Clin Biochem* 1998, **35 (Pt 1)**: 62-67.
39
- 40 444 24. English E, Idris I, Smith G, Dhatariya K, Kilpatrick ES, John WG: **The effect of anaemia and**
41 445 **abnormalities of erythrocyte indices on HbA1c analysis: a systematic review.** *Diabetologia*
42 446 2015, **58**: 1409-1421.
43
- 44 447 25. Flessa S, Zembok A: **Costing of diabetes mellitus type II in Cambodia.** *Health Econ Rev* 2014,
45 448 **4**: 24.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



86x71mm (300 x 300 DPI)

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)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	1
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3
	4	Study objectives and hypotheses	4
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
<i>Participants</i>	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
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	7
	15	How indeterminate index test or reference standard results were handled	7
	16	How missing data on the index test and reference standard were handled	7
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	7
	18	Intended sample size and how it was determined	6
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	17
	20	Baseline demographic and clinical characteristics of participants	15
	21a	Distribution of severity of disease in those with the target condition	15
	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	5
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	15
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	15
	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	10
	27	Implications for practice, including the intended use and clinical role of the index test	11
OTHER INFORMATION			
	28	Registration number and name of registry	NA
	29	Where the full study protocol can be accessed	NA
	30	Sources of funding and other support; role of funders	14

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

Diagnostic accuracy of self-administered urine glucose test strips as a diabetes screening tool in a low-resource setting in Cambodia

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2017-019924.R2
Article Type:	Research
Date Submitted by the Author:	16-Feb-2018
Complete List of Authors:	Storey, Helen; PATH, Diagnostics van Pelt, Maurits; MoPoTsyo Bun, Socheath; MoPoTsyo Daily, Frances; MoPoTsyo Neogi, Tina; PATH Thompson, Matthew; University of Washington, Department of Family Medicine McGuire, Helen; PATH, Noncommunicable Diseases Weigl, Bernhard; PATH
Primary Subject Heading:	Diabetes and endocrinology
Secondary Subject Heading:	Diagnostics, Global health
Keywords:	Low-resource settings, Diabetes, Diagnostics, Urine glucose test strip, Screening

SCHOLARONE™
Manuscripts

1 **Diagnostic accuracy of self-administered urine glucose test strips as** 2 **a diabetes screening tool in a low-resource setting in Cambodia**

3
4 Helen L. Storey; PATH, Seattle WA 98109, USA; hstorey@path.org

5 Maurits H. van Pelt; MoPoTsyo, Phnom Penh, Cambodia; mopotsyo@gmail.com

6 Socheath Bun; MoPoTsyo, Phnom Penh, Cambodia; mopo.medtrainer@gmail.com

7 Frances Daily; MoPoTsyo, Phnom Penh, Cambodia; fdaily@bigpond.com

8 Tina Neogi¹; PATH, Seattle WA 98109, USA; tinaneogi@gmail.com

9 Matthew Thompson; Department of Family Medicine, University of Washington, Seattle WA
10 98195, USA; mjt@uw.edu

11 Helen McGuire; PATH, Washington DC 20001, USA; hmcguire@path.org

12 Bernhard H Weigl²; PATH, Seattle WA 98109, USA; bweigl@intven.com

13
14 * Corresponding author: Helen McGuire, PATH, PO Box 900922 Seattle WA 98109,
15 hmcguire@path.org, 202-540-4360

16 ¹ Present Address: UW Neighborhood Ravenna Clinic, Seattle WA 98105

17 ² Present Address: Intellectual Ventures/Global Good, Bellevue WA 98005, USA

18
19
20 **Abstract** (word count: 287)

21 **Objective:** Screening for diabetes in low resource countries is a growing challenge, necessitating
22 tests that are resource and context appropriate. The aim of this study was to determine the

23 diagnostic accuracy of a self-administered urine glucose test strip compared to alternative
24 diabetes screening tools in a low resource setting of Cambodia.

25 **Design:** Prospective cross-sectional study

26 **Setting:** Members of the Borey Santhepheap community in Cambodia (Phnom Penh
27 Municipality, District Dangkao, Commune Chom Chao).

28 **Participants:** All households on randomly selected streets were invited to participate, and adults
29 at least 18 years of age living in the study area were eligible for inclusion.

30 **Outcomes:** The accuracy of self-administered urine glucose test strip positivity, HbA1c >6.5%,
31 and capillary fasting blood glucose measurement ≥ 126 mg/dL were assessed against a composite
32 reference standard of capillary fasting blood glucose measurement ≥ 200 mg/dL or venous blood
33 glucose 2 hours after oral glucose tolerance test ≥ 200 mg/dL.

34 **Results:** Of the 1289 participants, 234 (18%) had diabetes based on either capillary fasting blood
35 glucose measurement (74, 32%) or the oral glucose tolerance test (160, 68%). The urine glucose
36 test strip was 14% sensitive and 99% specific, and failed to identify 201 individuals with
37 diabetes, while falsely identifying 7 without diabetes. Those missed by the urine glucose test
38 strip had lower venous fasting blood glucose, lower venous blood glucose 2 hours after oral
39 glucose tolerance test, and lower HbA1c compared with those correctly diagnosed.

40 **Conclusions:** Low cost, easy to use diabetes tools are essential for low-resource communities
41 with minimal infrastructure. While the urine glucose test strip may identify persons with diabetes
42 that might otherwise go undiagnosed in these settings, its poor sensitivity cannot be ignored. The
43 massive burden of diabetes in low-resource settings demands improvements in test technologies.

44 **Keywords:** Diabetes, Low-resource settings, Diagnostics, Urine glucose test strip, Screening,
45

46 Article Summary (word count: 2261)

47 *Strengths and limitations of the study*

- 48 • This is one of the first studies to determine the prevalence of diabetes and report on the
49 screening accuracy of urine glucose test strips in Cambodia, which are commonly used as
50 screening tests in this setting.
- 51 • We used a prospective community-based design and had a large sample size with high
52 participation rate, though participation bias towards those able to miss a day of work to
53 attend a clinic visit may still have been an issue.
- 54 • Use of a composite reference test and not evaluating those with capillary fasting blood
55 glucose > 200 mg/dL by the oral glucose tolerance test, could have affected our study
56 results, though the use of oral glucose tolerance test allows comparison of our results to
57 those in a number of other studies.
- 58 • The urine glucose test was self-administered and self reported, which is pragmatic and
59 aligns with the practices at MoPoTyso and other clinical settings in Cambodia, however
60 errors in interpreting the test result could influence accuracy.

62 **Background**

63 According to the International Diabetes Federation, 415 million adults are living with diabetes
64 globally, almost half of which are undiagnosed, and this number is expected to increase to 642
65 million by 2040.[1] As is the case for most non-communicable diseases, three quarters of those
66 affected live in low- and middle-income countries. In Cambodia for example, there are an
67 estimated 230,000 people with diabetes, who are at risk for the associated micro- and
68 macrovascular complications of this disease, including cardiovascular disease.[1,2] Strategies to

1
2
3 69 reduce cardiovascular disease risk may also prevent and control diabetes, which would further
4
5 70 reduce rates of eye, kidney, and neural damage due to diabetes complications.[3] To facilitate
6
7 71 screening and monitoring for diabetes in these low- and middle-income countries, a low-cost,
8
9
10 72 point-of-care diagnostic test that is resource and context appropriate is needed.
11
12
13 73

14 74 In low-resource settings, urine glucose test strips have been used as diabetes screening tools
15
16 75 because they are inexpensive, noninvasive, and easy to use.[4,5] While these tests do not require
17
18 76 fasting and are user friendly, they can only detect glucose after it has exceeded the threshold for
19
20 77 reabsorption by the kidneys and appears in the urine. The reported threshold varies and is
21
22 78 affected by kidney function.[6] Although their low sensitivity makes them inadequate for use as
23
24 79 a screening tool,[7-9] the World Health Organisation acknowledges that they may have a place
25
26 80 in low resource settings where other tests are not possible and the prevalence of undiagnosed
27
28 81 diabetes may be high.[9] Currently many people are not diagnosed until severe complications
29
30 82 develop. Although the sensitivity of the urine test delays diagnosis relative to other methods, it
31
32 83 may provide an opportunity to reduce further advancement of complications.
33
34
35
36
37
38
39

40 84
41 85 MoPoTsyo, a nongovernmental organization, provides screening and care services to people with
42
43 86 diabetes and hypertension in Cambodia through an innovative, community-based peer educator
44
45 87 model.[10-12] MoPoTsyo uses urine glucose test strips issued in the community and self-
46
47 88 administered by patients as the initial method of diabetes screening, which has allowed them to
48
49 89 screen over 700,000 adults, followed by confirmation with blood glucose testing for those who
50
51 90 have a positive urine test. The aim of this study was to determine the diagnostic accuracy of a
52
53 91 self-administered urine glucose test strip compared to alternative diabetes screening tools in a
54
55
56
57
58
59

1
2
3 92 low resource setting of Cambodia. We also explored whether individuals with diabetes who were
4
5 93 detected by urine glucose test strips differed in health status compared to those who were missed
6
7
8 94 by this test but detected by blood glucose measurement. Greater understanding of the
9
10 95 performance of this test by the MoPoTsyo program will help to inform its optimal use.
11
12 96

14 97 **Methods**

16 98 *Study design and procedures*

17
18 99 A prospective cross-sectional study was performed among members of the Borey Santhepheap
19
20
21 100 community in Cambodia (Phnom Penh Municipality, District Dangkao, Commune Chom Chao)
22
23 101 from November 2013 to October 2014. All households on randomly selected streets were invited
24
25
26 102 to participate by a local peer educator, who described the study to all potential household
27
28 103 members. Adults at least 18 years of age living in the study area were eligible for inclusion.
29
30 104 Individuals were excluded if they had diabetes or hypertension or had taken medications for
31
32 105 diabetes and/or high blood pressure in the last 30 days, had kidney disease, or had received
33
34 106 dialysis. Written informed consent was obtained from all participants. The protocol was approved
35
36 107 by the PATH Research Ethics Committee and the National Ethics Committee for Health
37
38 108 Research (Cambodia Institutional Review Board). Study methods and results are reported in
39
40 109 alignment with the 2015 STARD recommendations.[13]
41
42
43
44
45 110

46
47 111 After enrollment, all participants were screened for diabetes using a self-administered and self-
48
49 112 reported urine glucose test strip (Sichuan Medicines and Health Products, Chengdu, China).
50
51 113 Participants were taught how to use the test strip and read the results with assistance of a color
52
53 114 chart, and were given several ways to report results to their peer educator. All participants were
54
55
56
57
58
59
60

1
2
3 115 then invited to attend the clinic following an 8-hour fast for laboratory confirmed tests for
4
5 116 diabetes and associated co-morbid risk factors. Upon arriving at the clinic all participants
6
7 117 provided a urine sample, a venous blood sample, and a finger stick blood sample for capillary
8
9 118 fasting blood glucose measurement (cFBG) (On Call Plus glucometer, Acon Laboratories, San
10
11 119 Diego, USA, <https://www.aconlabs.com/us/glucose/on-call/plus-bgms/>). If the cFBG was less
12
13 120 than 200 mg/dL they were asked to consume a 75g oral glucose load for the oral glucose
14
15 121 tolerance test (OGTT). The oral glucose load was ingested within 5 minutes of starting
16
17 122 consumption, and two hours after ingestion, further venous blood and finger stick blood samples
18
19 123 were obtained for glucose measurements. During the visit, a health history was completed based
20
21 124 on the WHO STEPS surveillance questionnaire [14] and blood pressure measured by trained
22
23 125 clinical staff using an electronic device (Omron Corporation, Tokyo, Japan). All devices used in
24
25 126 the study were owned and used previously by MoPoTsyo within the guidelines of the Cambodian
26
27 127 Ministry of Health; none of the devices were investigational. Additional laboratory tests
28
29 128 performed included HbA1c (DCA Vantage Analyzer, Siemens AG, Germany), serum creatinine,
30
31 129 glucose, total cholesterol, high-density lipoprotein cholesterol, and triglycerides (Humalyzer
32
33 130 3000 Chemistry Analyzer, Human Diagnostics, Germany), spot urine creatinine, protein, and
34
35 131 albumin tests (Combilyzer dipstick reader, Human Diagnostics, Germany).
36
37
38
39
40
41
42
43

44 133 A sample size of 1315 participants was calculated for a desired precision range of 10% and an
45
46 134 estimated sensitivity and specificity of the urine glucose test strip of 21% and 90%, respectively,
47
48 135 which is also sufficient for analysis of HbA1c, OGTT, and FBG as the test strip has the lowest
49
50 136 performance. The sample size for the study was calculated based on Buderer's formula [15],
51
52 137 accounting for a 3% drop-out rate and a 5% national prevalence of diabetes [16].
53
54
55
56
57
58
59
60

138

Data Analysis

The index tests of interest were a positive self-administered urine glucose test strip, HbA1c >6.5%, and cFBG \geq 126 mg/dL. Diagnostic accuracy was assessed against a composite reference standard, which was cFBG \geq 200 mg/dL, or venous blood glucose, 2 hours after OGTT \geq 200 mg/dL.[17,18] If the participant's cFBG was >200 mg/dL, the patient was considered to have diabetes and an OGTT was not performed. Other measures were defined as follows: Overweight (BMI \geq 25 or waist circumference >90cm for men or >80cm for women[19]), elevated blood pressure (systolic pressure \geq 140mmHg or diastolic pressure \geq 90mmHg), albuminuria (\geq 20 mg/L), and elevated albumin/creatinine ratio (\geq 30mg/g). We calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), negative LR (LR-),with 95% confidence intervals (CI).

150

Subgroup analyses were not prespecified, and therefore used to explore the performance of the urine glucose test strip in participants at increased risk for diabetes mellitus (DM), including age (\geq 50 years), BMI (\geq 25), gender, and waist circumference (>90cm for men or >80cm for women). Logistic regression analyses were also used to determine if the diagnostic accuracy of the index test was impacted by these clinical features. Prevalence of diabetes by subgroup was compared by chi-squared test. We also explored whether the individuals correctly classified by the urine glucose test strip had better or worse controlled diabetes than those misclassified by the test, as defined by various clinical and laboratory measures. Mean values of continuous variables were compared using Student's t-test while proportions of dichotomous values were compared using the chi-squared test. Data were analyzed using Stata/SE 13.1 (StataCorp LP, Texas, USA).

161

162 Results

163 Of 1328 eligible study subjects, 1316 participated in the study and 1289 were included in the
164 analysis (Figure 1). Participants were excluded from the analysis if they did not complete the
165 OGTT due to vomiting or other reasons (16), were not fasting prior to the clinic visit (5), or
166 reported taking medication for diabetes that day (6). Of the analyzed participants, 75%
167 (972/1289) were female, mean age was 51 years, 31% had high BMI, and 13% had elevated
168 blood pressure, although only 8% were taking antihypertensive medications. Characteristics of
169 the participants included in the analysis are presented in Table 1.

170

171 A total of 234 individuals had diabetes based on the composite reference standard of either cFBG
172 (70, 30%) or OGTT (164, 70%), corresponding to a prevalence of 18%. The 70 individuals with
173 cFBG ≥ 200 mg/dL, also all had HbA1c measurements $> 6.5\%$. Of the index tests evaluated, the
174 urine glucose test strip had lower sensitivity (14.1%, 95% CI: 9.90-19.2) than cFBG (73.9%,
175 95% CI: 67.8-79.4), and HbA1c (75.2%, 95% CI: 69.2-80.6). All three tests offered high
176 specificity (99.3%, 95% CI: 98.6-99.7; 96.8%, 95% CI: 95.5-97.8; and 98.5%, 95% CI: 97.5-
177 99.1; respectively) (Table 2). The urine glucose test strip failed to identify 201 individuals with
178 diabetes (false negatives) and falsely identified seven participants without diabetes (false
179 positives). The 201 patients with diabetes who were not identified by the urine test had
180 significantly lower venous FBG, lower 2 hr OGTT, and lower HbA1c compared to those
181 correctly diagnosed, but were similar in other characteristics (Table 3). The seven false positive
182 individuals had higher HbA1c, higher systolic BP, and higher proportion receiving treatment for
183 hypertension than those with true negative results (Table 3).

184
185 The prevalence of diabetes (diagnosed by the composite reference standard) was significantly
186 higher in participants who were 50 years of age or older compared to those under 50 years (24%
187 vs. 9.6%); those with high BMI compared to those with normal BMI (22% vs. 17%); and those
188 with greater waist circumference compared to those with normal waist (24% vs. 13%), but was
189 the same in males and females (Table 4). The diagnostic accuracy of the urine glucose test strip
190 was similar among subgroups of patients with various cofactors, with overlapping confidence
191 intervals (Table 4). Additionally, multivariate and univariate logistic regression analyses also
192 indicated that the diagnostic accuracy of the index test was not significantly impacted by these
193 cofactors.

194

195 Discussion

196 Urine glucose test strips had much lower sensitivity than either cFBG or HbA1c, but all three
197 tests offered high specificity. Patients who tested positive with the urine glucose test who were
198 confirmed to have diabetes by the reference standard (true positives) had higher FBG, higher
199 OGTT and higher HbA1c levels compared to the false negative group (urine test negative in
200 patients with diabetes), suggesting that the urine glucose test may identify individuals with poor
201 glycemic control. This suggests a subset of diabetes patients is being identified that may
202 potentially be at higher risk of advancing complications or comorbidities, and who may benefit
203 the most from further care [20]. In addition, testing for urine glucose was highly specific (99%),
204 with positive LRs in the 20s, indicating that when positive, this test is highly indicative of
205 diabetes.

206

207 The prevalence of diabetes in the MoPoTsyo population in Cambodia was 18%. This is much
208 higher than the national prevalence for Cambodia, which is reported at 3.0%.[1] This may be due
209 to the high proportion of individuals over 50 years of age in our study population, which could
210 be explained by a participation bias towards those who were able to miss a day of work to attend
211 a clinic visit. Additionally, our study took place in a rapidly changing urban population, which
212 had a 2.4 times higher diabetes prevalence in the STEP survey, country report from 2010.[21]

214 A wide range of sensitivities for the urine glucose test strip has been reported, and its use
215 remains controversial. A review in 2000 found six adequately designed studies that reported
216 performance of urine test strips for glucose.[8] Among these, sensitivities in two reports of
217 fasting patients were 16% and 35%; two using random samples found sensitivities of 18% and
218 64%; and three using postprandial and post-load measurements reported sensitivities between
219 39% and 48%. This review concluded that blood glucose measurements were preferred over
220 urinary glucose or HbA1c, and particularly, postprandial over fasting measures. Another review
221 found five studies reporting a range of sensitivity from 18% to 74% for urine glucose test
222 strips.[7] The review concluded that urine glucose test strips are not sufficient for screening for
223 diabetes.

225 This is one of the first studies to determine the prevalence of diabetes in Cambodia, and report on
226 the screening accuracy of urine glucose test strips which are commonly used as screening tests in
227 this setting. We used a prospective community-based design and had a large sample size with
228 high participation rate. The study had several limitations. Firstly, we used a composite reference
229 test and those with cFBG > 200 mg/dL were not evaluated by the OGTT. When evaluating the

1
2
3 230 index test of cFBG, the index test is included in the reference test, though at a different threshold.
4
5 231 This can cause incorporation bias resulting in an inflated test accuracy. Here the three different
6
7 232 index tests are included for comparison; however, the likely overestimation of diagnostic
8
9 233 accuracy for cFBG is important to keep in mind. While OGTT is considered the gold standard
10
11 234 reference test for assessing diagnostic accuracy, there has been some question of its performance.
12
13 235 Two studies in China, each on more than 200 participants, found that the reproducibility of the
14
15 236 OGTT was 56% [22] and 66% [23]. Though our choice of the reference standards, particularly
16
17 237 OGTT, could have affected our study results, its use allows comparison of our results to those in
18
19 238 a number of other studies. Second, the urine glucose test was self-administered and self reported.
20
21 239 While this was pragmatic, and aligns with the practices at MoPoTyso and other clinical settings
22
23 240 in Cambodia, errors in interpreting the test result could influence accuracy. We were not able to
24
25 241 repeat this test when patients attended their clinic visit as they were fasting at the clinic visit, and
26
27 242 thus their urine would not have been the random non-fasting urine test obtained at home. Third,
28
29 243 we were not able to obtain hemoglobin levels (or test for hemoglobin variants) as these tests are
30
31 244 not available in this setting, and hence cannot assess the impact of anemia or hemoglobinopathy
32
33 245 on test performance.[24] Fourth, glucose test strip accuracy may be subject to effects of heat and
34
35 246 humidity, we were not able to explore their possible impact on our results.
36
37
38 247
39
40
41
42
43
44 248 For clinicians working in settings similar to ours, the question is how useful is the urine glucose
45
46 249 test as a screening or diagnostic test, and is it “better than nothing”? The low sensitivity certainly
47
48 250 reduces the value of this test as a screening tool, but the high specificity means that positive tests
49
50 251 can be used to rule in patients with diabetes, suggesting that urine glucose may have some
51
52 252 diagnostic value in this setting. The false positive rate was extremely low, and only 7 patients
53
54
55
56
57
58
59
60

1
2
3 253 without disease were identified as positive by urine glucose test strip. From a population
4
5 254 perspective, the value of a low cost, poorly sensitive yet highly specific test for diabetes is
6
7
8 255 unclear in terms of balancing the opportunity to identify a subset of patients with less well
9
10 256 controlled diabetes who would not have been identified otherwise, with the downside of a high
11
12 257 false negative rate.[25]
13

14 258
15
16
17 259 Not surprisingly, usability parameters and cost make urine glucose test strips a highly desirable
18
19 260 test in this and other low-resource settings.[9] Product attributes such as low complexity and
20
21 261 infrastructure requirements, short time to results, and low participant burden greatly contribute to
22
23 262 the acceptability and desirability of the screening tool. The large patient burden and the frequent
24
25 263 inability to comply with fasting requirements reduce the feasibility of using OGTT or FBG tests.
26
27 264 While HbA1c testing does not require fasting, current tests are too expensive for use in most
28
29 265 low-income countries. The role of a poorly sensitive test like urine glucose in resource poor
30
31 266 settings such as Cambodia is debatable, on the one hand the test will identify some patients
32
33 267 previously undiagnosed, and assuming treatment can be initiated, reduce severity of
34
35 268 complications from this disease. On the other hand, the test will miss the majority of patients
36
37 269 with diabetes, thus risking a false reassurance, further postponement of diagnosis, and risking
38
39 270 patient's respect for the health care system.
40
41
42
43

44 271
45
46
47 272 There may be strategies to improve the performance (particularly sensitivity) of the urine glucose
48
49 273 test strip. First, using presence of risk factors such as high waist circumference or BMI, may
50
51 274 increase the pretest probability of diabetes and lead to improved performance. In our study, the
52
53 275 sensitivity of the urine glucose test strip among overweight men with high waist circumference
54
55
56
57
58
59

1
2
3 276 was twice the overall sensitivity (29% vs. 14% respectively). Second, using random,
4
5 277 postprandial, or glucose-loaded measurements may be superior than fasting because the renal
6
7 278 threshold for glucose is more often reached in non-fasting states.[8] Third, improving the limit of
8
9 279 detection may be possible by modifications in the test strip itself, or improvement in the way it is
10
11 280 read either manually (with trained users) or automatically (with electronic reading devices).
12
13
14 281 Finally, increasing screening frequency may be feasible in low resource settings, if the urine
15
16 282 glucose test strip truly does identify a smaller but more advanced fraction of diabetes patients.
17
18
19 283

21 284 **Conclusion**

22
23 285 Low cost, easy to use diabetes screening, diagnosis, and monitoring tools are essential for low-
24
25 286 resource communities with minimal infrastructure. While the urine glucose test strip has some
26
27 287 value as a screening test in these settings, its performance is far from optimal. Progress is
28
29 288 urgently needed to improve the performance, availability, and access of essential testing
30
31 289 technologies for diabetes.
32
33
34 290

35 291

37 292 **List of abbreviations**

38 293 urine glucose test strip (UGTS)
39
40 294 capillary fasting blood glucose measurement (cFBG)
41
42 295 oral glucose tolerance test (OGTT)
43
44 296 positive predictive value (PPV)
45
46 297 negative predictive value (NPV)
47
48 298 positive likelihood ratio (LR+)
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 299 negative likelihood ratio (LR-)

4
5 300 confidence intervals (CI)

6
7
8 301 diabetes mellitus (DM)

9
10 302

11
12 303 **Declarations**

13
14 304 *Ethical approval and consent to participate*

15
16
17 305 The protocol was approved by the PATH Research Ethics Committee and the National Ethics
18
19 306 Committee for Health Research (Cambodia Institutional Review Board). Informed consent was
20
21 307 obtained from all participants.

22
23
24 308 *Consent for publication*

25
26 309 Not applicable.

27
28 310 *Availability of data and material*

29
30 311 The datasets used during the current study are available from the corresponding author on
31
32 312 reasonable request.

33
34 313 *Competing Interests*

35
36 314 The authors declare that they have no competing interests.

37
38 315 *Funding*

39
40 316 This work was supported by a grant from Medtronic Foundation, and received additional support
41
42 317 from PATH and the University of Washington Department of Family Medicine. The funding
43
44 318 source had no involvement in study design, data collection, data analysis, data interpretation,
45
46 319 writing of the manuscript, or the decision to publish the results.

47
48
49 320 *Authors contributions*

50
51
52
53
54
55
56
57
58
59
60

321 MHP, SB, TN, HM and BW designed the study; MHP, SB, TN, and BW implemented the study;
 322 HLS, MT, HM, and BW analysed and interpreted the data; HLS, MHP, FD, MT, HM, and BW
 323 contributed to writing. All authors read and approved the final manuscript.

324 *Acknowledgements*

325 We would like to acknowledge the Borey Santhepheap community in Cambodia (Phnom Penh
 326 Municipality, District Dangkao, Commune Chom Chao) for participating in this study. We also
 327 acknowledge the input of Dr Annette Fitzpatrick and Dr Jim LoGerfo from the University of
 328 Washington.

329 *Authors' information*

330 Not applicable.

331

332

333 **Tables**

334

335 **Table 1.** Characteristics of included participants.

	Mean (SD) or % n=1289
Age, years	51.4 (14.9)
Female (%)	75.4
BMI ¹	23.2 (4.1)
High BMI (%)	30.5
Waist circumference above cutoff ² (%)	46.1
Systolic blood pressure, mmHg	123.5 (20.6)
Diastolic blood pressure, mmHg	80.8 (12.1)
Elevated blood pressure (%)	12.9
Take treatment for high blood pressure (%)	8.2

336 ¹n=1288

337 ²>90cm for men, >80cm for women. [19]

338

339

340 **Table 2.** Diagnostic accuracy of urine glucose test strip, capillary fasting glucose, and HbA1c determined
 341 by comparison with the composite reference standard (n=1289)¹.

342

	Urine glucose test strip positive	cFBG ≥126 mg/dL	HbA1c >6.5%
True positive (n)	33	173	176
False positive (n)	7	34	16
False negative (n)	201	61	58

True negative (n)	1048	1021	1039
True diabetes prevalence ² (95%CI)	18%, 234/1289 (16, 20.4)		
Sensitivity (95% CI)	14.1 (9.90, 19.2)	73.9 (67.8, 79.4)	75.2 (69.2, 80.6)
Specificity (95% CI)	99.3 (98.6, 99.7)	96.8 (95.5, 97.8)	98.5 (97.5, 99.1)
Positive PV (95% CI)	82.5 (67.2, 92.7)	83.6 (77.8, 88.3)	91.7 (86.8, 95.2)
Negative PV (95% CI)	83.9 (81.7, 85.9)	94.4 (92.8, 95.7)	94.7 (93.2, 96.0)
Positive LR (95% CI)	21.3 (9.50, 47.5)	22.9 (16.3, 32.2)	49.6 (30.3, 81.1)
Negative LR (95% CI)	0.90 (0.80, 0.90)	0.30 (0.20, 0.30)	0.30 (0.20, 0.30)

¹ Excludes individuals taking diabetes treatment that day (n=6), did not fast before OGTT as instructed (n=5), or did not complete the OGTT (n=16).

² Composite reference standard: OGTT \geq 200 mg/dL or cFBG \geq 200 mg/dL. 70 patients with cFBG \geq 200 were not tested by OGTT.

Table 3. Diagnostic accuracy of the urine glucose test strip by patient characteristics.

Patient characteristic: Mean (SD) or %	Diabetic ¹		Non-diabetic ¹	
	True Positive n=33	False Negative n=201	False Positive n=7	True Negative n=1048
Age	57 (9.3)	58 (10.5)	56 (11.9)	50 (15.5)
Female (%)	81.8	74.6	85.7	75.3
Venous fasting blood glucose	207 (75.3)	166 (73.2)	95 (16.9)	90 (13.1)
Venous blood glucose 2 hrs after OGTT	310 (60.8)	275 (62.2)	115 (43.2)	120 (31.0)
Change in venous blood glucose during OGTT	160 (50.8)	146 (49.8)	20 (47.7)	30 (30.0)
HbA1c	10 (2.3)	8 (2.4)	6 (0.7)	5 (0.5)
BMI	24 (3.9)	24 (3.9)	26 (3.2)	23 (4.1)
High BMI (%)	33.3	36.8	57.1	29.0
Waist circumference above cutoff (%)	60.6	61.7	71.4	42.8
Systolic blood pressure	132 (24.9)	130 (20.6)	146 (14.0)	122 (20.2)
Diastolic blood pressure	85 (9.6)	84 (11.7)	87 (6.5)	80 (12.1)
Elevated blood pressure (%)	15.2	20.9	14.3	11.3
Take treatment for high blood pressure (%)	18.2	11.4	28.6	7.1
Total Cholesterol	242 (62.3)	227 (69.8)	240 (63.1)	213 (56.3)
Proteinuria (n=1116) ² (%)	20.0	17.2	0	3.0
Albuminuria (%)	51.5	47.8	14.3	21.7
Abnormal albumin/creatinine ratio (%)	39.3	39.3	14.3	17.3

¹ Diagnosis by the composite reference standard: venous OGTT \geq 200 mg/dL or cFBG \geq 200 mg/dL. 70 patients with cFBG \geq 200 were not tested by OGTT.

² 4 missing values, 169 indeterminate measurements not included in analysis.

Bold = significantly different ($p \leq 0.05$) by Student's t-test or chi-squared test.

Table 4. Diagnostic accuracy of urine glucose test strip by participant cofactors (n=1289)¹.

Results	Cofactors
---------	-----------

	Age		BMI ³		Gender		Waist circumference ⁴	
	<50	≥50	<25	≥25	Male	Female	Normal	High
Number of participants	531	758	895	393	317	972	691	598
True positive (n)	8	25	22	11	6	27	13	20
False positive (n)	3	4	3	4	1	6	2	5
False negative (n)	43	158	127	74	51	150	77	124
True negative (n)	477	571	743	304	259	789	599	449
True diabetes prevalence ²	9.6% (7.2, 12.4)	24% (21.0, 27.4)	17% (14.0, 19.3)	22% (18.0, 26.0)	18% (14.0, 22.7)	18% (16.0, 20.8)	13% (11.0, 15.8)	24% (21.0, 27.7)
Sensitivity (95% CI)	15.7 (7.0, 28.6)	13.7 (9.0, 19.5)	14.8 (9.5, 21.5)	12.9 (6.6, 22.0)	10.5 (4.0, 21.5)	15.3 (10.3, 21.4)	14.4 (7.9, 23.4)	13.9 (8.7, 20.6)
Specificity (95% CI)	99.4 (98.2, 99.9)	99.3 (98.2, 99.8)	99.6 (98.8, 99.9)	98.7 (96.7, 99.6)	99.6 (97.9, 100)	99.2 (98.4, 99.7)	99.7 (98.8, 100)	98.9 (97.4, 99.6)
Positive PV (95% CI)	72.7 (39, 94.0)	86.2 (68.3, 96.1)	88 (68.8, 97.5)	73.3 (44.96, 92.2)	85.7 (42.1, 99.6)	81.8 (64.5, 93.0)	86.7 (59.5, 98.3)	80 (59.3, 93.2)
Negative PV (95% CI)	91.7 (89, 94)	78.3 (75.2, 81.3)	85.4 (82.9, 87.7)	80.4 (76.1, 84.3)	83.5 (78.9, 87.5)	84 (81.5, 86.3)	88.6 (86, 90.9)	78.4 (74.8, 81.7)
Positive LR (95% CI)	25.1 (6.9, 91.6)	19.6 (6.9, 55.7)	36.7 (11.1, 121)	10.0 (3.3, 30.5)	27.4 (3.4, 223)	20.2 (8.5, 48.2)	43.4 (10.0, 189)	12.6 (4.8, 33)
Negative LR (95% CI)	0.8 (0.8, 1.0)	0.9 (0.8, 0.9)	0.9 (0.8, 0.9)	0.9 (0.8, 1.0)	0.9 (0.8, 1.0)	0.85 (0.80, 0.91)	0.86 (0.79, 0.94)	0.87 (0.82, 0.93)

¹ Excludes individuals taking diabetes treatment that day (n=6), did not fast before OGTT as instructed (n=5), or did not complete the OGTT (n=16).

² Composite reference standard: OGTT ≥200 mg/dL or cFBG ≥200 mg/dL. 70 patients with cFBG ≥200 were not tested by OGTT.

³ n=1288.

⁴ High Waist circumference = >90cm for men, >80cm for women.[19]

Bold = significantly different ($p \leq 0.05$), chi-squared test.

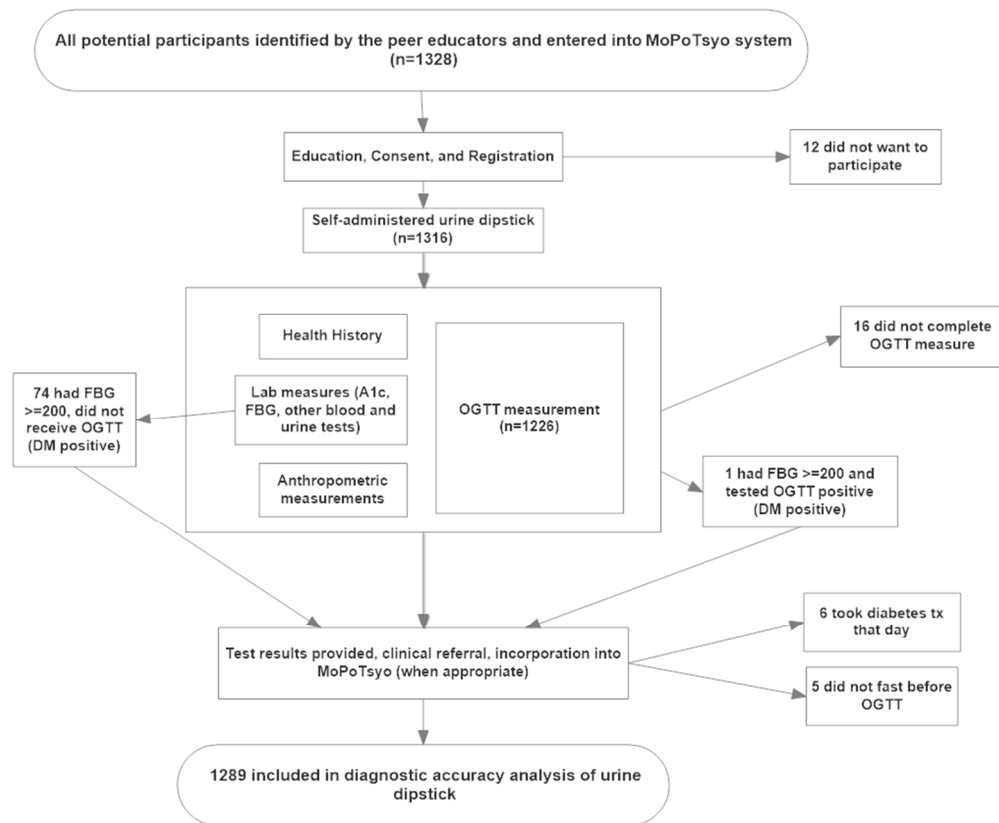
Figure legend

Figure 1: Study flow diagram.

Reference List

- 1
2
3 382 1. IDF Diabetes Atlas, 7th Edition. International Diabetes Federation . 2015. International Diabetes
4 383 Federation.
5 384 Ref Type: Electronic Citation
6
7 385 2. Forbes JM, Cooper ME: **Mechanisms of diabetic complications.** *Physiol Rev* 2013, **93(1)**: 137-
8 386 188.
9
10 387 3. Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, Ilanne-Parikka P *et al.*:
11 388 **Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired**
12 389 **glucose tolerance.** *N Engl J Med* 2001, **344**: 1343-1350.
13
14 390 4. Dyerberg J, Pedersen L, Aagaard O: **Evaluation of a dipstick test for glucose in urine.** *Clin Chem*
15 391 1976, **22**: 205-210.
16
17 392 5. van der Sande MA, Walraven GE, Bailey R, Rowley JT, Banya WA, Nyan OA *et al.*: **Is there a**
18 393 **role for glycosuria testing in sub-Saharan Africa?** *Trop Med Int Health* 1999, **4**: 506-513.
19
20 394 6. Taal MW, Chertow GM, Marsden PA, Skorecki K, Yu ASL, Brenner BM: **Laboratory**
21 395 **Assessment of Kidney Disease: Glomerular Filtration Rate, Urinalysis, and Proteinuria.** In
22 396 *Brenner & Rector's The Kidney.* Philadelphia: Elsevier; 2012.
23
24 397 7. Wei OY, Teece S: **Best evidence topic report. Urine dipsticks in screening for diabetes**
25 398 **mellitus.** *Emerg Med J* 2006, **23**: 138.
26
27 399 8. Engelgau MM, Narayan KM, Herman WH: **Screening for type 2 diabetes.** *Diabetes Care* 2000,
28 400 **23**: 1563-1580.
29
30 401 9. World Health Organization. Screening for Type 2 Diabetes. Report of a World Health Organization
31 402 and International Diabetes Federation meeting. 2003. Geneva, Switzerland, World Health
32 403 Organization.
33 404 Ref Type: Report
34
35 405 10. van Pelt M: **Improving access to education and care in Cambodia.** *Diabetes Voice* 2009, **54**.
36
37 406 11. Taniguchi D, LoGerfo J, van PM, Mielcarek B, Huster K, Haider M *et al.*: **Evaluation of a multi-**
38 407 **faceted diabetes care program including community-based peer educators in Takeo province,**
39 408 **Cambodia, 2007-2013.** *PLoS One* 2017, **12**: e0181582.
40
41 409 12. van Olmen J, Eggermont N, van Pelt M, Hen H, de Man J, Schellevis F *et al.*: **Patient-centred**
42 410 **innovation to ensure access to diabetes care in Cambodia: the case of MoPoTsyo.** *J Pharm*
43 411 *Policy Pract* 2016, **9**: 1.
44
45 412 13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L *et al.*: **STARD 2015: an**
46 413 **updated list of essential items for reporting diagnostic accuracy studies.** *BMJ* 2015, **351**:
47 414 h5527.
48
49 415 14. World Health Organization. WHO STEPS Surveillance Manual. World Health Organization . 6-13-
50 416 2008. World Health Organization. 4-30-2015.
51 417 Ref Type: Electronic Citation
52
53
54
55
56
57
58
59
60

- 1
2
3 418 15. Obuchowski NA: **Sample size calculations in studies of test accuracy.** *Stat Methods Med Res*
4 419 1998, **7**: 371-392.
5
6 420 16. King H, Keuky L, Seng S, Khun T, Roglic G, Pinget M: **Diabetes and associated disorders in**
7 421 **Cambodia: two epidemiological surveys.** *Lancet* 2005, **366**: 1633-1639.
8
9 422 17. **Standards of Medical Care in Diabetes-2016 Abridged for Primary Care Providers.** *Clin*
10 423 *Diabetes* 2016, **34**: 3-21.
11
12 424 18. Bartoli E, Fra GP, Carnevale Schianca GP: **The oral glucose tolerance test (OGTT) revisited.**
13 425 *Eur J Intern Med* 2011, **22**: 8-12.
14
15 426 19. World Health Organization. Waist circumference and waist-hip ratio: report of a WHO expert
16 427 consultation, Geneva, 8-11 December 2008. 2011. Geneva, Switzerland, World Health
17 428 Organization.
18 429 Ref Type: Report
19
20 430 20. **The effect of intensive treatment of diabetes on the development and progression of long-term**
21 431 **complications in insulin-dependent diabetes mellitus. The Diabetes Control and**
22 432 **Complications Trial Research Group.** *N Engl J Med* 1993, **329**: 977-986.
23
24 433 21. University of Health Sciences Cambodia, Ministry of Health Cambodia. Prevalence of non-
25 434 communicable disease risk factors in Cambodia; STEPS Survey Country Report. 2010.
26 435 Ref Type: Report
27
28 436 22. Liu M, Pan CY, Jin MM, Su HY, Lu JM: **[The reproducibility and clinical significance of oral**
29 437 **glucose tolerance test for abnormal glucose metabolism].** *Zhonghua Nei Ke Za Zhi* 2007, **46**:
30 438 1007-1010.
31
32 439 23. Ko GT, Chan JC, Woo J, Lau E, Yeung VT, Chow CC *et al.*: **The reproducibility and usefulness**
33 440 **of the oral glucose tolerance test in screening for diabetes and other cardiovascular risk**
34 441 **factors.** *Ann Clin Biochem* 1998, **35 (Pt 1)**: 62-67.
35
36 442 24. English E, Idris I, Smith G, Dhatariya K, Kilpatrick ES, John WG: **The effect of anaemia and**
37 443 **abnormalities of erythrocyte indices on HbA1c analysis: a systematic review.** *Diabetologia*
38 444 2015, **58**: 1409-1421.
39
40 445 25. Flessa S, Zembok A: **Costing of diabetes mellitus type II in Cambodia.** *Health Econ Rev* 2014,
41 446 **4**: 24.
42
43 447
44 448
45 449
46 450
47
48
49
50
51
52
53
54
55
56
57
58
59
60



86x71mm (300 x 300 DPI)

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)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	1
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3
	4	Study objectives and hypotheses	4
METHODS			
<i>Study design</i>	5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
<i>Participants</i>	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
<i>Test methods</i>	10a	Index test, in sufficient detail to allow replication	5
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6
<i>Analysis</i>	14	Methods for estimating or comparing measures of diagnostic accuracy	7
	15	How indeterminate index test or reference standard results were handled	7
	16	How missing data on the index test and reference standard were handled	7
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	7
	18	Intended sample size and how it was determined	6
RESULTS			
<i>Participants</i>	19	Flow of participants, using a diagram	17
	20	Baseline demographic and clinical characteristics of participants	15
	21a	Distribution of severity of disease in those with the target condition	15
	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	5
<i>Test results</i>	23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	15
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	15
	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	10
	27	Implications for practice, including the intended use and clinical role of the index test	11
OTHER INFORMATION			
	28	Registration number and name of registry	NA
	29	Where the full study protocol can be accessed	NA
	30	Sources of funding and other support; role of funders	14

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>.

