ORIGINAL ARTICLE

Reliability of Reagent Strips for Semi-quantitative Measurement of Glucosuria in a Neonatal Intensive Care Setting

Jolita Bekhof a,*, Boudewijn J. Kollen b, Sjef van de Leur c, Joke H. Kok d, Irma H.L.M. van Straaten a

a Amalia Children’s Center, Isala, Zwolle, The Netherlands
b Department of General Practice, University of Groningen, University Medical Centre Groningen, The Netherlands
c Department of Clinical Chemistry, Isala Clinics, Zwolle, The Netherlands
d Academic Medical Centre, Department of Neonatology, Amsterdam, The Netherlands

Received Oct 14, 2013; received in revised form Dec 2, 2013; accepted Dec 10, 2013
Available online 30 April 2014

Key Words
agreement; dipsticks; glucosticks; precision; reproducibility; urine; validity

Background: Glucosuria in preterm infants is often measured using a visually readable reagent strip, e.g., when monitoring total parenteral nutrition or during sepsis or when treating with corticosteroids. However, the specific circumstances in a neonatal intensive care unit (NICU), such as the use of diapers and the high temperature in incubators, could affect its reliability.

Objectives: To evaluate the reliability of the semi-quantitative measurement of glucosuria under the specific circumstances of a NICU setting.

Methods: Nine hundred assessments of artificially supplemented (contrived) urine samples, intended to simulate pathological specimens, were performed under the following varying conditions: environmental temperature (21°C, 24°C, and 34°C); different times of contact of the urine with the diaper; and using two different methods of collecting urine from the diaper. Each reagent strip was read independently by three observers. The test strips scores were categorized as 0, 1+, 2+, 3+, or 4+ in ascending degree of glucosuria.

Results: Agreement was excellent under all the different conditions (temperature, weighted kappa (κw) = 0.92; method of urine collection, κw = 0.88; time, p = 0.266). Inter-observer reliability was very good (multi-rater κ = 0.81). The deviation between the different conditions was seldom larger than one category (2.9%). The reagent strip readings were concordant with the true urinary glucose concentrations in 79.0% of assessments. The discordance was never larger than one category.

Conclusion: The reliability of the semi-quantitative measurement of glucosuria in newborn infants using reagent strips is good, even under the conditions of a NICU. Changes in the rating of reagent strips of more than one category are most likely to be beyond measurement error.

* Corresponding author. Princess Amalia Department of Pediatrics, Isala Klinieken, Dr van Heesweg 2, P.O. Box 10400, 8000 GK Zwolle, The Netherlands.
E-mail address: j.bekhof@isala.nl (J. Bekhof).

http://dx.doi.org/10.1016/j.pedneo.2013.12.008
1875-9572/Copyright © 2014, Taiwan Pediatric Association. Published by Elsevier Taiwan LLC. All rights reserved.
1. Introduction

Disturbed glucose homeostasis is often encountered in pre-term neonates. Measuring glucosuria in very premature infants can be useful, e.g., in monitoring the administration of total parenteral nutrition, insulin, corticosteroids, during sepsis, or for the evaluation of renal function. Glucosuria can be measured quantitatively in the laboratory by spectrophotometry or semi-quantitatively using reagent strips interpreted visually at the bedside. The use of reagent strips for urinalysis is quick, easy, and economical and thus they are often used as a point-of-care test. We previously showed good inter-rater reliability of reagent strips for the measurement of glucosuria (κ = 0.81).

However, the conditions in a neonatal intensive care unit (NICU) or high care unit differ substantially from those in the manufacturer’s product information. Contact of the urine with diapers and high environmental humidity and temperature in the incubator are factors that may influence the results obtained with the reagent strips. Another major issue is the very small volume of urine samples when considering urinalysis in (premature) neonates. It is often not possible to collect a single large sample of urine into which to dip the strip, as recommended by the manufacturer’s product information. As a consequence, reagent strips are often used by directly pressing the strip onto the wet diaper. To our knowledge, data on the reliability of using reagent strips for the measurement of glucosuria in neonates are lacking.

The aim of this study was to determine the reliability of the measurement of glucosuria by reagent strips read visually under the specific conditions of a NICU.

2. Methods

We used Combur reagent strips (Combur®Test, Roche Mannheim, Germany) to test artificially supplemented (contrived) urine samples, intended to simulate pathological specimens. The reagent strip scores were categorized as 0, 1+, 2+, 3+, and 4+ in ascending degree of glucosuria, resulting in five different test strip colors. After 60 seconds of contact between the urine and the test strip, the color of the test strip was compared with a standardized color scale, corresponding with the five different categories of glucosuria. According to the product information for the Combur strips, the corresponding ranges were 0.6–5.0 mmol/L for score 1+, 3.3–7.7 mmol/L for score 2+, 14.8–19.2 mmol/L for score 3+, and 52.8–57.2 mmol/L for score 4+. Five different quantities of glucose corresponding to the mid-range glucose concentrations (0, 2.8, 5.5, 17.0, and 55.0 mmol/L) were added to the freshly collected urine of healthy volunteers.

A total of 300 urine samples were used, consisting of 60 samples of five different glucose concentrations. The urine samples were distributed over disposable diapers with insert gauzes. Half of the diapers were kept at room temperature (21°C) and half were kept in an incubator at 34°C. The reagent strips were read from urine collected by two different non-invasive methods: the diaper-strip and the drip-strip methods. The drip-strip method uses a piece of cloth (gauze) in the diaper. After compression of the wet cloth, urine can be evacuated with a syringe and subsequently applied to the reagent strip. Gloves were worn during the procedure. Alternatively, reagent strips are frequently interpreted after pressing the strip in a wet diaper for a few seconds. This method is referred to as the diaper-strip method. The reagent strips were applied in a randomized order of glucose concentrations. Reagent strips were read at five different times of contact between the diaper and urine (0, 30, 60, 120, and 180 minutes). All reagent strips were independently read by three different observers, all intensive care neonatology nurses with at least 5 years of experience, from 60 to 120 seconds after contact with the urine sample.

Analysis of urinary glucose in the laboratory was carried out by the Gluco-quant glucose/HK method (Roche Diagnostics, Mannheim, Germany) on a Modular P800 instrument (Hitachi, Tokyo, Japan). This is an enzymatic hexokinase method and the rate of NADPH formation is measured photometrically. Urine collected by the diaper-strip method was used for this purpose. The laboratory staff was not aware of the reagent strip results.

2.1. Statistical methods

Assessments of the agreement between the results of the reagent strips were quantified by calculating the percentage agreement and weighted kappa (κw) values using quadratic weights. We interpreted the kappa statistics as follows: poor (κ = 0–0.40), fair (κ = 0.41–0.75), or excellent (κ = 0.76–1.0).

We used non-parametric tests for the analysis of the effect of time spent in the diaper on the urine samples (Friedman’s ANOVA) and the effect of environmental temperature (Wilcoxon) on glucose concentrations. All analyses, except for the calculation of κw, were performed with SPSS Version 18.0.

3. Results

The results of visual reading showed excellent agreement with the true concentration of glucose measured in the laboratory (Table 1; κw = 0.934, 95% confidence interval (CI) 0.93–0.94, raw agreement 79.0%, 95% CI 76.2–81.6). The results equally overestimated and underestimated the true degree of glucosuria (8.7% too low vs. 12.3% too high). Agreement was lowest for categories 1+ and 2+ (67.2% and 63.5%, respectively, vs. 94.0%, 77.9%, and 88.6% for
24.4% of samples after 0, 30, 60, 120, and 180 minutes, incubator temperature in 15.6%, 20.0%, 18.9%, 28.9%, and temperature were higher than the ratings of samples at the
of the reagent strips of the urine samples at room tem-

excellent for both environmental temperatures (room

The samples at room temperature were rated higher than
those in the incubator (higher in 21.6%, lower in 3.6%). This
difference was significantly influenced by the time the
urine spent in the diaper (p = 0.027, Friedman); the ratings
of the reagent strips of the urine samples at room tem-
perature were higher than the ratings of samples at the
incubator temperature in 15.6%, 20.0%, 18.9%, 28.9%, and
24.4% of samples after 0, 30, 60, 120, and 180 minutes,
respectively.
Agreement with the true glucose concentration was
excellent for both environmental temperatures (room
temperature \( \kappa_w = 0.919 \) with 78.0% raw agreement; incu-
bator temperature \( \kappa = 0.949 \) with 80.0% raw agreement).

3.1. Temperature
The results of the comparisons between the two different
environmental temperatures are presented in Table 2. The
agreement of the reagent strips used in the two different
environmental temperatures is excellent (\( \kappa_w = 0.921 \), 95%
CI 0.909–0.933, raw agreement 74.9%, 95% CI 70.6–78.8). The
samples at room temperature were rated higher than
those in the incubator (higher in 21.6%, lower in 3.6%). This
difference was significantly influenced by the time the
urine spent in the diaper (p = 0.027, Friedman); the ratings
of the reagent strips of the urine samples at room tem-
perature were higher than the ratings of samples at the
incubator temperature in 15.6%, 20.0%, 18.9%, 28.9%, and
24.4% of samples after 0, 30, 60, 120, and 180 minutes,
respectively.
Agreement with the true glucose concentration was
excellent for both environmental temperatures (room
temperature \( \kappa_w = 0.919 \) with 78.0% raw agreement; incu-
bator temperature \( \kappa = 0.949 \) with 80.0% raw agreement).

3.2. Urine collection method
The results of agreement between the different urine
collection methods are presented in Table 3, showing good
agreement between the different methods (\( \kappa_w = 0.877 
95\% CI 0.864–0.890, agreement 62.2\%, 95\% CI 58.9–66.7\%).
The results of the drip-strip method scored higher than the
diaper-strip method in 37.8% of assessments, and lower in
0.0%. The diaper-strip method shows better agreement
with the true glucose concentration than the drip-strip
method (diaper strip \( \kappa_w = 0.965 \) with 82.9% raw agree-
mint, drip strip \( \kappa = 0.905 \) with 75.1% raw agreement).
Overestimation of the true glucose concentration occurred
in 24.4% of drip-strip samples compared with 0.002% in
diaper-strip samples. Underestimation of the true glucose
concentration was found in 0.004% of drip-strip samples and
in 16.9% of diaper-strip samples.

3.3. Time
The accuracy of the reagent strips showed statistically
significant differences in categories 2+ and 4++; however,
no trend with time was seen (Table 4), meaning that the
accuracy did not change with the length of time that the
urine stayed in the diaper. The maximum change in the
time of the reagent strip readings was one category.
The glucose concentration measured by the laboratory
in urine samples collected through the drip-strip method

Table 1 Agreement in rating of glucosuria determined through reagent strips compared with true urinary glucose concentrations measured in a laboratory.

<table>
<thead>
<tr>
<th>Reagent strip</th>
<th>True glucosuria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>173</td>
</tr>
<tr>
<td>+</td>
<td>7</td>
</tr>
<tr>
<td>++</td>
<td>0</td>
</tr>
<tr>
<td>+++</td>
<td>0</td>
</tr>
<tr>
<td>++++</td>
<td>0</td>
</tr>
</tbody>
</table>

Values are numbers of samples. \( \kappa_w = 0.934 \), 95\% confidence interval 0.928–0.940.

Table 2 Agreement in rating of glucosuria determined through reagent strips at two different environmental temperatures.

<table>
<thead>
<tr>
<th>Incubator temperature (°C)</th>
<th>Room temperature (°C)</th>
<th>0</th>
<th>+</th>
<th>++</th>
<th>+++</th>
<th>++++</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>83</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>98</td>
</tr>
<tr>
<td>+</td>
<td></td>
<td>3</td>
<td>55</td>
<td>48</td>
<td>0</td>
<td>0</td>
<td>106</td>
</tr>
<tr>
<td>++</td>
<td></td>
<td>4</td>
<td>36</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>60</td>
</tr>
<tr>
<td>+++</td>
<td></td>
<td>0</td>
<td>0</td>
<td>76</td>
<td>14</td>
<td>94</td>
<td>Total 86</td>
</tr>
<tr>
<td>++++</td>
<td></td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>87</td>
<td>92</td>
<td></td>
</tr>
</tbody>
</table>

Values are numbers of samples. \( \kappa_w = 0.921 \), 95\% confidence interval 0.909–0.933.

Table 3 Agreement in rating of glucosuria determined through reagent strips between two methods of urine collection from the diaper.

<table>
<thead>
<tr>
<th>Diaper-strip method</th>
<th>0</th>
<th>+</th>
<th>++</th>
<th>+++</th>
<th>++++</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diaper-strip</td>
<td>83</td>
<td>15</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>101</td>
</tr>
<tr>
<td>+</td>
<td>0</td>
<td>36</td>
<td>83</td>
<td>10</td>
<td>0</td>
<td>129</td>
</tr>
<tr>
<td>++</td>
<td>0</td>
<td>16</td>
<td>30</td>
<td>0</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>+++</td>
<td>0</td>
<td>0</td>
<td>63</td>
<td>29</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>++++</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>82</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>83</td>
<td>51</td>
<td>102</td>
<td>103</td>
<td>111</td>
<td>450</td>
</tr>
</tbody>
</table>

Values are numbers of samples. \( \kappa_w = 0.877 \), 95\% confidence interval 0.864–0.890.
did not change significantly over the 180 minutes of time spent in the diaper (Friedman’s ANOVA, p = 0.266).

4. Discussion

The urinary glucose concentration measured by the reagent strips is not influenced by the time spent in the diaper and it is only slightly influenced by the temperature of the incubator and by the method of urine collection. The variability largely remains restricted to only one category difference, which is seen mainly in category 1+ and 2+. This can be explained by the fact that the ranges of glucosuria of categories 1+ and 2+ overlap (0.6–5.0 mmol/L for category 1+, 3.3–7.7 mmol/L for category 2+). We therefore conclude that the use of reagent strips for semi-quantitative measurement is reliable enough for use under NICU conditions, especially when categories 1+ and 2+ are taken together as one category. To our knowledge, this is the first study to address the reliability of reagent strips for the measurement of glucosuria in a NICU.

The two factors that slightly influenced the measurement of glucosuria were temperature and the collection method. We purposely used room temperature as a reference standard and chose the maximum temperature used in our NICU, arguing that this would reveal the maximum possible difference in results. Because we found only a slight, not clinically relevant, influence of these two extreme temperatures, we can safely conclude that lesser differences in temperature will not influence the results. With respect to the method of urine collection, the reagent strips are best applied by pushing the strip onto the wetted diaper (the diaper method). This is the most reliable and simple method; the contents of the diaper do not prevent misleading high or low results. Moreover, this measurement error may be largely avoided when the categories 1+ and 2+ are taken together as one category.

Conflicts of interest

The authors have no conflicts of interest relevant to this article.

Acknowledgments

The authors gratefully acknowledge Liesbeth Groot-Jebbink, Corrie Deiman, Carin Bunkers, Annelies Vogelszang, and Gerrie Veeneklaas for their help with the assessment of the reagent strips.