J Pathol Inform

Editor-in-Chief:
Anil V. Parwani , Pittsburgh, PA, USA
Liron Pantanowitz, Pittsburgh, PA, USA

OPEN ACCESS
HTML format

For entire Editorial Board visit : www.jpathinformatics.org/editorialboard.asp

Technical Note

Development and validation of an app-based cell counter for use in the clinical laboratory setting

Alexander C. Thurman, Jessica L. Davis, Max Jan[1], Charles E. McCulloch[2], Benjamin D. Buelow

Departments of Pathology and Laboratory Medicine, The University of California, San Francisco Medical Center, San Francisco, CA 94143-0102, [1]The University of California, San Francisco School of Medicine, San Franacisco, CA 94117, [2]Department of Epidemiology and Biostatistics, The University of California, San Francisco, San Francisco, CA 94107-1762, California, USA

E-mail: *Alexander C. Thurman - alexander.thurman@ucsf.edu
*Corresponding author

Received: 29 June 14 Accepted: 03 November 14 Published: 29 January 15

This article may be cited as:
Thurman AC, Davis JL, Jan M, McCulloch CE, Buelow BD. Development and validation of an app-based cell counter for use in the clinical laboratory setting. J Pathol Inform 2015;6:2.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2015/6/1/2/150252

Abstract

Introduction: For decades cellular differentials have been generated exclusively on analog tabletop cell counters. With the advent of tablet computers, digital cell counters – in the form of mobile applications ("apps") – now represent an alternative to analog devices. However, app-based counters have not been widely adopted by clinical laboratories, perhaps owing to a presumed decrease in count accuracy related to the lack of tactile feedback inherent in a touchscreen interface. We herein provide the first systematic evidence that digital cell counters function similarly to standard tabletop units. **Methods:** We developed an app-based cell counter optimized for use in the clinical laboratory setting. Paired counts of 188 peripheral blood smears and 62 bone marrow aspirate smears were performed using our app-based counter and a standard analog device. Differences between paired data sets were analyzed using the correlation coefficient, Student's *t*-test for paired samples and Bland–Altman plots. **Results:** All counts showed excellent agreement across all users and touch screen devices. With the exception of peripheral blood basophils (r = 0.684), differentials generated for the measured cell categories within the paired data sets were highly correlated (all $r \geq 0.899$). Results of paired *t*-tests did not reach statistical significance for any cell type (all $P > 0.05$), and Bland–Altman plots showed a narrow spread of the difference about the mean without evidence of significant outliers. **Conclusions:** Our analysis suggests that no systematic differences exist between cellular differentials obtained via app-based or tabletop counters and that agreement between these two methods is excellent.

Key words: App, apps, cell counter, digital, mobile, tablet, validation

Access this article online
Website: www.jpathinformatics.org
DOI: 10.4103/2153-3539.150252
Quick Response Code:

INTRODUCTION

The manual determination of cellular differentials in blood, bone marrow, and body fluids (including urine and pleural, peritoneal, and pericardial fluids) is a well-established diagnostic method useful in the evaluation of hematopoiesis, hematopoietic disorders, and infectious diseases.[1-3] Similar methods are also employed in establishing the grade of various solid tumors in tissue samples, a feature that often has a significant impact on prognosis.[4,5] For decades, such analyses have relied on the use of analog tabletop counting devices. In fact, despite the development of automated cell counters based on flow cytometric methods,[6,7] image recognition

software,[8] and other approaches, manual cell counting remains the gold standard for many of these applications.

Despite the importance of manual cell counting methods within the clinical laboratory, analog cell counters have remained essentially devoid of any significant improvement since their inception. As a result, these devices suffer from a number of substantial deficiencies. Not only are tabletop cell counters relatively expensive (often more than five hundred dollars for a single device), but this expense is compounded by the need to purchase units in sufficient quantities to accommodate each microscope used for cellular differentials. Furthermore, only 8 and 12 button configurations are widely available, frequently resulting in either a dearth or an excess of buttons than is required for a particular task. The buttons themselves are often closely spaced, and they cannot be distinguished without visual confirmation, increasing the risk for operational error. Importantly, analog counters also commonly lack editing capabilities, making error correction highly cumbersome. The inability to directly transfer these results to laboratory information systems increases the time needed to generate reports and introduces the possibility of typographical errors during data entry. Clearly, any erroneous results have the potential to jeopardize patient safety.

With the recent advent of tablet computers, digital cell counters – in the form of mobile applications ("apps") – have become an alternative to analog devices, having the potential to address many of their shortcomings. Relative to analog units, digital cell counters and their associated hardware are less expensive (at the time this article was written, Android-based tablets ranged in price from one to three hundred dollars, and cell counter apps were either free to download or, at most, five dollars per user) and offer customizable user-interface configurations for optimal performance of a particular task. As digital counters can be installed on any tablet computer or smartphone, they are significantly more portable than their tabletop counterparts. Importantly, data generated by digital counters can be edited during collection, and the final results are easily exported and/or digitally transferred without the need for manual entry of individual values.

Despite their reduced cost and broad availability, app-based cell counters have not been widely adopted by clinical laboratories. Although many factors likely impede the acceptance of digital cell counters within the laboratory setting, the most substantive concern – particularly in terms of its potential impact on patient care – is a perceived decrease in count accuracy, potentially owing to the lack of tactile feedback inherent to a touchscreen interface. Such concerns may have previously been understandable as validation data establishing the equivalency of digital cell counters to analog devices have not been reported. To address this deficiency, we herein provide the first systematic evidence that digital/app-based cell counters function equivalently to standard tabletop devices.

METHODS

App Design and Development

Our cell counter app was programmed using Java by computer science students at San Jose State University (San Jose, California: Amir Eibagi, Minh Dang, and Jake Karnes, under the supervision of Dr. Cay Horstmann) based on our design [Figure 1]. Following initial troubleshooting by our group, the app was distributed to faculty and staff at the University of California, San Francisco (UCSF) Medical Center for additional testing. Further improvements (incorporating end-user feedback from UCSF faculty and staff) were implemented through a similarly iterative process. After preliminary data had suggested that no systematic differences existed between digital and analog device-generated data, the app was made publicly available at: https://play.google.com/store/apps/details?id=edu.sjsu.hemepathcounter.

Case Selection

188 peripheral blood smears and 62 bone marrow aspirate smears (250 total samples) were selected, each for paired analysis using both the app-based counter and the tabletop unit. To ensure that a wide range of differential values for each cellular subset would be available for analysis, peripheral blood and bone marrow aspirate smears were selected from both the UCSF hematopathology teaching archives and from among

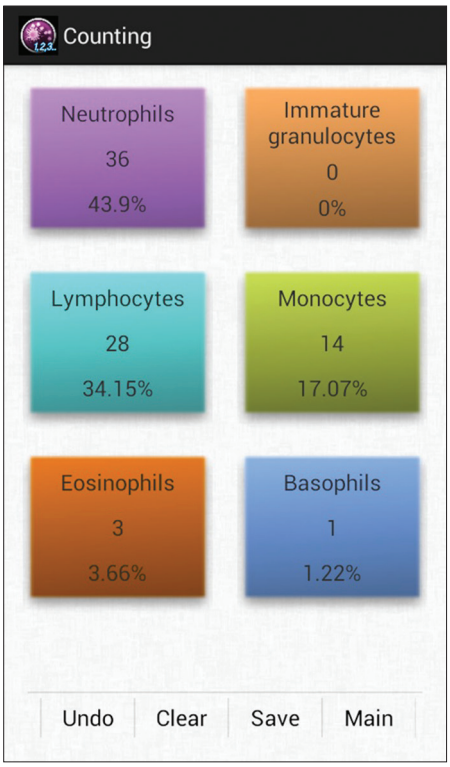

Figure 1: Screen capture demonstrating the user interface for a custom 6 key peripheral blood configuration. Additional images are available at https://play.google.com/store/apps/details?id=edu.sjsu.hemepathcounter

de-identified patient samples. All cases were selected randomly, and the reviewer was blinded to the final diagnosis as well as the reported differential count.

Cell Counting

Four authors (AT, BB, JD, and MJ) performed paired 100-cell differentials on a subset of the selected peripheral blood and/or bone marrow aspirate smears, once using our app-based cell counter (cell counter for android, installed on one of three mobile devices: Samsung Galaxy S3, Samsung Galaxy Tab 4 7.0, or Samsung Galaxy Note 10.1; Samsung Electronics), and once with a with a standard tabletop unit (Modulus Data Systems DIFFCOUNT). To streamline data acquisition and presentation, a total of six cell categories were quantified for each sample type: Neutrophils (including band forms), immature granulocytes, lymphocytes, monocytes, eosinophils, and basophils for peripheral blood smears; mature granulocytes (including band forms), immature granulocytes, lymphocytes, monocytes, erythroid precursors, and blasts for bone marrow aspirate smears.

Statistical Analysis

Differences between paired data sets were evaluated using the correlation coefficient, Student's *t*-test for paired samples, and Bland–Altman plots, according to the well-established method described by Bland and Altman for establishing agreement.[9,10] Data analysis was performed in Microsoft Excel (Microsoft Corporation). Primary data were aggregated across users [Figures 1-3] with individual user data analyzed separate [Tables 1 and 2].

RESULTS

Cases Selected for Analysis Span a Wide Range of Cellular Differentials

Inspection of the differential results generated by the tabletop unit confirmed that the selected material encompassed a wide range of cellular differentials associated with a variety of neoplastic and nonneoplastic disorders. These findings were considered to be representative of the range of differential results typically encountered in the

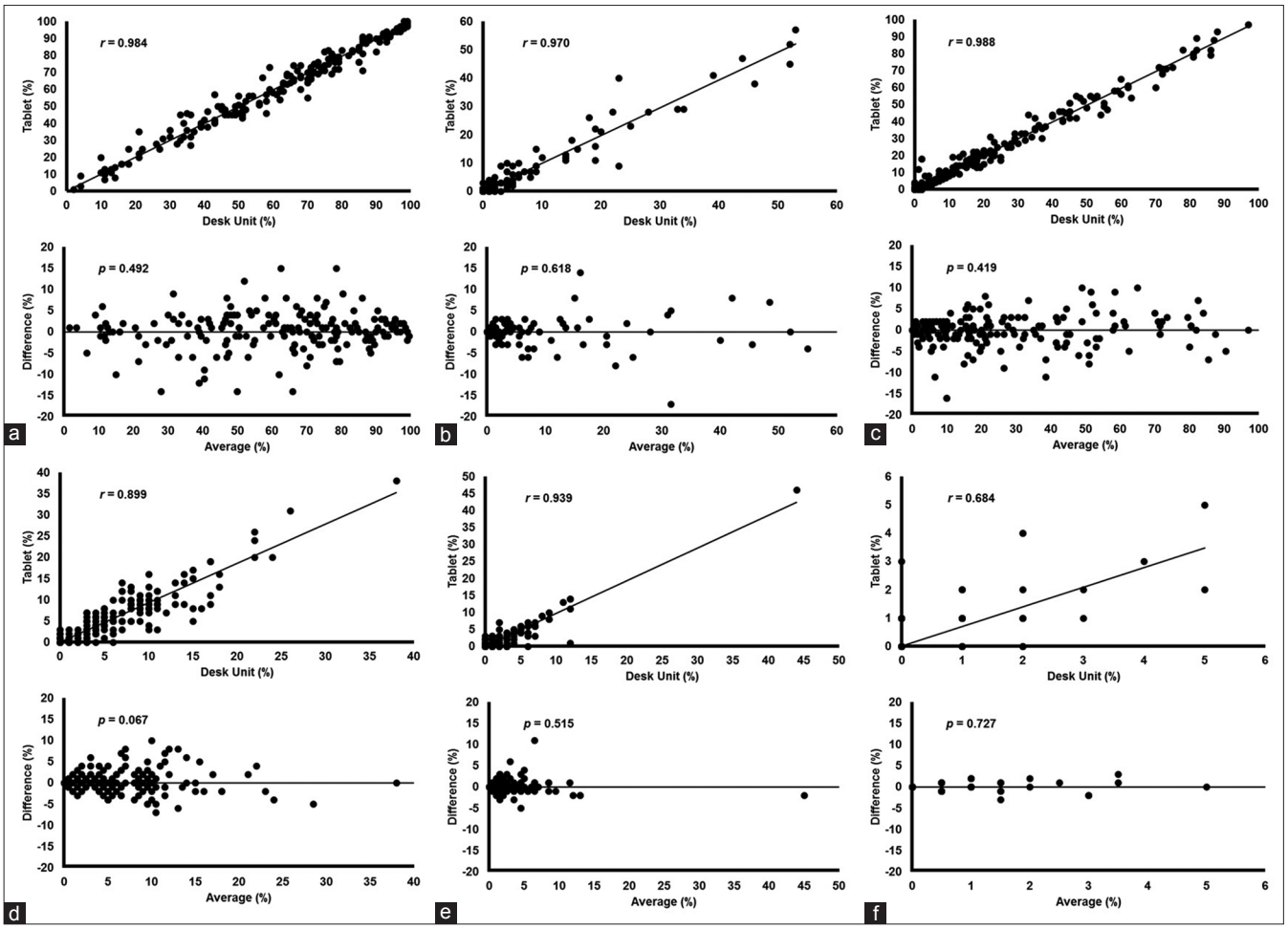

Figure 2: Scatterplots with lines of linear regression (upper) and Bland–Altman plots (lower) generated following paired counts of one hundred and 88 peripheral blood smears. The relative proportions of neutrophils, (a) immature granulocytes, (b) lymphocytes, (c) monocytes, (d) eosinophils and (e) basophils, (f) were measured. Correlation coefficients (*r*-values) and results of paired *t*-tests (*P* values) are shown for each data set

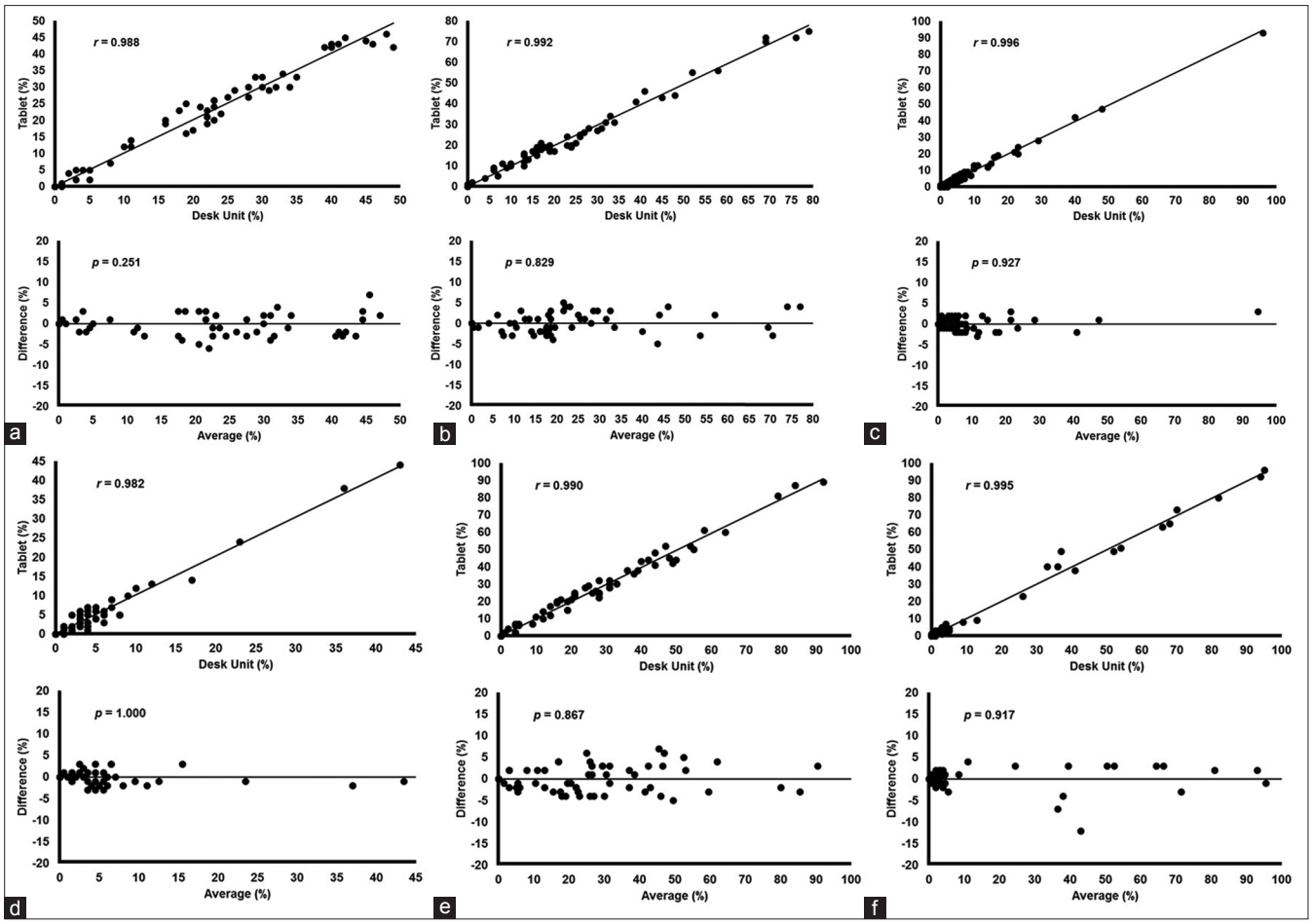

Figure 3: Scatterplots with lines of linear regression (upper) and Bland–Altman plots (lower) generated following paired counts of 62 bone marrow aspirate smears. The relative proportions of mature granulocytes (a), immature granulocytes (b), lymphocytes (c), monocytes (d), erythroid precursors (e), and blasts (f) were measured. Correlation coefficients (*r*-values) and results of paired *t*-tests (*P* values) are shown for each data set

clinical laboratories at UCSF Medical Center, an academic tertiary care hospital. As expected, a significant proportion of the patient samples taken from the hematopathology service were, upon unblinding, found to have no significant pathologic abnormalities; these smears exhibited differential counts within established age-matched references ranges by both the standard and app-based methods.

Manual Differentials Generated on Peripheral Blood Smears Using Tabletop Versus App-Based Cell Counters Show No Statistically Significant Differences

Having established that our selected cases spanned a wide range of values for each cellular subset under investigation, we turned our attention to the evaluation of similarity between the paired analog- and app-based counts of peripheral blood smears. With the exception of basophils ($r = 0.684$), peripheral blood differential cell counts were found to be highly correlated within each cell category (mean $r = 0.956$; Figure 2, upper panels). Despite their relatively low correlation coefficient, basophils showed a high level of agreement between values generated by the two methods, and the values from the analog- and app-based methods never differed by an absolute value >3%. Bland–Altman plots uniformly exhibited a narrow spread about the mean difference without evidence of significant outliers [Figure 2, lower panels]. Paired *t*-tests demonstrated *P* values which did not reach statistical significance for any cell type (all $P > 0.05$, including basophils), arguing against the presence of any systematic deviation from the tabletop unit by the app-based counter. To assess for inter-user variability in count accuracy, we also calculated *r* and *P* values separately for each user [Table 1]. Among 24 comparisons, only a single nominally statistically significant result was observed (user MJ, lymphocytes, $P = 0.028$).

Manual Differentials Generated on Bone Marrow Aspirate Smears Using Tabletop Versus App-Based Cell Counters Show No Statistically Significant Differences

Given the consistent similarity of differentials generated by the two methods in peripheral blood, we proceeded with the same analysis in bone marrow aspirate smears. Comparable results were obtained for bone marrow

Table 1: Individualized correlation coefficients (*r*-values) and results of paired *t*-tests (*P* values) for all six peripheral blood cell categories

User	AT (*n*=50)	BB (*n*=50)	JD (*n*=38)	MJ (*n*=50)
Neutrophils				
r	0.995	0.983	0.986	0.963
p	0.083	0.400	0.741	0.248
Immature granulocytes				
r	0.988	0.952	0.997	0.956
p	0.837	0.367	0.850	0.807
Lymphocytes				
r	0.996	0.986	0.988	0.962
p	0.893	0.159	0.835	0.028
Monocytes				
r	0.974	0.929	0.879	0.678
p	0.330	0.123	0.761	0.066
Eosinophils				
r	0.994	0.938	0.820	0.663
p	0.105	0.242	0.385	0.616
Basophils				
r	0.519	0.938	0.615	0.436
p	0.371	1.000	0.676	0.735

Table 2: Individualized correlation coefficients (*r*-values) and results of paired *t*-tests (*P* values) for all six bone marrow aspirate cell categories

User	AT (*n*=50)	JD (*n*=12)
Mature granulocytes		
r	0.981	N/A
p	0.252	N/A
Immature granulocytes		
r	0.977	0.998
p	0.810	1.000
Lymphocytes		
r	0.989	1.000
p	0.699	0.175
Monocytes		
r	0.981	N/A
p	1.000	N/A
Erythroid precursors		
r	0.987	0.997
p	0.859	1.000
Blasts		
r	0.998	0.987
p	0.157	0.279

aspirate differential cell counts: The two data sets were highly correlated (mean $r = 0.990$; Figure 3, upper panels); showed a tight, random dispersion about the mean difference on Bland–Altman plots [Figure 3, lower panels]; and were not found to significantly differ by paired *t*-tests (all $P > 0.05$). No substantial differences between users were observed [Table 2].

CONCLUSIONS

To our knowledge, a formal evaluation of the concordance between digital/app-based cell counters and analog tabletop units – considered to be the gold standard instrument for use in the determination of manual cellular differentials – has never been performed. Our work addresses this deficiency by providing the first evidence that cellular differentials generated by an app-based cell counter do not significantly differ from those of an analog device. This conclusion is supported by the tight correlation ($r \geq 0.899$ for all measured cell categories, excluding peripheral blood basophils) and lack of statistically significant differences observed between paired data sets (aggregated $P > 0.05$ for all measured cell categories) as well as by Bland–Altman analysis, which demonstrated a narrow spread of the difference about the mean without evidence of systematic bias between paired results. Analysis of individual user-device pairs showed similar results, save for a single nominally significant difference observed for one user (MJ) within one cell category (peripheral blood lymphocytes). However, as this value was not corrected for multiple comparisons (according to the method of Bonferroni),[11] we do not consider it to be a contradictory datum.

Although statistical measures of difference (such as the Chi-squared test, Student's *t*-test, etc.) are familiar to pathologists and laboratorians alike, the approach to demonstrating similarity between data sets is often less well-defined. Some studies rely exclusively on the correlation coefficient to establish statistical similarity; however,[12-17] the combination of the correlation coefficient with Bland–Altman analysis is a more stringent approach to demonstrating functional equivalence between a novel and an established method and is subject to fewer sources of error.[9,10,12,13,16] The superiority of a combined approach is evidenced by our results for peripheral blood basophils: Although only moderate correlation was observed for this cell category ($r = 0.684$; Figure 2f, upper panel), the accompanying Bland–Altman analysis [Figure 2f, lower panel] demonstrated that this lack of correlation does not represent a systematic discrepancy between the two methods. The nonsignificant difference demonstrated by the *t*-test further argues against any systematic deviation within this cell category. We therefore hypothesize that this finding is attributable to the enhanced effect of background variance on this low-frequency event.

Taken together, the findings suggest that our app-based cell counter is functionally similar to an analog tabletop unit for the purpose of generating manual cellular differentials – a conclusion which is based on data

generated by four different users on three separate touchscreen devices. Accordingly, our findings do not support the hypothesis that the loss of tactile feedback inherent to a digital/app-based cell counter leads to a decrease in count accuracy. Our app provides alternative modes of user feedback through differential coloring and distinctive sound effects for each button, which may compensate for the lack of button tactility (future versions of our app may also include haptic/vibrational feedback with each keystroke). Additionally, the screen sizes of the devices used in our study ranged from 4.8 inches (Samsung Galaxy S3) to 10.1 inches (Samsung Galaxy Note 10.1), suggesting that touchscreen dimensions do not significantly impact count accuracy. Nevertheless, future studies to formally examine the effect of screen size, resolution, and/or orientation (horizontal versus vertical) are warranted to confirm this impression. Although our results cannot be directly extrapolated to all cell counting apps and/or touchscreen devices, our study provides a framework for the validation of other software/hardware combinations.

Unlike digital counters, tabletop devices are ubiquitous, durable, require little technical support, and are unlikely to be stolen. Thus, it seems doubtful that app-based counters will entirely supplant analog devices, at least in the immediate future. However, given the low total cost of app-based counters (100 dollars at minimum, including hardware) and their ease of validation (as laid out in this work), digital cell counters may represent an appealing alternative for small-sized and/or newly established clinical laboratories. Although not formally investigated in this study, subjectively, none of the users noted an increase in the time required to perform cellular differentials when using our app-based counter. Nonetheless, the potential impact of digital cell counters on laboratory workflow is a topic in need of further study.

In summary, we have presented here the first statistical evidence that digital/app-based cell counters are functionally similar to analog tabletop devices in the generation of manual cellular differentials.

ACKNOWLEDGMENTS

We wish to acknowledge Minh Dang, Amir Eibagi, Jake Karnes, and Dr. Cay Horstmann (San Jose State University Computer Science) for their programming assistance and Garin Mazaika for his artistic contributions to our app.

REFERENCES

1. Houwen B. The differential cell count. Lab Hematol 2001;7:89-100.
2. Novak RW. The beleaguered band count. Clin Lab Med 1993;13:895-903.
3. Tatsumi N, Pierre RV. Automated image processing. Past, present, and future of blood cell morphology identification. Clin Lab Med 2002;22:299-315, viii.
4. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. Br J Cancer 1957;11:359-77.
5. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: Experience from a large study with long-term follow-up. Histopathology 1991;19:403-10.
6. Buttarello M, Plebani M. Automated blood cell counts: State of the art. Am J Clin Pathol 2008;130:104-16.
7. Meintker L, Ringwald J, Rauh M, Krause SW. Comparison of automated differential blood cell counts from Abbott Sapphire, Siemens Advia 120, Beckman Coulter DxH 800, and Sysmex XE-2100 in normal and pathologic samples. Am J Clin Pathol 2013;139:641-50.
8. Briggs C, Longair I, Slavik M, Thwaite K, Mills R, Thavaraja V, *et al.* Can automated blood film analysis replace the manual differential? An evaluation of the CellaVision DM96 automated image analysis system. Int J Lab Hematol 2009;31:48-60.
9. Altman DG, Bland JM. Measurement in medicine: The analysis of method comparison studies. Statistician 1983;32:307-17.
10. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.
11. Bland JM, Altman DG. Multiple significance tests: The Bonferroni method. BMJ 1995;310:170.
12. Akahira-Azuma M, Yonemoto N, Ganzorig B, Mori R, Hosokawa S, Matsushita T, *et al.* Validation of a transcutaneous bilirubin meter in Mongolian neonates: Comparison with total serum bilirubin. BMC Pediatr 2013;13:151.
13. Akyalcin S, Cozad BE, English JD, Colville CD, Laman S. Diagnostic accuracy of impression-free digital models. Am J Orthod Dentofacial Orthop 2013;144:916-22.
14. Chen Z, Shin D, Chen S, Mikhail K, Hadass O, Tomlison BN, *et al.* Histological quantitation of brain injury using whole slide imaging: A pilot validation study in mice. PLoS One 2014;9:e92133.
15. Lobmaier SM, Cruz-Lemini M, Valenzuela-Alcaraz B, Ortiz JU, Martinez JM, Gratacos E, *et al.* Influence of equipment and settings on myocardial performance index repeatability and definition of settings to achieve optimal reproducibility. Ultrasound Obstet Gynecol 2014;43:632-9.
16. Solomon C, Baryshnikova E, Tripodi A, Schlimp CJ, Schöchl H, Cadamuro J, *et al.* Fibrinogen measurement in cardiac surgery with cardiopulmonary bypass: Analysis of repeatability and agreement of Clauss method within and between six different laboratories. Thromb Haemost 2014;112:109-17.
17. van Hooijdonk RT, Winters T, Fischer JC, van Dongen-Lases EC, Krinsley JS, Preiser JC, *et al.* Accuracy and limitations of continuous glucose monitoring using spectroscopy in critically ill patients. Ann Intensive Care 2014;4:8.