Filter Paper-collected Blood Lead Testing in Children

To the Editor

The Editorial “Filter Paper Lead Testing” by Moyer et al. (1) repeats previously raised hypothetical and unsubstantiated claims on filter paper (FP collection-based) blood lead testing (1–4). All the theoretical charges were appropriately rebutted in the past based on published evidence (5–9). The peer-reviewed scientific literature supports the clinical usefulness of FP-collected blood samples in childhood lead poisoning screenings (10–19).

The Editorial’s authors (1) used references that do not support their statements: one of the references lacks entirely the topic of lead testing for which it was used. Worse yet, available references were omitted from the Editorial that support FP collection-based methods. The following comments and references attempt to correct these mistakes and provide accurate documentation from the scientific literature on FP.

Reference 9 in the Editorial (14) is cited for support of “low extraction recoveries” of lead from FP. This study is based on a method that is a direct analysis that does not require an extraction (14). The same reference is cited to support the statements that contamination occurs as a result of “high and variable lead content of the FP or by contamination encountered in the sampling process”. The reference did not address FP blank lead concentration issues or that contamination caused any unusual problems during sample collection.

The wrong volume and page numbers are cited for reference 8 in the Editorial (18). The work of Cernik (18) and O’Broin et al. (20) are used to support “nonreproducibility of blood spotting because of differences in hematocrit [Hct] and hemoglobin [Hb]”. (1). Cernik (18) studied neither Hct nor Hb. His article investigated the effects of paper punch size on reproducibility of blood lead measurement on FP. He reported that punches <4.0 mm in diameter caused variation in reproducibility but that punches >4.0 mm correlated excellently with volumetric measurements. Laboratories in the 1990s use 6.0 mm (¼-inch) or 4.5 mm (⅛-inch) punch sizes with excellent reproducibility (11–17).

Similarly, the study by O’Broin et al. (20) (reference 11 in the Editorial) did not examine blood lead measurement on FP. Their conclusion was that: “Specific studies on the distribution in a paper matrix of all analytes of interest are desirable”. The cited study (20) is applicable only to ferritin, cobalamin, and folates. Lead was not an analyte in their study. However, an earlier study by O’Broin (21) has some relevance to FP measurement of heavy metals that distribute mainly into erythrocytes. Lead is 94% distributed into erythrocytes (22). Two analytes were examined from FP blood samples for accuracy at a wide range of Hct values (19–63%). [125I]-labeled thyroxine was seriously influenced, but chromium-51 was hardly affected by the large Hct variation. O’Broin’s conclusion was that analytes that distribute into plasma are affected by the wide variation of Hcts but analytes in erythrocytes are not significantly influenced (21).

O’Broin’s findings (21) were specifically confirmed for blood lead measurement on FP in recent studies that measured Hb or Hct along with FP and liquid blood lead concentrations (13, 17). Schonfeld et al. (13) in a study of >1500 children reported an average Hb of 120 g/L [12 g/dL, Hb (range, 8.3–18.7 g/dL)]. They stated, “Both fingerstick lead screening tests had excellent correlation with the confirmatory VPb [venous lead]”. The authors concluded that “screening for lead poisoning in urban areas is feasible by direct measurement of fingerstick CPb [capillary/Microtainer lead] or FPb [FP lead]”.

Yee and Holtrop (17) specifically studied the Hct effects on blood lead quantification. Fingerstick FP blood and simultaneous venous specimens were collected (n = 59) and analyzed by graphite furnace atomic absorption spectrometry. Hct values ranged between 27% and 44%. The variation did not affect the excellent correlation between FP and venous blood lead results.

The Editorial (1) implies that one cannot achieve accurate results using FP collected samples because of spotting difficulties. On the precision of FP blood lead measurement, Morrow et al. (23) of the CDC stated, “Filter paper had excellent correlation to target values (r = .998) with precision similar to that of liquid samples.”

The Editorial (1) states that “Bias may be introduced by high and variable lead content of the FP...”. In 1992, Wang et al. (12) reported, “A random sample of 30 blank filter paper spots from our neonatal sampling sheet was tested for lead. No detectable levels of lead were found in any of the blank filter paper spots tested”. Morrow et al. (23) stated, “lead in the filter paper was generally low (<2 mcg/dL) during the last 5 years (1988–1993). Therefore these lots appear to be suitable for measuring low levels of lead”.

Yee and Holtrop (17) in 1997 tested Schleicher and Schuell 903 FP (S&S 903). They reported that “filter paper blank values were <2 mcg/dL [20 mcg/L] with the use of aqueous standards without correction and 0.1 mcg/dL with a paper blank correction; and during the collection of 163 samples from 6 different urban sites over a period of 1 year no environmental lead contamination was encountered, based upon the good agreement between all of the paired FP and venous lead results”.

Verebey (6) stated, “in Leadtech as part of the QA and QC program all new lots of filter paper (Schleicher & Schuell 903 filter paper) are tested for background lead. Of the 156 lots examined (n = 780 tests) in the past 4 years, not one blank filter paper was found with blood lead equivalent ≥1 µg/dL (10 µg/L).”

Wong et al. (24) (reference 6 in the Editorial) did report high blank FP lead values in an abstract, which was never published in a peer-reviewed scientific communication. Their methodological problem may have led to the very high background
blank filter paper results. Yet this was the sole study chosen by the authors of the Editorial to support their statement. Contrary to the findings of Wong et al. (24), four different laboratories found the blank FP (S&S 903) values low and acceptable for blood lead screening procedures (6, 12, 17, 23).

The Editorial (1) suggests that contamination of blank FP collectors caused imprecision during the first two cycles (trial period) in the Wisconsin State pilot proficiency testing (PT) program (19). The suggestion is not feasible because all laboratories participating successfully used the same (S&S 903) filter papers routinely in their practice and in the PT program. The initial performance variances were likely attributable to calibration and methodological adjustments to the new PT program.

Reading the Editorial (1), one gets the notion that FP has not been successfully used in large clinical studies and that, therefore, we have no idea how this “new procedure” would perform in the “field”. Thus, the authors of the Editorial suggest that FP collection is uniquely predisposed to “preanalytical contamination” and analytical variations. The preanalytical contamination and analytical concerns are ones that any competent analyst must deal with in the performance of trace metal analysis. There is no need to insinuate that these concerns are unique or apply only to FP lead testing. In two of the worst examples with Microtainer capillary collectors, as many as 68% and 70% of positive results were false positives attributed to preanalytical contamination (25, 26). In Texas, Medicaid released a provider alert (Bulletin No. 115) indicating that a batch of blood collection tubes used for fingerstick collection were contaminated with the equivalent of as much as 250 µg/L lead.

It is clear that contamination theoretically can cause preanalytical problems with FP collection as well as other capillary blood collections, but it was not noted as a unique problem in several major studies that used FP in the 1990s.

Another criticism in the Editorial (1) claimed that the good performance of FP methods in the Wisconsin State Pilot PT program (19) was attributable to the controlled “artificial environment” and that it would not be expected to perform similarly well in a day-to-day practice. This statement is purely hypothetical. FP has been used successfully in several large studies in clinical settings in the United States and other countries without any unusual analytical or contamination problems (10, 12–17, 27, 28): Schonfeld et al. (13) studied >1500 children in Connecticut; Schlenker et al. (27) studied 5350 children in Salt Lake City County; Pfitzner et al. (28) studied 218 children in Nigeria; and Verebey et al. (5) studied >94 964 children across the United States. The average blood lead concentrations of children in the Verebey study (5) was 30 µg/L (3 µg/dL), closely matching the CDC’s NHANES national average (29). Piomelli et al. (10) studied 106 people in the Himalayas, using FP-collected blood. The average blood lead in the Himalayas study was also 30 µg/L (3 µg/dL). These low concentrations (5, 10) could not be measured if the FP itself had “high and variable lead content”, as suggested by the Editorial (1) and Wong et al. (24).

The clinical significance of screening tests is to identify subjects with blood lead concentrations >100 µg/L (10 µg/dL) and to avoid false-negative tests. Yee and Holtrop (17) compared 119 field samples in which venous blood and fingerstick FP collections were made simultaneously. At the 100 µg/L (10 µg/dL) cutoff, the false-positive and false-negative rates were 0.8% and 1.7%, respectively. The correlation was y (FP) = 1.37 + 0.83x (venous). Thus, the FP-based screening tests do fulfill the basic requirements needed to identify children with increased blood lead concentrations.

The NHANES statistical data (29) indicate that in the United States, ~95.7% of the screening test results are expected to be <100 µg/L (10 µg/dL). Each clinical laboratory screening report states that a higher lead result must be confirmed by a venous specimen. No clinical treatment decisions are made based on an increased screening result alone.

The excellent precision shown by the FP screening test in numerous field studies (12, 14, 16, 17) and in the new Wisconsin State PT (19) program is comparable in performance to confirmatory methods. Therefore, the FP-based methods are more than adequate to serve the screening functions they are intended to perform. The advantages of FP blood collection is well stated in the Editorial (1): “(a) a relatively small amount of blood is required; (b) blood collection via fingerstick is generally easier and less traumatic to children; (c) collection can be performed by a relatively untrained collector; and (d) appropriately prepared FP specimens are stable and can be easily transported to the laboratory”.

References
The authors of the editorial cited above respond:

To the Editor:

We appreciate the opportunity to comment on the letter of Verebey regarding our Editorial (1) on filter paper lead (FPb) testing. The FPb technique has been critically assessed before (2) with a similar response (3).

We take exception to Verebey’s statement that “All the theoretical charges were appropriately rebutted in the past based on published evidence”. The citations noted by Verebey as published evidence are letters to editors; letters may not undergo the same degree of peer review for scientific content as would a scientific article, and they often represent the opinion of the author rather than scientific evidence. Although some peer-reviewed articles have been published that support the use of the FPb technique, we feel that the technique needs further improvement to meet CLIA performance standards.

Our Editorial raised concern about the accuracy and utility of FPb testing (1). The arguments we used are neither hypothetical nor unsubstantiated. The latest work of Stanton et al. (4) confirms our position. Our view is that the clinical laboratory should approach this technology with caution until it can be shown that the technique is capable of analytical performance equivalent to blood lead testing, the accepted procedure for diagnosis of lead toxicity defined in the 1991 guidelines from the United States Center for Disease Control and Prevention. Stanton et al. (4) reported that even when all laboratories used the same Schleicher & Schuell No. 903 paper and laboratory-prepared blood spots on that paper, three of the six participating laboratories failed minimum CLIA standards for blind proficiency testing.

Our Editorial contained two errors in the references. In the first, the work by Verebey et al. (5) should not have appeared in the discussion of “low extraction recoveries” or filter paper blanks; the works by Yee and Holtrop (6) and Wong et al. (7) were adequate to support this point. Second, the volume and page numbers for the Cernik citation were in error (8). We regret these errors. It is our opinion that all other references used in our original Editorial are correct and, more importantly, appropriate.

The rebuttal from Verebey addresses the specific issues of hematocrit-hemoglobin, filter paper blanks, collection device contamination, and accuracy requirements. We will respond to each issue briefly as they have been addressed many times before.

Verebey states that we are incorrect in using the references to Cernik (8) and O’Brien et al. (9) in our discussion of effect of hematocrit and hemoglobin on the reproducibility of blood spotting. Although Cernik (8) does not specifically measure hemoglobin or hematocrit, he certainly addresses the issues of the “build-up of red cells” at the periphery of the blood spot and states that “variation in spread is the most important factor involved and it must be reduced to a minimum”. Furthermore, Cernik (8) states that “when the sample of blood applied was between 0.02 and 0.04 mL, maximum but not complete uniformity of spread for blood spot size was achieved”. To determine the red cell migration the blood was tagged with 51Cr. Cernik (8) does state that increasing the punch disc size produced better precision for the lead measurement, which is consistent with comments made by Verebey.

We recognize that O’Brien et al. (9) did not study blood lead in this particular work; however, their detailed study of dried blood is pertinent, particularly when they state that “the blood in each DBS (dried blood spot) punch (n = 234) was greatly influenced by both sample hematocrit (r = 0.63) and hemoglobin concentration (r = 0.63). The volume in identical punches (n = 57) also differed significantly when measured independently using either 1251 human serum albumin or hemoglobin relative to the original sample”. Furthermore, the authors state that “DBS punch volumes should be predeter-
made only for a subset of 35 samples for which hematocrit, FPb, and venous blood lead results were presented. The least-squares equation for a plot of FPb vs venous lead was: 
\[ y = 0.79x + 0.8 \text{ \mu g/dL} \] with a correlation of 0.98. Even with a very tightly controlled hematocrit (range, 27–44%) and considering that this was a highly selected subset of the study population, results by the FPb technique were only ~79% of those measured in venous blood.

Recognizing a point made by Verebey, Schonfeld et al. (12) were able to show good results with the FPb technique for blood samples with hemoglobin concentrations of 83–187 g/L (8.3–18.7 g/dL). No correlations of hematocrit with results were presented, but excellent agreement of both capillary blood lead and filter paper blood lead with venous blood lead was observed. We have never said that the FPb technique does not work. We have only advised caution in its application.

On the subject of FPb blank concentrations, many laboratories (including our own) agree that Schleicher & Schuell No. 903 paper contains very low lead. Indeed, our Editorial notes that a “radical improvement in imprecision was observed” when the filter paper substrate was switched to this single source (1). It is conjecture and hypothesis on Verebey’s part, however, to suggest that the high and variable results reported by Wong et al. (7) were attributable to “methodological problems”.

Although lead content of filter paper has been determined to be low by most of the reporting laboratories, the bias of the technique when compared with standard analytical lead techniques has not. In reviewing the data compiled by Verebey et al. (5), Wong et al. (7), Schonfeld et al. (12), Yee and Holtrop (6), and Stanton et al. (4), it can be seen that although low concentrations of lead have been found in the FPb paper substrate, significant positive and negative biases of FPb results are evident. Four of the laboratories participating in the study by Stanton et al. (4) had small biases (~0.77 to 0.55 \text{ \mu g/dL Pb}), but two had a much larger bias (2.7 \text{ \mu g/dL Pb}). Other investigators (5, 7, 12) reported unacceptably large bias (~1.6 to 5.4 \text{ \mu g/dL Pb}). Schonfeld et al. (12) observed excellent agreement of both capillary blood lead (\( y = 0.99x \)) and FPb (\( y = 0.92x \)) with venous blood lead values but noted that the bias for capillary blood lead was only 0.8 \text{ \mu g/dL Pb}, whereas the bias for FPb was 5.4 \text{ \mu g/dL Pb}.

Verebey suggests that the FPb technique is not predisposed to more preanalytical contamination than the capillary collector technique. Regarding the latter, he notes studies in which 68% and 70% of positive results were false positives. In one of the cited studies (13), however, the authors state, “Although the false-positive rate of the capillary lead screening test was 70% (21/30) in this setting, only 2% of the total sample had a false-positive screening test (an average of fewer than one false-positive per month per practice)”. In this study, 1085 children were surveyed. Of the positive (>15 \text{ \mu g/dL Pb}) capillary samples recorded, follow-up venous analysis indicated that only 9 of the 30 cases were positive. In fact, the false-positive rate for capillary blood taken for lead analysis is extremely low. In studies by Johnson et al. (14), Parsons et al. (15), and Schonfeld et al. (12), capillary collection into Microtainer collection devices added virtually no bias to the lead determination when compared with venous blood lead measurement. Verebey et al. (5) conceded that “capillary sampling is a viable alternative to venipuncture for lead poisoning screening in young children”.

The remaining issue to be addressed is lead recovery by the FPb technique. We acknowledge that any number of studies can be cited to show that this technique works, but demonstration of accuracy requires either comparison with an accepted reference method or analysis of reference materials. The FPb technique has been evaluated by both open
proficiency testing (4, 16) and by reference method comparison (5–7, 12). In all of the reference method comparison studies that directly compare the FPb technique to either graphite furnace (5, 6, 12) or ICP-MS (7), the slopes were significantly less than 1 (0.88, 0.83, and 0.86, respectively). The only exception is the work of Schonfeld et al. (12) ($y = 0.92x + 5.4 \mu g/dL$), where the slope was >0.9 but significant bias was evident. In a proficiency study of blind vs overt lead testing by Reilly et al. (16), the FPb technique produced results from blind samples divergent from the overt samples by >35%. In the open study by Stanton et al. (4), three laboratories produced slopes near 1 with very small bias. These three laboratories, exhibiting slopes of 0.94, 1.07, and 1.01, successfully met the CLIA requirements for blind proficiency testing. But the performance of three other participants did not meet CLIA standards. We cannot endorse a technology that allows 50% of participating laboratories to fail performance standards.

In summary, we restate the original concern expressed in our Editorial—can the FPb technology be implemented in the field to provide clinically useful results? We suggest that this question will be laid to rest only by a blinded study involving samples collected by healthcare workers in the field.

References

Thomas P. Moyer1*
David E. Nixon1
Owen K. Ash2

1 Department of Laboratory Medicine and Pathology
Mayo Clinic
Rochester, MN 55905
2 Department of Pathology
University of Utah/ARUP Laboratories
Salt Lake City, UT 84108

*Author for correspondence.

More on Filter Paper Lead Testing

To The Editor:

The recent Editorial on filter paper (FP) lead testing (1) provides a needed discussion of preanalytical variables in FP testing. These issues were not investigated in our report on FP proficiency testing results (2) that provided the basis for the Editorial. We would like to comment on two statements made in the Editorial.

First, the authors assert that improvements in participant performance that were observed following an initial testing period are attributable to a change in the source of FP sheets during the study (1). We believe a close examination of the data does not support this hypothesis. To review, our study (2) reported FP data from the period October 1997 through June 1998. FP specimens were grouped in monthly testing events, each consisting of five specimens. Events in October and November 1997 were considered preliminary, and the data from these events showed wide scatter. Beginning with the December 1997 testing event, the FP results were evaluated using CLIA acceptability criteria for blood lead. We reported that a substantial improvement in accuracy coincided with the institution of this formal evaluation and continued throughout the rest of the study period. We attributed this improvement to changes made by the participating laboratories during or immediately following the preliminary events (2).

As the Editorial correctly notes, beginning with the January 1998 test event, the FP sheets used were from a single lot obtained by the Wisconsin State Laboratory of Hygiene for study purposes. Before this, the FP sheets had been provided by the individual participating laboratories. The Editorial then suggests that the improvement in performance reflected the change in FP source (1). However, this hypothesis does not reconcile data from the December 1997 testing event (the first formally evaluated event), which were provided on FP sheets supplied by the participants, that showed the improved performance consistent with the later period (2). All of the FP sheets used in the study “shared the common characteristics of Schleicher & Schuell no. 903 composition, designated spotting and handling areas, and prior suitability testing by participants” (2). These shared characteristics are also contrary to the hypothesis stated in the Editorial. In addition, we received reports from some participants that they had made changes to their procedures...