Journal of Clinical Laboratory Analysis 19:267–275 (2005)

Evaluation of a Standardized Procedure for Counting Microscopic Cells in Body Fluids

Jane F. Emerson,[1]* and Scott S. Emerson[2]
[1]*Department of Pathology, University of California–Irvine, Irvine, California*
[2]*Department of Biostatistics, University of Washington, Seattle, Washington*

A standardized urinalysis and manual microscopic cell counting system was evaluated for its potential to reduce intra- and interoperator variability in urine and cerebrospinal fluid (CSF) cell counts. Replicate aliquots of pooled specimens were submitted blindly to technologists who were instructed to use either the Kova system with the disposable Glasstic slide (Hycor Biomedical, Inc., Garden Grove, CA) or the standard operating procedure of the University of California–Irvine (UCI), which uses plain glass slides for urine sediments and hemacytometers for CSF. The Hycor system provides a mechanical means of obtaining a fixed volume of fluid in which to resuspend the sediment, and fixes the volume of specimen to be microscopically examined by using capillary filling of a chamber containing in-plane counting grids. Ninety aliquots of pooled specimens of each type of body fluid were used to assess the inter- and intraoperator reproducibility of the measurements. The variability of replicate Hycor measurements made on a single specimen by the same or different observers was compared with that predicted by a Poisson distribution. The Hycor methods generally resulted in test statistics that were slightly lower than those obtained with the laboratory standard methods, indicating a trend toward decreasing the effects of various sources of variability. For 15 paired aliquots of each body fluid, tests for systematically higher or lower measurements with the Hycor methods were performed using the Wilcoxon signed-rank test. Also examined was the average difference between the Hycor and current laboratory standard measurements, along with a 95% confidence interval (CI) for the true average difference. Without increasing labor or the requirement for attention to detail, the Hycor method provides slightly better interrater comparisons than the current method used at UCI. J. Clin. Lab. Anal. 19:267–275, 2005. © 2005 Wiley-Liss, Inc.

Key words: Poisson variability; hemacytometer; urinalysis; cerebrospinal fluid; manual examination

INTRODUCTION

Multiple sources of variation in the microscopic evaluation of urine and other body fluids contribute to a lack of reproducibility within and between institutions. This in turn contributes to problems with reference value determinations and clinical interpretations. Because laboratory quantification of cells and formed elements usually involves examining a small volume taken from a clinical specimen, such measures are subject to sampling variability. Theory would dictate that the actual number of elements in a sample drawn from a homogeneous specimen would follow a Poisson distribution according to the concentration in the entire specimen and the volume of the sample measured. This variability cannot be removed by any means (though making measurements on larger volumes can increase the precision of estimates of concentrations).

In practice, additional (extra-Poisson) sources of variation arise due to specimen preparation. These include the degree to which the specimen is adequately homogenized prior to sampling, slide preparation techniques that may reintroduce heterogeneity

Grant sponsor: Hycor Biomedical, Inc.

*Correspondence to: Jane F. Emerson, M.D., Ph.D., Chief of Clinical Pathology, Department of Pathology, Rt. 38, University of California–Irvine Medical Center, 101 The City Dr. S., Orange, CA 92868. E-mail: jemerson@uci.edu

Received 18 March 2005; Accepted 16 August 2005
DOI 10.1002/jcla.20089
Published online in Wiley InterScience (www.interscience.wiley.com).

(e.g., edge effects), accurate measurement of volume used for examination (drops vs. chamber volumes), selection of microscopic fields that may reflect observer bias, human error in the identification and enumeration of cells and elements, and variable practices for reporting results.

Systematic trends in these additional sources of extra-Poisson variation across observers (laboratory technologists) lead to less reliable measurements and ultimately to less clinically useful markers of disease or prognosis.

As a step toward the goal of identifying laboratory procedures that might remove the above technical sources of variation, we evaluated a candidate standardized procedure, the Hycor Kova (Hycor Biomedical, Inc., Garden Grove, CA) system for microscopic examination of urine and cerebrospinal fluid (CSF). The Hycor system addresses sources of variability by 1) providing a mechanical means of obtaining a fixed volume of fluid in which to resuspend the sediment, 2) fixing the volume of specimen to be microscopically examined by using capillary filling of a counting chamber, and 3) dictating the examination of predefined counting grids within the plane of the fluid. Previous reports have evaluated this system with respect to correlations with glass hemacytometers (1), another standardized system (2), flow cytometry (3), automated hematology analyzers (4), and an automated urine analysis system (5), and as part of an integrated algorithm along with automated cell counting to evaluate urine sediment (6). These reports did not address the ability of a system to reduce the high variability in manual cell counts.

We compared this procedure with the current standard method used at a university (University of California–Irvine (UCI))-based hospital with respect to intra- and interobserver variability of urine and CSF manual cell counts. Specifically, we considered 1) tendencies toward higher or lower counts using the Hycor system, 2) tendencies toward lower or higher extra-Poisson variability across replicated measurements made by the same observer when using the Hycor system compared to the current laboratory standards, and 3) tendencies toward greater variability across replicated measurements made by different observers using the same criteria for comparison.

MATERIALS AND METHODS

CSF and urine specimens were examined microscopically for cell counts and formed elements using two different processing and microscopic methods. The methods were compared by correlation of laboratory measures across methods and by intra- and interoperator reproducibility within methods.

Our general method used discarded laboratory specimens, either from a single individual or pooled from multiple specimens, which were then divided into two or more aliquots for repeated measurements by the same or different operators. In all cases the specimens were acquired, pooled (if applicable), divided into aliquots, and labeled by nontesting personnel so that the individuals performing the tests were blind to the origin and matching of the aliquots. The urine specimens were 24 hr old when they were pooled and immediately tested. In keeping with IRB standards for this type of study, no identifying or clinical data regarding the patients are available.

For urine, 15 single-patient specimens and 90 aliquots from five pooled specimens were evaluated. The single-patient specimens (minimum volume = 30 mL) were selected randomly from samples submitted to the clinical laboratory for routine testing. Each specimen was divided into two aliquots prior to processing. One aliquot from each specimen was processed and microscopically examined by current standard operating procedures for microscopic urine analysis at our institution. First, 10 mL of a well-mixed specimen were poured into a conical centrifuge tube and spun at 2,000 rpm for 5 min. All but 0.5 mL, as measured visually by tube reference marks, was removed by pipette. The sediment was resuspended by swirling and examined by loading a drop onto a plain glass slide, which was then covered with a glass slip. The cell counts are reported in number per high-power field.

The other aliquot was processed by the Hycor Kova standardized system and examined microscopically using the Glasstic slide. First a defined volume of well-mixed specimen (12 mL) was spun in Kova conical centrifuge tubes for 5 min at 1,500 rpm. A Kova pipetter with a lock-tip was inserted in the tube, trapping 1.0 mL of sediment, while the rest of the sample was decanted. One drop of resuspended sediment was loaded into one well of a 10-well Kova Glasstic slide, which drew a defined of volume of 6.6 μL. Casts were quantified at 100 ×, and all other elements were quantified at 400 ×. The counting grid was such that the ratio of areas of a high-power field to a small square was 3.12. The total number of cells of one type within 10 small squares was counted and converted to cells/μL using the value table supplied by the vendor.

A single technologist performed all of the analyses on the single-patient specimens to assess operator-independent differences between the two methods.

Five pooled specimens (minimum combined volume = 220 mL) were formed from a random collection of patient specimens that had been routinely analyzed. Each pooled specimen was divided into 18 aliquots, and six aliquots were assigned to each of three technologists. Each technologist processed and examined three of the aliquots by the current institutional method, and three

of the aliquots by the Hycor system. The aliquots from one of the pooled specimens were used to assess sample-independent variability in the methods due to operator differences. Then the results from each pooled specimen were compared to evaluate the effects of sample differences on intra-and interoperator variability.

Data recorded for the urine specimens included enumeration and classification of cells, organisms, casts, and crystals. For data analysis, the geometric factor of the ratio of areas of a Hycor counting grid to a high-power field was applied to the counts.

For CSF, the sampling and test performance scheme was carried out as described above for urine, except for the processing step, which was not required. The specimens and aliquots were compared using a standard bright-field hemacytometer and the Glasstic slide.

The entire study was completed over a 3-month period.

Statistical Methods

Tendencies toward higher or lower cell counts across specimens were examined descriptively with a scatterplot of measurements from the Hycor methods vs. measurements obtained from the current laboratory standards. Also examined was the average difference between the Hycor and current laboratory standard measurements for the 15 matched samples, along with a 95% confidence interval (CI) for the true average difference. Tests for systematically higher or lower measurements obtained with the Hycor methods were performed using the Wilcoxon signed-rank test for the 15 paired urine specimens and 15 paired CSF specimens.

Using the five urine and CSF specimens formed from pooling individual specimens, the variability of replicate Hycor measurements made on a single specimen by the same observer was compared with that predicted by a Poisson distribution. Tests were performed for each of three operators using both Hycor and current laboratory standard methods on each of the five pooled urine and five pooled CSF specimens. If random sampling is the only source of variation for measurements across replicate measurements made on the same specimen, the j-th replicate cell count measurement X_{ij} on the i-th sample would obey a Poisson distribution with mean $\lambda_i V_j$, where λ_i represents the mean concentration of cells in a standard unit of volume, and V_j represents the volume examined in that measurement. We tested for departures from the Poisson distribution by examining the conditional distribution of cell counts across the replicate measurements, where we conditioned on the total counts across all the replicates. Letting $X_{i+} = X_{i1} + \ldots + X_{iJ}$ represent the observed total number of cells across J replicate samples having a total volume of $V = V_1 + \ldots + V_J$, the conditional distribution of the cell counts for each replicate is multinomial with parameters $n = X_{i+}$ and proportion vector $(p_1 = V_1/V, \ldots p_J = V_J/V)$. The chi-squared goodness-of-fit test was used to test for samples exhibiting greater variation among the replicate measurements than could be explained by Poisson variation. Each specimen was examined in this manner for replicate measurements made by the same observer. No adjustment was made for the multiple comparisons involved in analyzing results from the multiple samples examined by each operator.

These same measurements were used to assess interoperator agreement among the three laboratory technicians. The measurements were displayed graphically, separately for each sample, with points labeled by the individual technologist (A, B, or C). For each type of cell count (red (RBC) or white (WBC) blood cells), a Poisson regression model was fit separately for the Hycor and current laboratory standard measurements, with separate analyses performed for the urine specimens and the CSF specimens (giving a total of eight distinct Poisson regression models). In each regression model, terms were included for each of the three technicians and each of the five pooled samples. A linear contrast of the regression parameters was then used to test for systematic trends across the three laboratory technologists.

RESULTS

Figure 1 shows the counts of RBC (Fig. 1a), WBC (Fig. 1b), and squamous cells (Fig. 1c) for the 15 paired urine measurements made by a single laboratory technologist using both the Hycor and current laboratory standard methods.

Figure 2 shows the analogous results for RBC (Fig. 2a) and WBC (Fig. 2b) counts for the 15 paired CSF measurements made by a single (though different) laboratory technologist.

Table 1 presents the average observed difference between the Hycor and current laboratory standard methods, along with a 95% CI for the mean difference and a two-sided P-value from the Wilcoxon signed-rank test of a systematic difference in measurements under the two laboratory methods. From these figures and tables, it is clear that while the measurements made with the Hycor and current standard methods are highly correlated, there is some suggestion of a systematic trend toward lower measurements using Hycor for the technologist making urine measurements, and a possible trend toward higher measurements using Hycor for the technologist making CSF measurements.

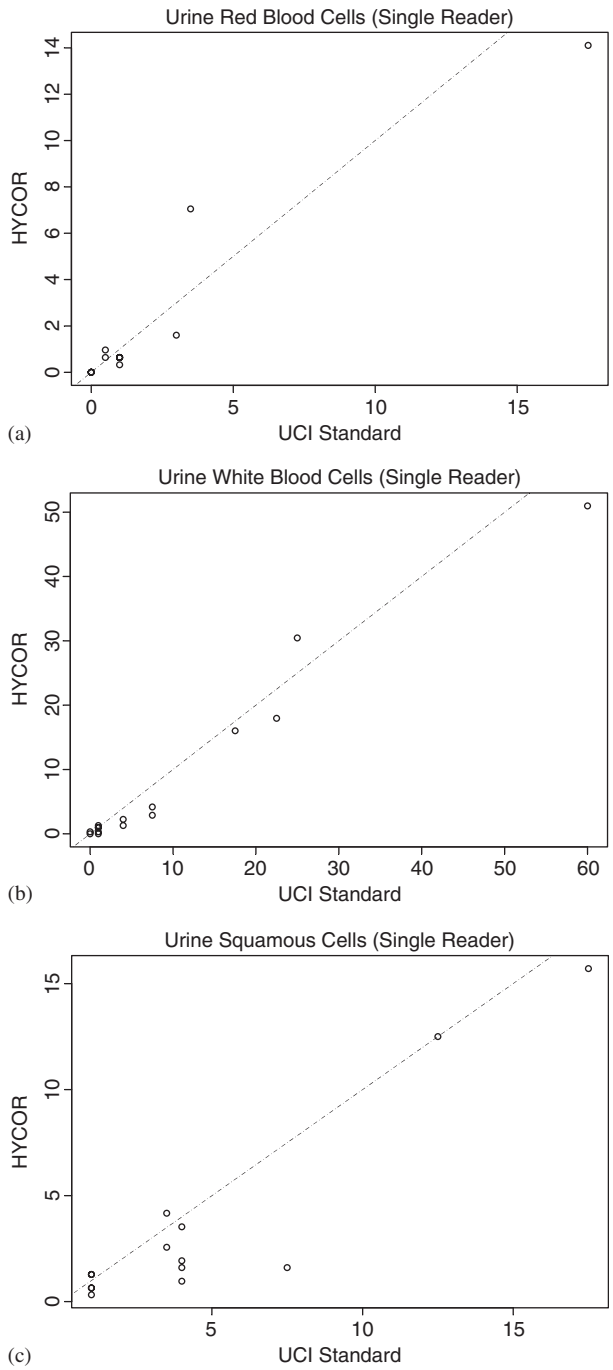

Fig. 1. Counts by one technologist using the current laboratory method (x-axis) and Hycor method (y-axis) for 15 different urine specimens. **a**: RBC count; linear regression gives a slope of 1.08 with an intercept of 0.08 and $R^2 = 0.9288$. **b**: WBC count; linear regression gives a slope of 1.08 with an intercept of 1.0 and $R^2 = 0.9635$. **c**: Squamous cell count; linear regression gives a slope of 1.08 with an intercept of 0.08 and $R^2 = 0.9288$.

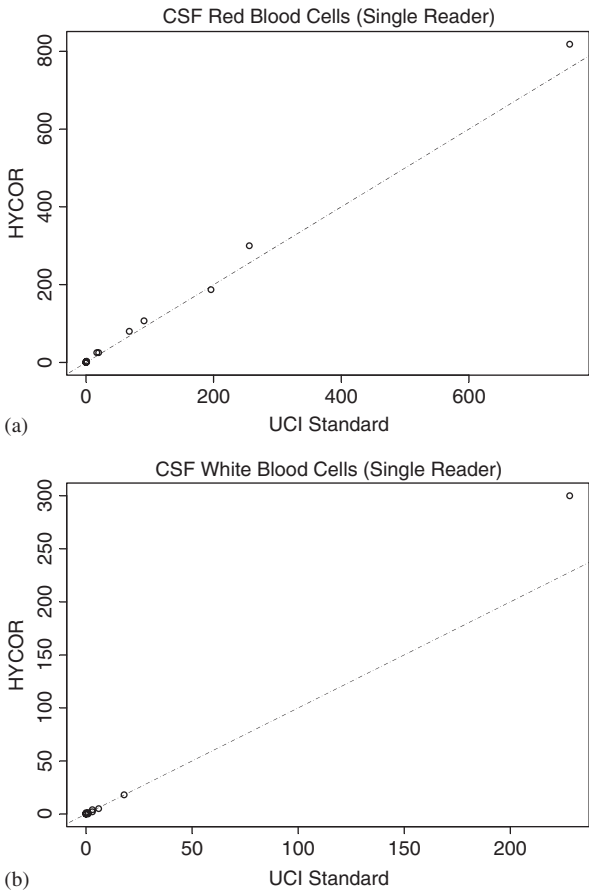

Fig. 2. Counts by one technologist using the current laboratory method (x-axis) and Hycor method (y-axis) for 15 different CSF specimens. **a**: RBC count; linear regression gives a slope of 0.92 with an intercept of –1.64 and $R^2 = 0.998$. **b**: WBC count; linear regression gives a slope of 0.76 with an intercept of 0.53 and $R^2 = 0.9995$.

TABLE 1. Average observed difference between Hycor and current standard methods

Specimen type	Type of cells	Mean diff (95%CI)	*P* val
Urine	RBC	–0.16(–0.93, 0.61)	0.344
	WBC	–1.55(–3.32, 0.22)	0.041
	Squamous Cells	–1.14(–2.08,–0.20)	0.009
CSF	RBC	9.40(–0.84,19.64)	0.017
	WBC	4.80(–5.50,15.10)	0.661

Five pooled urine samples and five pooled CSF samples were each measured in triplicate by each of three laboratory technicians using both the Hycor and current laboratory standard methods. Figures 3 and 4 present the results graphically, and measurements made by the same technologist are labeled accordingly. (Note: The RBC count on specimen HYU17 was off-scale for Fig. 3a.)

Table 2 presents the mean and standard deviation (SD) of results for urinary RBC, WBC, and squamous cell counts, and for CSF RBC and WBC counts by

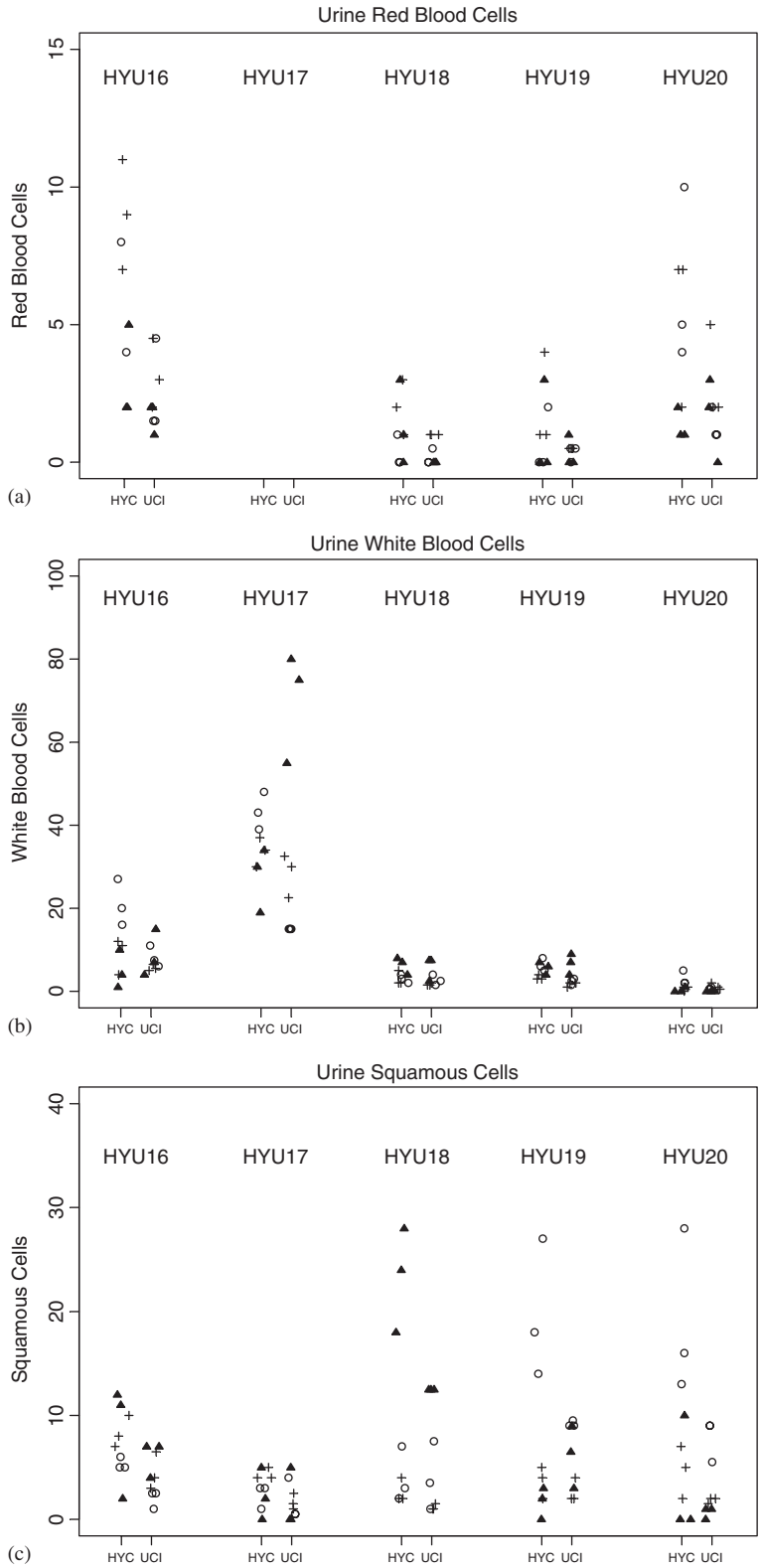

Fig. 3. Interrater comparisons are demonstrated for each pooled urine sample labeled HYU16–HYU20. Symbols +, ▲, and ○ represent the results obtained by three different technologists performing the counts. The column labeled HYC indicates the counts obtained using the Hycor method, and UCI indicates those obtained using the current standard method. **a**: Urine RBC counts; (note that counts for HYU17 are off scale). **b**: Urine WBC counts. **c**: Urine squamous cell counts.

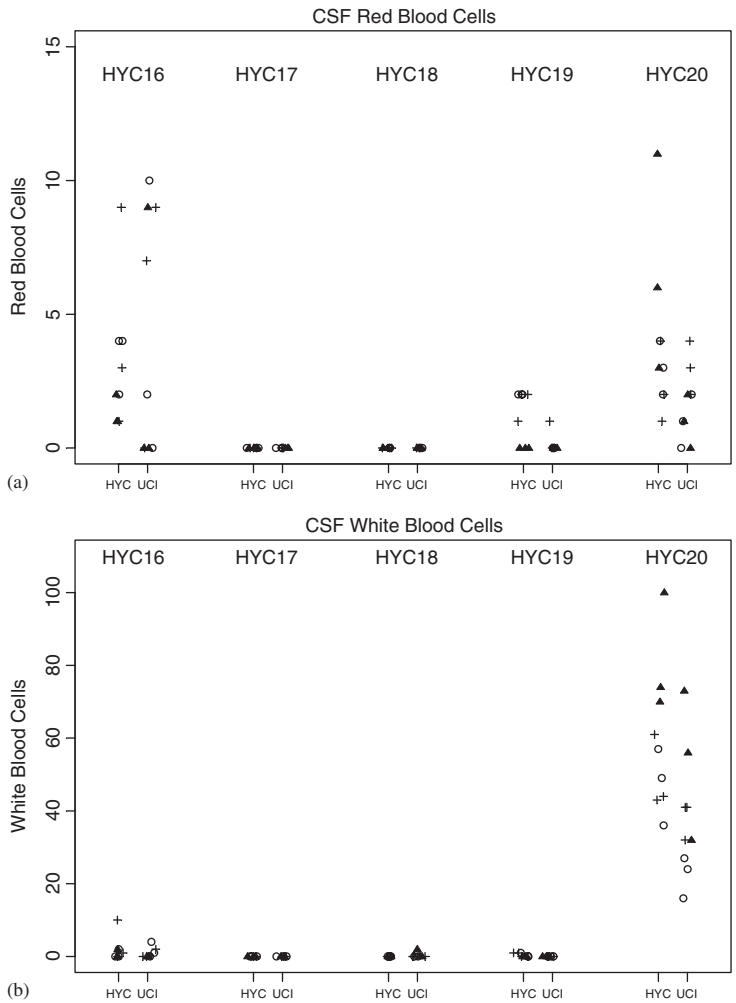

Fig. 4. Interrater comparisons are demonstrated for each pooled CSF sample labeled HYC16–HYC20. Symbols +, ▲, and ○ represent the results obtained by three different technologists performing the counts. The column labeled HYC indicates the counts obtained using the Hycor method, and UCI indicates those obtained using the current standard method. **a**: CSF RBC counts. **b**: CSF WBC counts.

sample and laboratory technologist. Also presented for each combination of sample and laboratory technologist is a test for extra-Poisson variation among samples.

The two methods are similar with respect to the number of triplicate measurements that exhibit extra-Poisson variation. It should be noted that under the current laboratory standards, some measurements were given as a range. We used the midpoint of the range, resulting in an apparent increase in precision for the current method.

The results in Table 2 were also analyzed with respect to variation across laboratory technologists. In Poisson regressions fit to the individual cell counts, highly statistically significant ($P<0.0001$) interrater variability was observed for all cell counts except the CSF RBC, for which the P-values were 0.98 and 0.05 for the Hycor and current laboratory standard methods, respectively. The Hycor methods generally resulted in test statistics that were slightly lower than those for the laboratory standard methods, indicating a trend toward decreasing the effects of various sources of variability. The scientific relevance of this finding is uncertain in the face of such high interrater variation with either method.

DISCUSSION

Variables in the microscopic evaluation of urine and other body fluids contribute to a lack of reproducibility within and between institutions, which in turn contributes to problems with reference value determinations and clinical interpretations. An accurate assessment of precision is difficult to achieve because of the number of pertinent parameters involved and the instability of the samples. An examination of College of American Pathologists (CAP) proficiency testing survey data on hemacytometer fluid counts can be used to estimate

TABLE 2. Tests for extra-Poisson variability in replicate counts for each laboratory method

		Hycor methods			Current standard methods		
		Mean	SD	*P* value (extra-poisson)	Mean	SD	*P* value (extra-poisson)
Urine RBC							
HYU16	Tech A	10.0	7.2	0.006	2.5	1.7	0.301
	Tech B	3.0	1.7	0.368	1.7	0.6	0.819
	Tech C	9.0	2.0	0.641	3.2	1.3	0.607
HYU17	Tech A	431.7	30.2	0.120	150.0	0.0	1.000
	Tech B	188.7	18.6	0.160	325.0	52.2	0.000
	Tech C	281.3	70.9	0.000	43.3	7.6	0.260
HYU18	Tech A	0.3	0.6	0.368	0.2	0.3	0.607
	Tech B	1.3	1.5	0.174	0.0	0.0	1.000
	Tech C	2.0	1.0	0.607	1.0	0.0	1.000
HYU19	Tech A	0.7	1.2	0.135	0.3	0.3	0.779
	Tech B	1.0	1.7	0.050	0.3	0.6	0.368
	Tech C	2.0	1.7	0.223	0.5	0.0	1.000
HYU20	Tech A	6.3	3.2	0.196	1.3	0.6	0.779
	Tech B	1.3	0.6	0.779	1.7	1.5	0.247
	Tech C	5.3	2.9	0.210	3.0	1.7	0.368
Urine WBC							
HYU16	Tech A	21.0	5.6	0.229	8.2	2.6	0.447
	Tech B	5.0	4.6	0.015	8.7	5.7	0.024
	Tech C	9.0	4.4	0.121	5.7	0.8	0.902
HYU17	Tech A	43.3	4.5	0.626	15.0	0.0	1.000
	Tech B	27.7	7.8	0.113	70.0	13.2	0.082
	Tech C	33.7	3.5	0.693	28.3	5.2	0.385
HYU18	Tech A	3.0	1.0	0.717	2.7	1.3	0.552
	Tech B	6.3	2.1	0.505	5.8	2.9	0.240
	Tech C	3.0	1.7	0.368	1.7	0.3	0.951
HYU19	Tech A	6.3	1.5	0.692	2.3	0.8	0.779
	Tech B	5.7	1.5	0.663	6.7	2.5	0.387
	Tech C	3.3	0.6	0.905	1.7	0.6	0.819
HYU20	Tech A	3.0	1.7	0.368	0.7	0.3	0.883
	Tech B	0.3	0.6	0.368	0.0	0.0	1.000
	Tech C	0.7	0.6	0.607	1.2	0.8	0.607
Urine squamous cells							
HYU16	Tech A	5.3	0.6	0.939	2.0	0.9	0.687
	Tech B	8.3	5.5	0.026	6.0	1.7	0.607
	Tech C	8.3	1.5	0.756	4.5	1.8	0.486
HYU17	Tech A	2.3	1.2	0.565	1.7	2.0	0.086
	Tech B	2.3	2.5	0.066	1.7	2.9	0.007
	Tech C	4.3	0.6	0.926	1.7	0.8	0.705
HYU18	Tech A	4.0	2.6	0.174	4.0	3.3	0.068
	Tech B	23.3	5.0	0.338	12.5	0.0	1.000
	Tech C	2.7	1.2	0.607	1.2	0.3	0.931
HYU19	Tech A	19.7	6.7	0.105	9.2	0.3	0.991
	Tech B	1.7	1.5	0.247	6.2	3.0	0.229
	Tech C	3.7	1.5	0.529	2.7	1.2	0.607
HYU20	Tech A	19.0	7.9	0.036	7.8	2.0	0.594
	Tech B	3.3	5.8	0.000	0.7	0.6	0.607
	Tech C	4.7	2.5	0.257	1.8	0.3	0.956
CSF RBC							
HYC16	Tech A	3.3	1.2	0.670	4.0	5.3	0.001
	Tech B	1.3	0.6	0.779	3.0	5.2	0.000
	Tech C	4.3	4.2	0.018	5.3	4.7	0.015
HYC17	Tech A	0.0	0.0	1.000	0.0	0.0	1.000
	Tech B	0.0	0.0	1.000	0.0	0.0	1.000

TABLE 2. Continued

		Hycor methods			Current standard methods		
		Mean	SD	*P* value (extra-poisson)	Mean	SD	*P* value (extra-poisson)
	Tech C	0.0	0.0	1.000	0.0	0.0	1.000
HYC18	Tech A	0.0	0.0	1.000	0.0	0.0	1.000
	Tech B	0.0	0.0	1.000	0.0	0.0	1.000
	Tech C	0.0	0.0	1.000	0.0	0.0	1.000
HYC19	Tech A	2.0	0.0	1.000	0.0	0.0	1.000
	Tech B	0.0	0.0	1.000	0.0	0.0	1.000
	Tech C	1.7	0.6	0.819	0.3	0.6	0.368
HYC20	Tech A	3.0	1.0	0.717	1.0	1.0	0.368
	Tech B	6.7	4.0	0.086	1.0	1.0	0.368
	Tech C	2.3	1.5	0.368	3.0	1.0	0.717
CSF WBC							
HYC16	Tech A	0.7	1.2	0.135	1.7	2.1	0.074
	Tech B	0.7	1.2	0.135	0.0	0.0	1.000
	Tech C	3.7	5.5	0.000	0.7	1.2	0.135
HYC17	Tech A	0.0	0.0	1.000	0.0	0.0	1.000
	Tech B	0.0	0.0	1.000	0.0	0.0	1.000
	Tech C	0.0	0.0	1.000	0.0	0.0	1.000
HYC18	Tech A	0.0	0.0	1.000	0.3	0.6	0.368
	Tech B	0.0	0.0	1.000	0.7	1.2	0.135
	Tech C	0.0	0.0	1.000	0.0	0.0	1.000
HYC19	Tech A	0.3	0.6	0.368	0.0	0.0	1.000
	Tech B	0.0	0.0	1.000	0.0	0.0	1.000
	Tech C	0.7	0.6	0.607	0.0	0.0	1.000
HYC20	Tech A	47.3	10.6	0.093	22.3	5.7	0.235
	Tech B	81.3	16.3	0.038	53.7	20.6	0.000
	Tech C	49.3	10.1	0.126	38.0	5.2	0.491

the lower limit of irreproducibility among institutions. Data from over 2,500 laboratories in the first quarter of 2004 CAP survey show that the between-laboratory coefficients of variation (CVs) for all methods range from 21% to 42% for RBC counts, and from 15% to 58% for WBC counts (7). However, because a standardized, plasma-like substance is used, the irreproducibility of measurements on patient-derived specimens with variable volumes and cell/element concentrations is likely to be greater.

In the current study we used a method to compare variability in manual cell counts obtained by two different procedures. According to the measures reported above, the candidate method performed slightly better than the current standard method in terms of reducing intra- and interoperator variability. However, the standard procedure for performing manual cell counts allowed the results to be reported in ranges. The selection of the midpoint as the result to use in the statistical analysis resulted in an overestimation of the precision in the standard method compared to the Hycor method. Additionally, effects that were due to differences in training periods between the two methods were not accounted for.

The technologists favored the use of the Hycor system for several reasons: Decanting to the required volume in which to resuspend the urine sediment was a faster and more accurate process, and allowed for batching. For CSF cell counting, the technologists preferred to use the Glasstic slides because of perceived time savings, ease of counting with the grid, and simple calculations. The disadvantage is the likelihood that 10 chambers on each slide will not be fully utilized, since the slides are not retained for more than one shift after they are used.

The Hycor system addresses several factors that contribute to imprecision in manual microscopic examinations of body fluids. Without increasing labor or the requirement for attention to detail, the Hycor method provides slightly better interrater comparisons than the current method used at this institution. Further work is needed to demonstrate the scientific relevance of this finding in the face of the high interrater variation observed with either method.

REFERENCES

1. Bertholf MF, Kao KJ. Disposable plastic and reusable glass hemacytometers for cell counts. Lab Med 1991;22:864–868.
2. Mahon CR, Smith LA. Standardization of the urine microscopic examination. Clin Lab Sci 1990;3:328–332.
3. Ottiger C, Regeniter A, Kochli HP, Huber AR. Standardized counting of particles in the urine: a comparison between flow cytometry, cell chamber counting and traditional sediment analysis. Schweiz Rundsh Med Prax 2004;93:15–21.
4. de Jonge R, Brower R, Smit M, et al. Automated counting of white blood cells in synovial fluid. Rheumatology 2004;43:170–173.
5. McGinley M, Wong LL, McBride JH, Rodgerson DO. Comparison of various methods for the enumeration of blood cells in urine. J Clin Lab Anal 1992;6:359–361.
6. Ottiger C, Huber AR. Quantitative urine particle analysis: integrative approach for the optimal combination of automation with UF-100 and microscopic review with Kova cell chamber. Clin Chem 2003;49:617–623.
7. College of American Pathologists. Surveys 2004 (hematology): participant summary. Northfield, IL: College of American Pathologists; 2004. p 1.