Supplementary Materials for

Point-of-Care microchip electrophoresis for integrated anemia and hemoglobin variant testing

Ran An[1], Yuncheng Man[1], Shamreen Iram[2], Erdem Kucukal[1], Muhammad Noman Hasan[1], Ambar Solis-Fuentes[1], Allison Bode[1], Ailis Hill[1], Kevin Cheng[1], Sanjay P. Ahuja[3], Jane A. Little[4], Michael Hinczewski[2], Umut A. Gurkan[1,5,*]

[1] Department of Mechanical and Aerospace Engineering, Case Western Reserve University, Cleveland, OH, USA

[2] Department of Physics, Case Western Reserve University, Cleveland, OH, USA

[3] Division of Hematology and Oncology, Rainbow Babies and Children's Hospital, Case Western Reserve University, Cleveland, OH 44106, USA

[4] Division of Hematology/Oncology, Department of Medicine, University of North Carolina, Chapel Hill, NC, USA

[5] Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, USA

*** Corresponding author:**

Umut A. Gurkan, PhD

Warren E. Rupp Associate Professor

Case Western Reserve University

Office: Glennan 616B, 10900 Euclid Ave., Cleveland, OH 44106.

Telephone: +1 (216) 368-6447

E-mail: umut@case.edu

1. Supplementary Tables

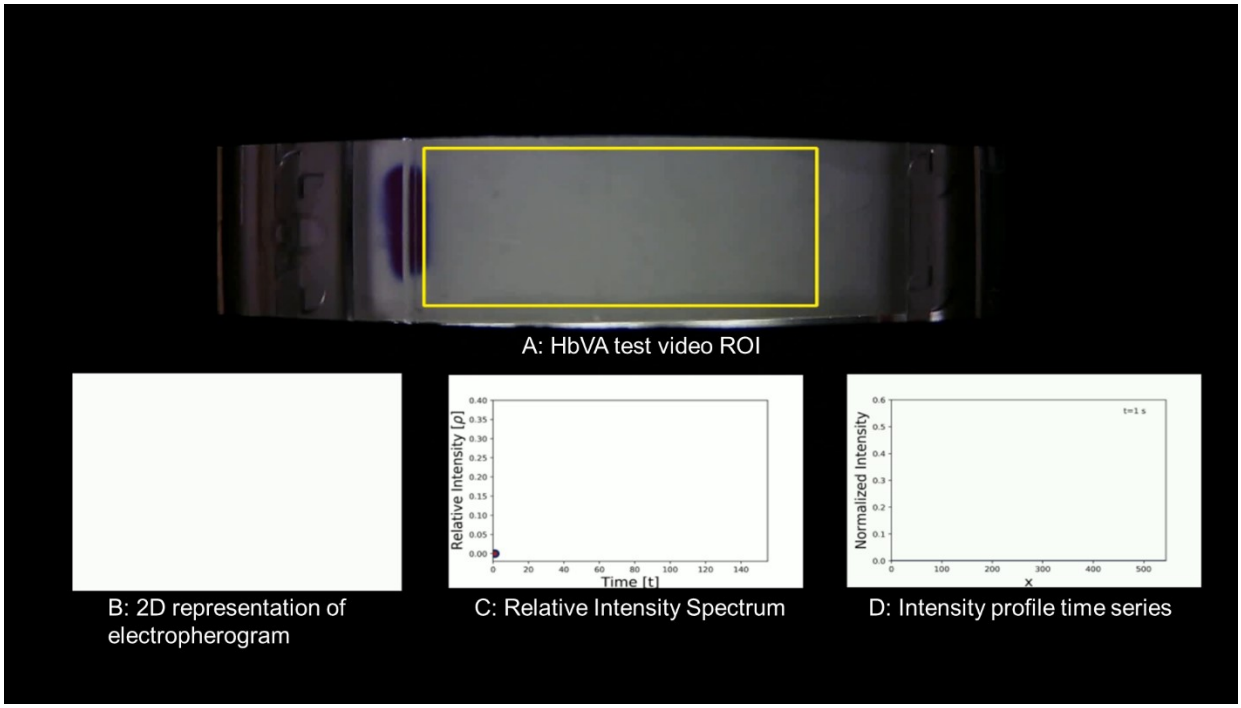
Table S1. Clinical Laboratory Standard Institute (CLSI) suggested test subjects for Hb level measurement using new technologies

	Total	≤9.0 g/dL	9.1-12.0 g/dL	12.1-17.0 g/dL	≥17.1 g/dL
Sample percentage (CLSI recommended)	100%	15%	25%	50%	10%
Number of Samples (CLSI recommended)	≥40	≥6	≥10	≥20	≥4
Number of Samples (Tested)	46	17	19	10	0
Sample percentage (Tested)	100%	37%	41%	22%	0

Table S2. Artificial Neural Network Machine Learning Training

Iteration #	Mean Relative Error (MRE)%
1	6.66312
2	6.13888
3	7.35896
4	6.95064
5	7.38383
6	7.15175
7	7.12065
8	6.66197
9	5.70314
10	7.15175
Average Mean Absolute Error (MAE)	6.83±0.17

2. Supplementary Video



Video S1. Overview of HbVA integrated Hb level determination, anemia detection and Hb variant identification

3. Efficacy of ANN Based Processing Pipeline

For image processing and data analyses of video data generated from the HBVA tests, we utilized an automated machine learning based protocol. Our goal was to look for signatures encoded in the relative intensity spectrum obtained from the HbVA data (Hemoglobin vs standard calibrator dye), and correlate that to the corresponding Hb level for the sample. A detailed breakdown of the analyses workflow and network details can be found in Methods. To test the efficacy of our Hb level predictor ANN, we used repeated random sub sampling validation. This involves repeatedly randomly splitting the available dataset into training and testing groups of fixed sizes. Our initial data set consisted of 68 sample videos, out of which we used a set of 27 samples for training, and the remaining 41 as the testing set holdout. This choice of training set size (which is smaller than test set) was made after running many initial training runs on various set sizes, comparing average training results for each set size, and selecting the smallest set size that delivered performance comparable to the best achievable performance. We wanted our workflow to be adaptable to a scenario where availability of data is limited, and optimize for minimal data set size. For each iteration of the split, the ANN was fit to the training data and evaluated on the testing data. Overall network performance and a measure of the quality of its predictive power was then obtained by averaging over the evaluation metrics obtained from each independent training iteration. While k fold cross validation is a more popular choice, for our problem we needed to design a predictor ANN with a non-traditional training: testing split. So we wanted to fix our split ratio to this value instead of one determined by the number of iterations à la k fold cross validation. The quality of each training iteration was evaluated using the resulting mean absolute relative error (MAE) achieved on the testing set. MAE is computed from the mean of absolute relative difference between HBVA predicted and observed CBC values i.e.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_{HbVA} - y_{CBC}|/y_{CBC}$$

for a test set of size n. The repeated random sub sampling was run for 10 iterations, yielding an average MAE of 6.83% in testing, which matches/beats human/manual performance estimates. Results from our repeated sub sampling validation are summarized in Supplementary Table.S2.

The testing set was later augmented by a further 5 sample videos from a follow up data set, which played no part in the sub sampling validation. The network corresponding to the run with MAE closest to the mean MAE was chosen as the final best version network, and tested on the combined testing set of 46 samples. Test results for CBC vs HbVA Hb level determination for the combined testing set (completely unseen by the best version ANN) are shown in **Fig. 5A** in the main text. The solid diagonal line indicates the line of perfect agreement between the two testing methods. We see how HbVA quantified Hb levels as predicted by the ANN strongly and accurately correlate with the CBC values, with MAE of 6.2%.