University of Groningen

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

Bekhof, Jolita; Kollen, Boudewijn J.; van de Leur, Sjef; Kok, Joke H.; van Straaten, Irma H. L. M.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)
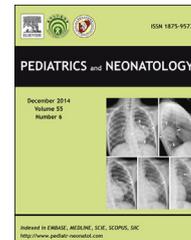
ORIGINAL ARTICLE

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

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 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 (κ_w) = 0.92; method of urine collection, κ_w = 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.

1. Introduction

Disturbed glucose homeostasis is often encountered in pre-term neonates.[1−5] Measuring glucosuria in very premature infants can be useful, e.g., in monitoring the administration of total parenteral nutrition,[6] insulin,[7] corticosteroids,[8] during sepsis,[9] or for the evaluation of renal function.[10,11] 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.[12] We previously showed good inter-rater reliability of reagent strips for the measurement of glucosuria ($\kappa = 0.81$).[13]

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.[14] 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 (Combur³Test, 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 wetted 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 (κ_w) values using quadratic weights.[15] We interpreted the kappa statistics as follows: poor ($\kappa = 0−0.40$), fair ($\kappa = 0.41−0.75$), or excellent ($\kappa = 0.76−1.0$).[16]

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 κ_w, 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; $\kappa_w = 0.934$, 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

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

Reagent strip	True glucosuria					
	0	+	++	+++	++++	Total
0	173	11	0	0	0	184
+	7	121	52	0	0	180
++	0	48	94	6	0	148
+++	0	0	34	152	9	195
++++	0	0	0	22	171	193
Total	180	180	180	180	180	900

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

categories 0, 3+, and 4+, respectively). As we showed previously,[13] inter-observer reliability was good (85% overall agreement and multi-rater $\kappa = 0.81$). Inter-observer scores for the reagent strips never deviated more than one category from each other.

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 temperature 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; incubator temperature $\kappa = 0.949$ with 80.0% raw agreement).

Overestimation of the true glucose concentration occurred in 17.3% of samples at room temperature compared with 7.3% for those at the incubator temperature. Underestimation of the true glucose concentration was found in 4.7% of samples at room temperature and in 12.7% of samples at incubator temperature. The laboratory measured glucose concentration in the samples retrieved through the diaper-stick method was also higher at room temperature than at incubator temperature (mean difference 0.93 mmol/L, 95% CI 0.52–1.35 mmol/L; Wilcoxon $p < 0.001$). However, all measured laboratory concentrations, from both room and incubator temperature samples, remained within the ranges corresponding to the categories of the reagent strip as provided by the manufacturer.

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

Incubator temperature (°C)	Room temperature (°C)					
	0	+	++	+++	++++	Total
0	83	15	0	0	0	98
+	3	55	48	0	0	106
++	0	4	36	20	0	60
+++	0	0	4	76	14	94
++++	0	0	0	5	87	92
Total	86	74	88	101	101	450

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.

Diaper-strip method		Drip-strip method					
		0	+	++	+++	++++	Total
Diaper-strip	0	83	15	3	0	0	101
	+	0	36	83	10	0	129
	++	0	0	16	30	0	46
	+++	0	0	0	63	29	92
	++++	0	0	0	0	82	82
	Total	83	51	102	103	111	450

Values are numbers of samples. $\kappa_w = 0.877$, 95% confidence interval 0.864–0.890.

Table 4 Agreement for glucosuria between reagent strips read visually and true glucosuria (laboratory measurement) after different periods of contact.

True glucose concentration	Time (min)					
	0	30	60	120	180	p^*
0	91.7	100	100	94.4	94.4	0.061
+	77.8	63.9	58.3	72.2	63.9	0.184
++	44.4	63.9	44.4	47.2	61.1	0.003
+++	83.3	83.3	77.8	88.9	88.9	0.056
++++	75.0	92.7	100	100	100	<0.001
Total	74.4	80.6	76.1	80.6	82.7	<0.001

$n = 900$. Values are % of samples.
 * Friedman's ANOVA.

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 appear to influence the glucose concentrations, which is in accordance with earlier reports.[14,17,18] When too little urine is available, e.g., with extremely absorbent diapers, urine can be collected by putting a gauze in the diaper. The urine collected by this drip-stick method tends to overestimate the degree of glucosuria. This may be explained by the fact that more water could have evaporated from the gauze compared with urine collected in the diaper. However, the largest part of the variation in the reagent strips result can be explained by inter-observer

reliability ($\kappa_w = 0.81$).[13] This inter-observer variation is a factor that cannot be neglected, as in daily clinical practice it is unavoidable that different staff will use the reagent strips.

The strength of our study is the use of a large number of samples covering the complete range of glucose concentrations. In particular, when using κ for the evaluation of reliability, it is of vital importance to be cautious in cases of low or high prevalence.[19] By the use of equally distributed prevalence of all degrees of glucosuria, we aimed to prevent misleading high or low κ values. Another strength is the fact that we studied the reliability under the specific condition of a NICU, which has not been reported previously.

This study has limitations. The use of predefined, categorized glucose concentrations, scored under optimum controlled study conditions, may affect the generalizability of our results to real-life clinical practice. In clinical settings, the reagent strip is read under variable circumstances, with variations in room lighting and, importantly, variously trained staff. Future studies using urine samples from patients in real-time clinical practice with a wide spectrum of glucosuria are needed to elucidate this matter.

In summary, the semi-quantitative measurement of glucosuria using reagent strips read visually seems to be a reliable and simple bedside method, suitable for use in a NICU. Changes in the rating of reagent strips of more than one category are almost certainly beyond measurement error. 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

References

1. Hey E. Hyperglycaemia and the very preterm baby. *Semin Fetal Neonatal Med* 2005;**10**:377—87.
2. Ward Platt M, Deshpande S. Metabolic adaptation at birth. *Semin Fetal Neonatal Med* 2005;**10**:341—50.
3. Farrag HM, Cowett RM. Glucose homeostasis in the micropremie. *Clin Perinatol* 2000;**27**:1—22.
4. Beardsall K. Measurement of glucose levels in the newborn. *Early Hum Dev* 2010;**86**:263—7.
5. Beardsall K, Vanhaesebrouck S, Ogilvy-Stuart AL, Vanhole C, Palmer CR, Ong K, et al. Prevalence and determinants of hyperglycemia in very low birth weight infants: cohort analyses of the NIRTURE study. *J Pediatr* 2010;**157**:715—9.
6. Falcão MC, Leone CR, Ramos JL. Is glycosuria a reliable indicator of adequacy of glucose infusion rate in preterm infants? *Sao Paulo Med J* 1999;**117**:19—24.

7. Decaro MH, Vain NE. Hyperglycaemia in preterm neonates: what to know, what to do. *Early Hum Dev* 2011;**87**:S19—22.

8. Mitanchez D. Glucose regulation in preterm newborn infants. *Horm Res* 2007;**68**:265—71.

9. Manzoni P, Castagnola E, Mostert M, Sala U, Galletto P, Gomirato G. Hyperglycaemia as a possible marker of invasive fungal infection in preterm neonates. *Acta Paediatr* 2006;**95**:486—93.

10. Coulthard MG, Hey EN. Renal processing of glucose in well and sick neonates. *Arch Dis Child Fetal Neonatal Ed* 1999;**81**:F92—8.

11. Drukker A, Guignard JP. Renal aspects of the term and preterm infant: a selective update. *Curr Opin Pediatr* 2002;**14**:175—82.

12. Wilson LA. Urinalysis. *Nurs Stand* 2005;**19**:51—4.

13. Bekhof J, Kollen BJ, Groot-Jebbink LJ, Deiman C, Van de Leur SJ, Van Straaten HL. Validity and interobserver agreement of reagent strips for measurement of glucosuria. *Scand J Clin Lab Invest* 2011;**71**:248—52.

14. Burke N. Alternative methods for newborn urine sample collection. *Pediatr Nurs* 1995;**21**:546—9.

15. Richard Lowry 2001—2014. Kappa as a Measure of Concordance in Categorical Sorting. Available at http://vassarstats.net/kappa.html. Accessd April 17, 2014.

16. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;**76**:378—82.

17. Muratore C, Dhanireddy R. Urine collection from disposable diapers in premature infants: biochemical analysis. *Clin Pediatr (Phila)* 1993;**32**:314—5.

18. Roberts SB, Lucas A. Measurement of urinary constituents and output using disposable napkins. *Arch Dis Child* 1985;**60**:1021—4.

19. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol* 2012;**8**:23—34.