The effect of hematocrit on metabolomic profiles of dried blood spot cards

Scott D. McCulloch[1], Edward D. Karoly[2], Matthew W. Mitchell[3], Anne M. Evans[4], Kelli Goodman[5]

[1] Senior Study Director, Discovery and Translational Sciences; [2] Senior Director, Discovery and Translational Sciences; [3] Senior Director of Statistics, Data Science; [4] Director, Research and Development; [5] Staff Scientist, Research and Development

Metabolon Inc, Morrisville, North Carolina, USA
December 15, 2021

Introduction

One key difference of dried blood spot (DBS) cards in comparison to plasma or serum is the inclusion of the cell component of whole blood (*i.e.*, hematocrit; HCT). Under normal circumstances these will comprise ~35-50% of total volume in adults (1, 2). The overwhelming majority of these cells are erythrocytes (red blood cells; RBC), with ~1% or less of volume taken up by leukocytes and platelets. The remaining ~50-65% is the liquid portion that remains after uncoagulated blood cells have been removed (plasma) or after the blood has coagulated (serum) and been separated. Various factors can affect the hematocrit level, including but not limited to age, anemia, dehydration, altitude, disease state, and drug treatment, with values ranging from 20-70% being observed (**Figure 1**) (3-6). Thus, the range of RBC proportion in a standard DBS punch being profiled using metabolomics is potentially large, and differences in HCT can affect DBS in multiple ways. HCT can impart a dilution effect on biochemicals that are more concentrated or solely detected in either the liquid portion or RBCs (*i.e.*, higher HCT causes greater dilution of liquid biochemicals and greater concentration of RBC-bound biochemicals). Additionally, as HCT increases and the proportion of liquid decreases, the viscosity of whole blood increases. This causes slower spread on the filter paper when spotted and ultimately results in smaller spots that contain a greater volume of blood per sub-punch area. It is therefore of interest to investigate what effect the hematocrit has on the overall metabolite profile when analyzing DBS cards. Metabolon has validated the use of Whatman 903 cards as a viable matrix for global untargeted metabolomics using our Precision Metabolomics™ platform (HD4). In doing so we have analyzed the effects of HCT on DBS profiles using internal studies. These analyses guide our recommendations for designing DBS-based research studies, the results of which are presented below.

Study Design and results

All study samples were processed using the standard extraction protocol developed by Metabolon for DBS cards, with all proceeding steps following Metabolon's normal workflow for sample analysis (UPLC-MS/MS, automated peak identification, data curation, statistics) (7).

Preliminary biochemical screening

We first wanted to identify biochemicals that may be impacted by varying HCT levels under controlled conditions. Experimental blood samples were manually adjusted by adding or removing plasma to

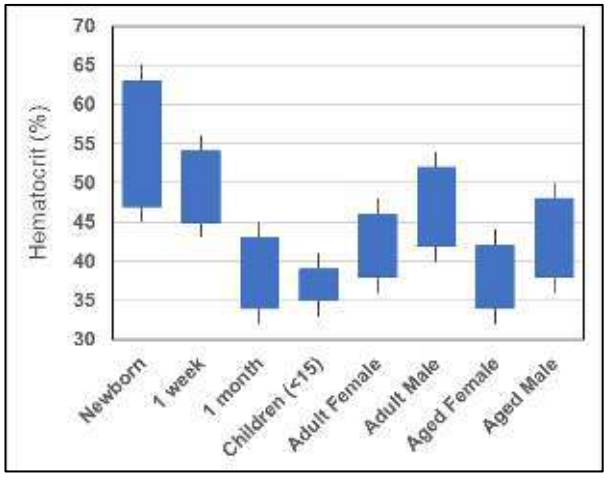

Figure 1 - Ranges of hematocrit levels typically found at various ages of life in human populations.

Metabolon, Inc.
617 Davis Drive, Suite 100, Morrisville, NC 27560

ENLIGHTENING LIFE™

achieve HCT of 30% and 60%, which were compared to an unadjusted control sample that was ~50% HCT. Median peak areas were used to compare each experimental HCT to the control. The *relative percent difference* (RPD) was calculated as the absolute difference divided by the average and was used to determine HCT effects for all biochemicals with 100% fill in all groups. The term "HCT Effect" here is defined as RPD >25% with a $p<0.05$. The coefficient of variance (CV) of technical replicates was also assessed. *This experiment does not demonstrate the effects in real-world studies where each group is expected to have a range of HCT levels.*

Affected	Outcome	% Diff 65:25	CV (25 – 65)	Correlation (r)	RPD		
					25:65 (low:high)	25:45 (low:mid)	65:45 (high:mid)
No	No Effect	< 20%	< 15%	*	< 20%		
Yes	Minor	25 – 50%	< 20%	> 0.7 (±)	25 – 40%	< 25% in **both**	
	Moderate					> 25% in **one** (< 15% in the other)	
	HCT Effect	> 35%	> 20%	> 0.75 (±)	> 35%	NA	NA

Table 1 - Criteria for HCT Effects at extreme values (25% and 65% HCT). NA: not applicable (Hematocrit Effect was defined only for differences between the extreme values, so comparisons with the Control were not performed).

Of the 547 biochemicals analyzed, only 4% (20) had HCT effects when 60% HCT was compared to control, while 25% of biochemicals (137) had HCT effects when 30% HCT was compared to control. Of those 137, 83 changed by more than 40%. When we plotted the HCT effect frequency versus HCT, the number of HCT effects decreased linearly as the experimental HCT got closer to the control HCT, as seen by a coefficient of determination (r^2) of 0.993. The data suggests that less than ~20% of biochemicals may be affected by HCT when the low and high ends of the normal range are compared directly (~37 vs 60% HCT). However, considering that the average adult population of men and women ranges from ~37–54% HCT, the percentage of affected biochemicals would be expected to be less in a well powered Case vs Control study. As demonstrated in the next section, *the overall impact on metabolomic results will likely be minimal, as the chance that aberrant HCT samples will be dispersed across groups is greater.*

Comparison of extreme HCT

To more directly compare the results when analyzing low, medium, and high HCT levels on DBS cards, a second experiment was performed in which five lots of whole blood were manually adjusted to 25% (low), 45% (Control), and 65% (high) HCT by adding or removing RBCs (performed by BioIVT). Four DBS cards were prepared and extracted per condition. Batch-normalized means were used to perform *t*-tests, calculate relative percent differences (RPD), precision, %CV, and correlation coefficients (*r*-values) across all HCT comparisons (25:45, 45:65, 25:65), on each biochemical detected in ≥1 of 5 lots in at least one HCT level. Outcomes were assigned as shown in **Table 1**.

A total of 671 biochemicals were analyzed, with over half (366; 54%) showing no HCT effects, 59 (9%) showing minor effects, and 26 (4%) with moderate effects (**Figure 2**). Thus, 67% of biochemicals were largely unaffected by the HCT extremes. There were 220 biochemicals (33%) which did exhibit some level of HCT effects in which precision across the low, mid, and high HCT samples was poor (CV > 20%). Of these, 85 changed by more than 2-fold when either the low and/or high HCT extreme was compared to mid-HCT, the majority of which (58 of 85) were only

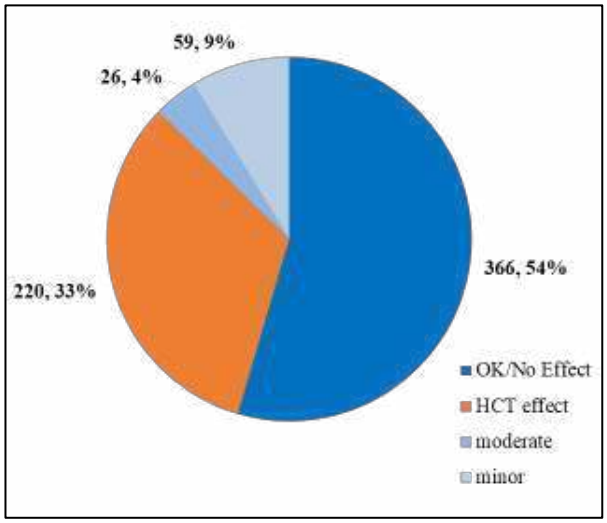

Figure 2- Numbers and percent of total compounds (n=671) of each outcome when comparing low (25% HCT) or high (65% HCT) extreme values to normal (45%).

Metabolon, Inc.
617 Davis Drive, Suite 100, Morrisville, NC 27560

affected when the aberrant HCT was low. Most biochemicals (90%) exhibited a positive correlation between response and HCT (*i.e.*, higher levels at higher HCT, lower levels at lower HCT); 10% (22 of 220) were found to decrease as HCT increased. This could be due to some stability effects caused by longer drying time associated with higher HCT levels. For example, we have previously observed this to be true for diacylglycerols (DAGs) in both short-term and long-term stability experiments. *Most importantly, only 1% of the 671 total biochemicals were affected at both extremes.* This indicates that even if a group were to be made entirely of samples that had HCT values ±20% compared to normal adult values, a very limited number of biochemicals would be expected to have significantly biased results.

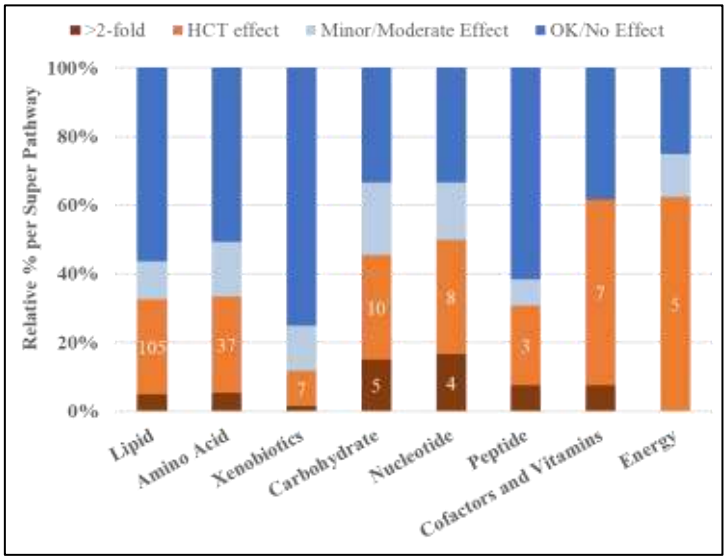

Figure 3 - Percentage of each Super pathway affected by HCT extremes values (65% vs 25%). Numbers of compounds are shown in the bars for HCT effects and >2-fold difference.

Looking deeper at which classes and biochemicals were affected by HCT extremes, we found that over half of biochemicals in Super Pathways (excepting Energy and Cofactors/Vitamins) showed either No Effects or only Minor Effects (**Figure 3**). Less than 20% of the biochemicals in each Super Pathway had greater than 2-fold changes between the HCT extremes (25% compared to 65%). Relative to other Super Pathways, Carbohydrates and Nucleotides appear to have a higher relative proportion (% per Pathway) with these large differences. However, only 4–5 biochemicals were affected in each, representing a very small number (0.7%) of the total biochemicals detected. Xenobiotics contained the lowest relative frequency of HCT effects at 1%.

Hemotcrit as normalization factor

We also investigated whether the hematocrit level (if available) would be useful as a normalization factor. The previous experiments utilized manually adjusted HCT levels from a single lot of whole blood, so we also wanted to investigate "real world" DBS specimens with known HCT values. A cohort of DBS cards and matched plasma from 100 mothers and their infants was utilized for this purpose. This allowed assessment of the HCT ranges that might be present within a study, and to investigate whether normal biological processes affect the distribution of biochemicals into plasma and RBC components in ways that differ from those observed from blood with artificially created HCT levels. Given the larger range in HCT levels in the infant set, it was not surprising that the overall correlations of biochemical levels to HCT were stronger in this group (**Table 2**). *However, for both Mother and Infant groups there were large numbers of metabolites that did not show any correlation to HCT level.* To test the effects of normalization directly, we analyzed by Welch's *t*-Test two subgroups, both with and without HCT normalization, and found very little difference in the *p*-values of the top 20

Group	N	HCT Range	Median HCT	Top 5 Positive correlation range	Top 5 Negative correlation range
Mother	100	20.0 – 39.1	32.6	0.52 to 0.31	-0.36 to -0.30
Infant	100	25.8 – 70.0	48.0	0.82 to 0.52	-0.49 to -0.40

Table 2 - Overview of the group composition and ranges (*r* value) for the Top 5 Positive and Negatively correlated metabolites in each group. With the larger range in HCT values, the Infant group showed better overall correlations, while both groups had large numbers of biochemicals that were not well correlated with HCT.

Metabolon, Inc.
617 Davis Drive, Suite 100, Morrisville, NC 27560

METABOLITE	Welch's t-Test *p*-Value	
	No Normalization	HCT Normalized
lamivudine	5.80×10^{-22}	4.18×10^{-22}
tryptophan betaine	7.54×10^{-10}	1.17×10^{-9}
5alpha-pregnan-3beta,20beta-diol monosulfate	8.03×10^{-9}	1.65×10^{-8}
pregnenetriol sulfate*	2.31×10^{-7}	3.20×10^{-7}
5alpha-pregnan-diol disulfate	2.54×10^{-6}	5.53×10^{-6}
pregnenediol sulfate (C21H34O5S)*	3.36×10^{-6}	4.64×10^{-6}
pregnenediol disulfate (C21H34O8S2)*	3.64×10^{-6}	4.12×10^{-6}
pregnanediol-3-glucuronide	8.56×10^{-6}	1.37×10^{-5}
hyocholate	1.98×10^{-5}	3.50×10^{-5}
1-palmitoyl-2-arachidonoyl-GPI (16:0/20:4)*	1.06×10^{-4}	7.01×10^{-4}

Table 3 - Comparison of *p*-Values of the top 10 most significant metabolites from a Welch's *t*-Test performed on non-normalized and HCT-normalized data from the Infant dataset. Results show that normalization to HCT has very little effect on the metabolites with strong, statistically significant differences between groups in pairwise comparisons.

most significant named biochemicals (a subset of which are shown in **Table 3**). This suggests that for the most highly significant differences, normalization does not improve or strengthen the quality of the data. Additionally, for the metabolites mentioned above with no correlation to HCT, normalization would be expected to confound results. *Thus, we currently advise that normalization to hematocrit is unnecessary for most well powered studies*. Lastly, we note that for datasets in which HCT is unavailable but normalization is deemed warranted, we have identified a possible surrogate for HCT; a chemically unnamed but reliably detected peak that correlates well with known HCT values ($r = 0.82$ for the Infant dataset and 0.52 for the Mother dataset).

Conclusions

Given the wide range of potential hematocrit levels in human populations of differing ages or disease states, it is not unexpected that extreme values (low or high) that differ from the normal range of values in adults will be encountered. The role of carbohydrates, nucleotides, cofactors/vitamins, and energy metabolites in red blood cells, coupled with the fact that higher/lower HCT means more/less RBCs per volume of blood extracted makes it unsurprising that a higher proportion of these pathways (>40% of each) exhibited a significant change between the two HCT extremes (25% vs 65%). But overall, the findings presented here are consistent with the idea that in a sufficiently well-powered study, the effects of varying hematocrit levels among samples will likely be minimal. Only if extreme levels exist between various groups within a study would HCT effects potentially confound the analysis for a limited set of biochemicals. For example, in a Case *vs* Control study where the mean HCT of one group is >30% different than the other with relatively small ranges of values for each group. Such a situation could also occur if the age ranges of groups differed (*i.e.*, newborns compared to children). In these cases, choosing a study design that takes these factors into account, or increasing group sizes sufficiently to account for normal variability in the sample population is suggested. It is important to note that even in scenarios that might cause a large difference in the HCT of various groups, (*i.e.*, age group differences, disease states, etc), real differences in the metabolomic profiles are likely to overpower the relatively modest HCT effects we have presented here. There is currently limited evidence to suggest that normalization of the entire dataset to HCT (or surrogates of it) is of utility in sufficiently powered, well-designed studies. However, we continue to explore methods of normalization of a limited subset of biochemicals that are impacted by this value.

Metabolon, Inc.
617 Davis Drive, Suite 100, Morrisville, NC 27560

References

1. Mondal, H. and D.P. Budh, *Hematocrit*, in *StatPearls*. 2021, StatPearls Publishing: Treasure Island (FL).
2. Billett, H.H., *Hemoglobin and Hematocrit*, in *Clinical Methods: The History, Physical, and Laboratory Examinations*, H.K. Walker, W.D. Hall, and J.W. Hurst, Editors. 1990, Butterworths: Boston.
3. Bashir, B.A. and S.A. Othman, *Neonatal polycythaemia.* Sudan J Paediatr, 2019. **19**(2): p. 81-83.
4. Moore, C.A. and A. Adil, *Macrocytic Anemia*, in *StatPearls*. 2021, StatPearls Publishing: Treasure Island (FL).
5. Badireddy, M. and K.M. Baradhi, *Chronic Anemia*, in *StatPearls*. 2021, StatPearls Publishing: Treasure Island (FL).
6. Hall, E.M., S.R. Flores, and V.R. De Jesús, *Influence of Hematocrit and Total-Spot Volume on Performance Characteristics of Dried Blood Spots for Newborn Screening.* International journal of neonatal screening, 2015. **1**(2): p. 69-78.
7. Evans, A.M., et al., *High Resolution Mass Spectrometry Improves Data Quantity and Quality as Compared to Unit Mass Resolution Mass Spectrometry in High-Throughput Profiling Metabolomics.* Metabolomics, 2014. **4**(2): p. 132.

Metabolon, Inc.
617 Davis Drive, Suite 100, Morrisville, NC 27560