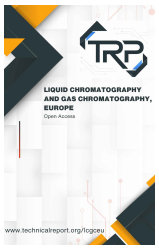


Journal of Liquid Chromatography and Gas Chromatography, Europe
https://technicalreport.org/lcgceu/

LC GC Eu., Vol. 37, No.2, 16-24, 2024

Technical Reports Press

RESEARCH ARTICLE

Residual Speckle Dose-Path Calibration for Label-Free CEACAM5 Quantification with Antibody-Functionalized Tapered Multimode Microfibers

C. Ruan[1*], **Y. Leo**[1] **and W. Mei**[1]

[1] *Jiangsu Key Laboratory of Advanced Structural Materials and Application Technology, School of Materials Science and Engineering, Nanjing Institute of Technology, Nanjing, 211167 P. R. China*
[*] *Corresponding author*

Corresponding author:
198776576@163.com

Abstract
CEACAM5 is a known glycoprotein antigen for which reliable measurement relies on repeatable concentration tracking through repeated measurements. This study proposes a label-free optical sensor based on the CEACAM5-induced modal redistribution analysis of tapered multimode microfibers in terms of their residual optical signature as opposed to the label-dependent identification of fixed images. The problem posed here concerns whether the CEACAM5 concentrations in the range of $1\,\mathrm{ng\,mL^{-1}}$ to $1000\,\mathrm{ng\,mL^{-1}}$ can be determined from the residual speckle signature with the retention of the initial steepness in the lower concentration part and the subsequent saturation at higher antigen concentrations. The data used include 900 specklegrams obtained at 9 different concentrations with respect to a tapered waist of 7.46 µm, tapering length of 15 000 µm, near-infrared illumination wavelength of 1550 nm, 6 milliseconds exposure, and separate calibration, validation, and independent testing sets of the acquired images. In each case, a residual optical field with respect to the reference state is determined and characterized in terms of entropy, radial Fourier energy, gray-level co-occurrence texture analysis, angular anisotropy, lacunarity, and wavelet energy. The residual fields are used as descriptors to establish a CEACAM5 concentration path along which branch-aware Gaussian processes calibration takes place. For the lower concentrations between $1\,\mathrm{ng\,mL^{-1}}$ and $50\,\mathrm{ng\,mL^{-1}}$, a slope of -0.00123 response units per $\mathrm{ng\,mL^{-1}}$ with a coefficient of determination of 0.99221 is achieved, while the saturating trend across the whole interval is described with a lower response constant near 0.82648 with $R^2 = 0.99173$. The findings answer the research question affirmatively: residual speckle dose-path calibration converts CEACAM5-induced modal redistribution into an uncertainty-qualified concentration estimate without reducing the assay to categorical image recognition.

Keywords: CEACAM5; tapered multimode microfiber; label-free biosensor; specklegram spectroscopy; residual texture; optical calibration; Gaussian process; analytical instrumentation

Received: 18 January 2024
Accepted: 29 May 2024
Available online: 30 June 2024

1. Introduction

Carcinoembryonic antigen-related cell adhesion molecule 5 (CEACAM5), often quantified as part of the carcinoembryonic antigen family, is an important diagnostic tumor-associated glycoprotein applied in colorectal cancer follow-ups, treatment assessments, and post-surgery diagnostics. This marker does not have much value as a screening tool but still allows for repeat measurements, as fluctuations in levels may help to assess risks, treatment response, or progression based on clinical observation [1–3]. Discussed in terms of guidelines for GI tumor markers, CEACAM family tests should be understood as requiring context, repeatable measurements, and accuracy, rather than just end-point readings [4, 5]. Such a diagnostic use poses significant analytical challenges. The required test must be selectively measuring this target at very low levels, have a broad dynamic range, and be capable of accurately reporting uncertainty. It cannot only indicate

Cite as: C. Ruan, Y. Leo & W. Mei (2024). Residual Speckle Dose-Path Calibration for Label-Free CEACAM5 Quantification with Antibody-Functionalized Tapered Multimode Microfibers. LC GC Eu., 37(2) (2024) 16-24.

high antigen level; it also needs to reflect ordering in concentration, sensitive region, and saturation onset.

Standard immunoassays utilize antigen-antibody binding as the means for biochemical specificity; however, the use of labels and multi-reagent steps, the need for end-point readout, and laboratory-based analysis are common drawbacks in most cases. Label-free optical techniques mitigate these concerns by converting specific binding events to changes in refractive index, intensity, phase, wavelength, or spatial distribution [6–8]. Biosensor types that exploit refractive-index sensing, including surface plasmons, resonators, interferometers, waveguides, and optical fibers, have been shown to be highly analytically useful in biomolecular studies [9–11]. While integrated photonic devices have excellent possibilities in miniaturization and multiplexing, optical fiber geometry offers advantages in compact alignment and remote detection as well as fluid sampling capabilities. Out of the optical fiber-based methods, the tapered optical fibers are especially attractive because of the shrinking of the core diameter which results in enhanced sensitivity towards the surrounding chemical medium [12, 13]. It follows that the CEACAM5 layer deposited on the taper would perturb the field propagation without labeling.

Speckles formed by multimode microfiber have a detailed analytic signal due to recording many interfered modes into an output pattern. The grain distribution depends on phase interaction, coupling, bending, surface disturbance, and propagation of guided waves that agree with guided wave and Fourier optics concepts of imaging [14–16]. Speckle is commonly considered a random optical fluctuation, but it has great sensitivity to tiny changes and is useful in sensor applications due to a number of interferences that can be observed within a single camera exposure [17–19]. The binding of CEACAM5 antigens changes the index of refraction at the interface of the fiber with an antibody-coated taper. As a result, the output speckle changes its properties in terms of reorganized intensity distribution, new frequencies, local textural changes, and directional bias of the modal interference. Those changes are too scattered for a simple brightness measurement but appropriate for spectra and textures.

Calibration is inherently challenging since the speckle response contains significant information. The direct image classification approach can discriminate between the nine precalibrated concentrations; however, it is different from the process of quantitative estimation of CEACAM5 levels. In this case, the sample is considered to have one value on the continuum rather than being categorized into any of the image groups. A classifier can provide a certain label corresponding to the intermediate or dominant saturation response, although the true value should have a rather wide interval in case of the continuous variable. Chemometric analysis is thus another story because the response must be translated into physical parameters, ranked according to the increasing antigen concentrations, and estimated with confidence intervals. This is possible with such approaches as texture statistics, wavelets, and nonlinear manifold learning since they reveal the structure of the speckle distribution at different scales without obscuring the actual data through complicated image transformations [20–22]. It is useful to utilize neighbor-preserving maps in case of the continuously changing analytical response [23].

The current manuscript studies the CEACAM5 microfiber specklegram analysis using residual speckle dose-path calibration. The steps in the method are clear: establish an optical reference state, capture antigen-induced specklegrams for different CEACAM5 levels, calculate the residual mode rearrangement, derive spectral-textural features, place the features on a concentration path, and calibrate using branch-specific Gaussian process to get the CEACAM5 concentrations with confidence intervals. The core scientific inquiry here is that the protocol would be capable of quantifying CEACAM5 over 1 ng mL^{-1} to 1000 ng mL^{-1}, while ensuring that the low concentration range remains an independent region for sensitivity. The inquiry will be conducted on 900 specklegrams, with split training/test sets, two branches, and full interpretation of optics, chemistries, features, calibration parameters, and decision rule.

2. Materials and Methods

2.1. Design of the assay and specklegram collection

For the analysis, we employed a tapered multimode fiber probed with an excitation wavelength of 1550 nm. In this study, the concentration range of CEACAM5 covered 1, 5, 10, 20, 30, 50, 100, 500, and 1000 ng mL^{-1}. The lower values of this series were measured at an elevated density since early binding contributes most to the marginal optical signal prior to approaching the saturation limit. The upper range checks whether the reaction remains organized in the course of achieving the saturation limit of the recognition layer. The total set of specklegrams comprises 900 images, where 450 were collected for calibration, 225 for validation, and 225 for testing.

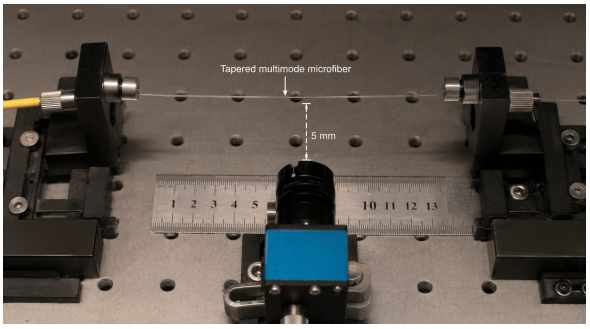


Figure 1: Optical acquisition bench.

Optical system shown in Figure 1 specifies the experimental setting where the assay takes place. A tapered microfiber is placed between mounts and the far-field speckle image is captured by the camera at a constant working distance. Optical setup is compact yet highly sensitive to the quality of light coupling, proper positioning of the fiber, distance from the camera, and potential instrumental drifts. In order to minimize the influence of instrumental effects on our measurements, we treat each frame as one observation within replicated concentration groups.

These parameters, which are encapsulated in Table 1, determine the assay's instrumental identity. Specifically, the 7.46 µm waist enhances the evanescent field's interaction with its surroundings, and 1550 nm interrogation means that our measurements take place in an established near-infrared region of fiber optics. In addition, the short duration allows for quick re-measurements, and a data set containing 900 frames contains enough repetitions to investigate within-run reproducibility. Just as crucial is the list of chemicals in Table 1: piranha cleaning, silanization with APTES, glutaraldehyde activation, antibody conjugation, and BSA blocking will determine

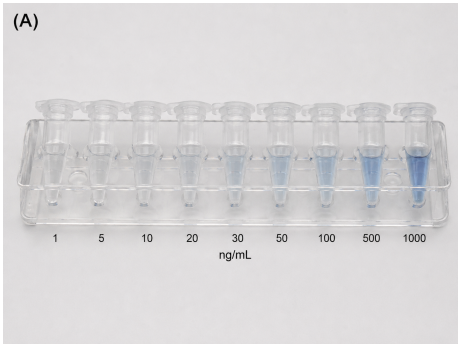
The design presented in Figure 2 demonstrates the measurement procedure well. The concentration series includes the high sensitivity low concentration line and the high concentration saturation line; repeat images account for variation at the level of optics; and calibration-validation-test design prevents the last test from being tuned to the frames that were used for measurement reporting. The latter is particularly significant in the case of speckle detection since the visually indistinguishable images may differ at the descriptor level.

2.2. Surface functionalization and transduction pathway

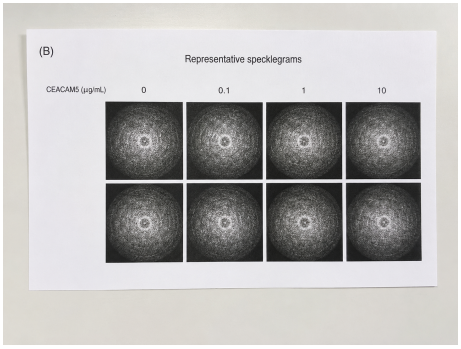
The surface of the tapered fiber was modified via steps including hydroxylation, silanization, aldehyde derivatization, antibody immobilization, and blocking. The process starts with piranha treatment

Table 1: Assay conditions.

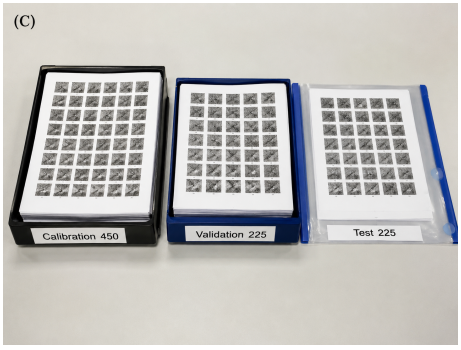
Category	Condition or numerical value
Fiber structure	Tapered multimode microfiber with 7.46 µm waist width and 15 000 µm tapering length
Optical interrogation	Near-infrared laser centered at 1550 nm
Image acquisition	CMOS camera recording with 6 ms exposure time
Camera geometry	5.3 µm by 5.3 µm pixel size; 5 mm fiber-to-camera distance
Surface preparation	Piranha treatment for 60 min; APTES treatment for 60 min; glutaraldehyde activation for 40 min; antibody immobilization for 60 min; BSA blocking for 30 min
Recognition reagent	CEACAM5 antibody at 1 µg mL^{-1} in PBS at pH 7.4
Analyte levels	1, 5, 10, 20, 30, 50, 100, 500, and 1000 ng mL^{-1} CEACAM5
Measurement protocol	1 h measurement at each concentration and 10 min PBS rinsing between concentration groups
Specklegram corpus	900 frames across nine concentration levels
Image allocation	450 calibration images, 225 validation images, and 225 independent test images



(a) Dose sequence



(b) Replicate fields



(c) Image allocation

Figure 2: Specklegram allocation.

resulting in the introduction of hydroxyl groups to the surface. Subsequent silanization using APTES results in amino-silane functionality. Glutaraldehyde provides aldehyde moieties which can enable further covalent conjugation of CEACAM5 antibodies. Blocking is done using BSA to avoid non-specific adsorption. Such modification pathway has been extensively applied to biosensing, as it involves creation of a surface layer responsible for selective interaction with the analyte of interest [24]. In this assay, the binding of the CEACAM5 alters the local refractive index surrounding the taper.

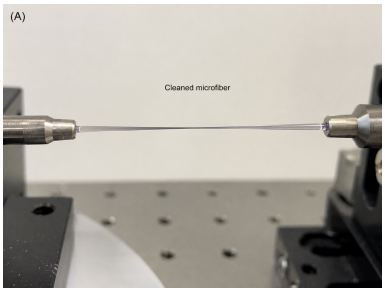
Each surface treatment in the sequence of Figure 3 corresponds to its functional purpose in the analysis. The clean taper provides an interface with reactivity, APTES and glutaraldehyde create the chemical bonding, the antibody layer introduces the CEACAM5 specificity, and the last layer prevents non-specific binding. In that sense, the specklegram decoding requires not a separate physical process but a chain of biochemical reactions. An unstable recognition layer may lead to optical changes but with no correlation to the target analyte, whereas with stable antibodies, the residual speckle can be considered as a measurement signal.

This figure helps understand why residual extraction is important before descriptor calculation. This is because the specklegram has high interference structures, which may hinder any alteration due to the target molecule. The importance of residual field calculation is that it focuses more on the redistribution of structures after the antigen treatment, and thus makes the descriptors more relevant to the changes caused by CEACAM5.

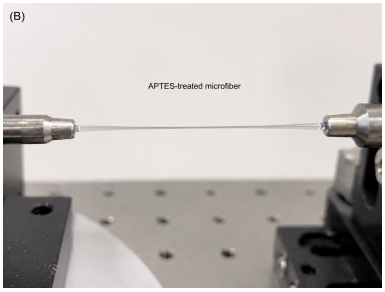
2.3. Speckle residual dose calibration

This method of analysis is deliberately left free from equations in the main discussion as the intention is to provide the interpretation and reproducibility of the optical measurement as opposed to introducing any unnecessary math. The specklegram images are initially normalized in order to remove fluctuations arising from fluctuations in laser power or changes in camera gain settings. Then, the method uses a reference optical state to determine the residual speckle pattern for each of the frames exposed to antigens. The following quantities are determined from the residual speckle pattern: entropy (global), radial Fourier band energy, gray level co-occurrence contrast, gray level homogeneity, angular anisotropy, lacunarity, and wavelet band energy. Each was chosen on the basis of its visual or physical meaning.

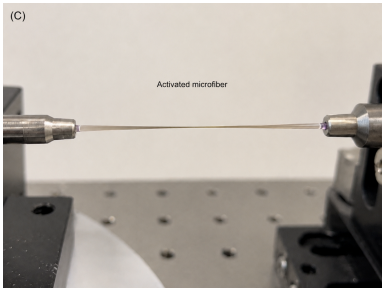
The descriptors are ordered on a dose path according to the concentration order of the analyte. Images that have similar residual spectral-texture characteristics are placed adjacent to each other, and an increase in the CEACAM5 concentration will result in a movement along the path in the same direction. This is not an aesthetically pleasing image mapping scheme. Rather, it is a geometry based on a calibration scheme, constrained by the order of the analytes, and the physics that more binding cannot randomly turn around the direction of the optical response. The concentration estimate is made using a Gaussian-process model because the relationship between the response and the predic-



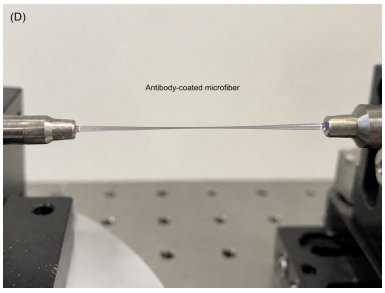
(a) Clean taper



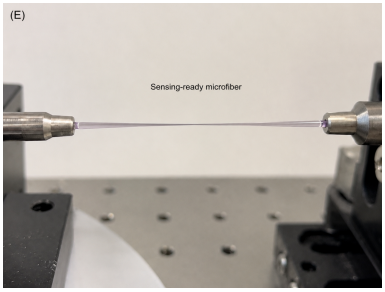
(b) Silanized taper



(c) Aldehyde surface



(d) Antibody layer



(e) Blocked sensor

Figure 3: Recognition-layer preparation.

tor is non-linear and smooth. In addition, any good analytical result must be accompanied by the associated uncertainty [25, 26]. In fact, statistical learning texts suggest separation between training, feature selection, and final evaluation of prediction for high dimensional data [27].

The route depicted in Figure 5 illustrates how continuous calibration

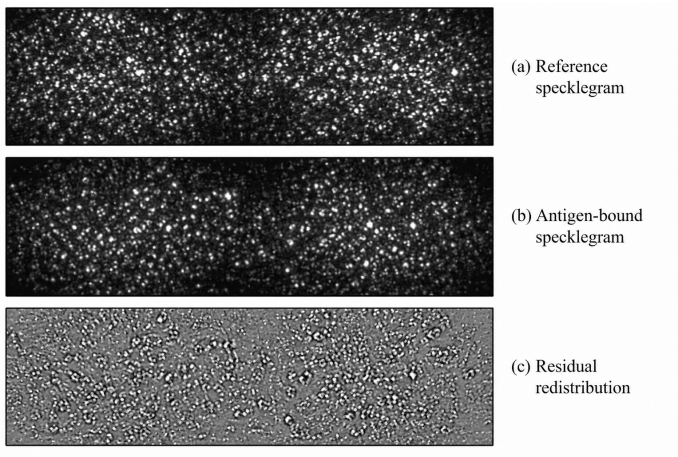


Figure 4: Residual speckle extraction.

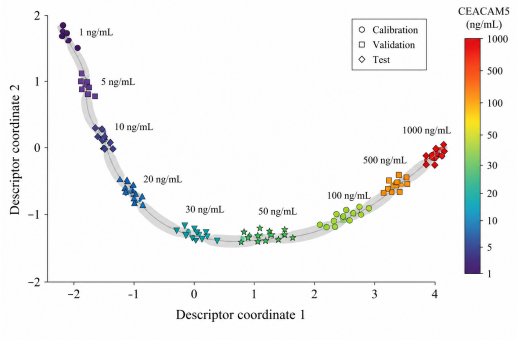


Figure 5: Ordered dose path.

can be achieved. CEACAM5 images that correspond to low concentrations appear on one side of the coordinate system, those that fall within intermediate concentrations create the bend, while those that are high concentrations tend towards a saturated region. The structure itself allows for better interpolation between concentrations rather than using a direct class-label for an image. It also provides information on when uncertainties in the data are likely to occur: around the bend from the low-linear concentration branch and at the high end.

2.4. Calibration constants and branch interpretation

The response is analyzed through two mutually supportive branches. First, the branch of 1 ng mL^{-1} to 50 ng mL^{-1}, where the normalized response has a linear dependence. Second, the branch for the entire range with saturation rather than trying to force a linear function over higher concentrations. Two branches make chemical sense due to early antigen adsorption producing a larger marginal change in refractive index, while late adsorption produces a smaller one.

Table 2: Calibration constants.

Response branch	Parameter	Numerical value
Linear branch, 1–50 ng/mL	Intercept	0.98496 ± 0.00139
Linear branch, 1–50 ng/mL	Slope	-0.00123 ± 5.43291e-5
Linear branch, 1–50 ng/mL	Pearson coefficient	-0.99610
Linear branch, 1–50 ng/mL	R-squared	0.99221
Linear branch, 1–50 ng/mL	Adjusted R-squared	0.99026
Full-range saturating branch	Upper response constant	0.99054 ± 0.00454
Full-range saturating branch	Lower response constant	0.82648 ± 0.00819
Full-range saturating branch	Transition concentration	65.84244 ± 9.52209
Full-range saturating branch	Shape factor	1.20221 ± 0.19525
Full-range saturating branch	Reduced chi-square	4.52487e-5
Full-range saturating branch	R-squared	0.99173
Full-range saturating branch	Adjusted R-squared	0.98760

From the constants found in Table 2, we can say that each branch solves a specific question analytically. The low branch evaluates the sensitivity since the negative slope implies consistency in a reduction in similarity when CEACAM5 varies from 1 to 50 ng mL^{-1}. The complete branch evaluates the operational range since the constants in the upper and lower branches define the extent of bending towards saturation of the assay. The transitional concentration of 65.84 ng mL^{-1} plays an essential role in that it falls slightly above the low branch.

2.5. Rule for validating and reporting

The principle of the validation algorithm is different from the principle of the image recognition approach. In analytical-calibration paradigm, calibration frames define the response pathway, validation frames define pre-processing and calibration parameters, and test frames determine the final behavior of the response. The result of inference is accepted if the inferred concentration belongs to the right branch, if the uncertainty interval is not too wide to differentiate neighboring concentration intervals, and if replicates corresponding to the same nominal concentration belong to the same cluster. An excessively wide interval, an inconsistency between the branch and the inferred concentration, and the wrong descriptor path require acquiring new images.

3. Results and Discussion

3.1. Sensor response and compatibility with analysis

The narrow waist of the taper creates an interaction between the waveguide mode and the outer coating due to the small diameter. The wavelength range of interrogation from 1550 nm is suitable for instrumental stability, and the duration of illumination of 6 ms is sufficient to allow repeated image capturing with minimal waiting time. This aspect is critical in spectroscopic analysis since it determines not the sensitivity but the reproducibility of the response creation process. Sequential chemistry is also favorable for the test since the layer of antibodies creates biochemical selectivity of the optical response, while the blocking solution minimizes non-specific responses.

This measurement matrix provides sensitivity and range. High resolution measurements in the range of 1-50 ng mL^{-1} make it possible to analyze the early binding range where every change of the concentration causes significant modal distribution changes. Including 100, 500, and 1000 ng mL^{-1} allows ensuring that the calibration curve does not rely on the low branch alone and whether the analytical function will remain monotonic after including the curved range. Such combination is necessary due to the fact that the analysis of CEACAM5 levels can require sensitivity in certain situations and large changes in others.

The reproducibility of instrumental measurements benefits from interpretational analysis at the level of descriptors as opposed to direct visual inspection of the images. Visual inspection by a human observer may overlook changes due to differences in concentration because they are diffused among multiple grains and various spatial frequencies. However, the residual dose-path methodology transforms this diffuse information into parameters that carry optical significance. Any changes to Fourier energy, co-occurrence contrast, or lacunarity are meaningful changes to the structure of the interference field; this is not the case with categorical labels.

3.2. Low concentration sensitivity

This low-concentration sensitivity represents the main analytical zone of the test. The estimated slope of −0.00123 response units per ng mL^{-1} proves that there exists a noticeable linear dependency between a decrease in the normalized similarity ratio from 1 to 50 ng mL^{-1} due to binding of the analyte. The high value of the Pearson coefficient (−0.99610) and R^2 equal to 0.99221 demonstrate that such a response cannot be explained by an artificial distinction between several groups of images.

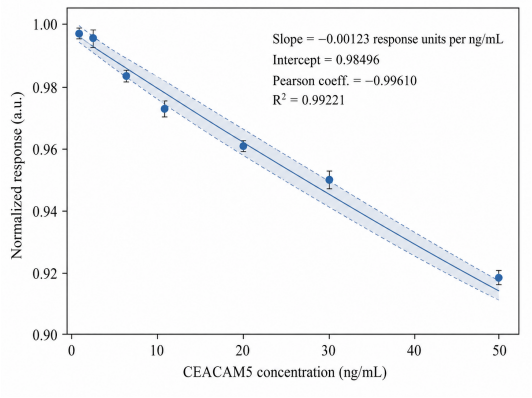


Figure 6: Low-range linearity.

The graph in Figure 6 proves that the initial segment of the dose-response can indeed be considered a high sensitivity calibration range. The values are still on the regression line and have maintained the same response trend at the low concentration levels. Thus, it is recommended to adopt an analysis technique wherein low concentrations of CEACAM5 can be reported based more on the slope value than the saturated high concentration range. Mechanistically, the taper detects early antibody-antigen binding via the modulation of the surface response.

The negative response trend can be reasonably explained from an optics perspective. As CEACAM5 fills up the antibody spots, the refractive index environment surrounding the taper gets modified and results in dissimilar interference patterns relative to the reference case. The resultant speckle intensity field, thus, tends to develop a more pronounced spatial shift. The descriptor-based path accounts for the dissimilarity rather than just the increase/decrease in intensity, which is a more favorable scenario since the latter is extremely sensitive to any changes in laser power or camera gain.

Low-branch calibration also demonstrates why direct concentration classification is not the appropriate output. A classifier could claim an image belongs to 10 or 20 ng mL^{-1}, but it would be unnatural to state that the image is situated somewhere between these concentrations and provide uncertainty intervals. Dose-path calibration allows one to think about the nine concentrations as anchors in terms of continuous measurement, thus fitting better into the usage paradigm of CEACAM5 concentration.

3.3. Saturation behavior within the wide-range interval

Saturation occurs with the upper constant being 0.99054, the lower constant being 0.82648, and the transition concentration of approximately 65.84 ng mL^{-1}. The extremely high value of R^2, namely 0.99173, suggests that the saturated response remains orderly despite the lack of global linearity. Such non-linearity should not be regarded as analytical deficiency. Instead, one needs to comply with the physical response function.

Figure 7 depicts the curvature which comes after the low curve. The presence of the lower asymptote implies that the residual effect continues changing but at reduced marginal differences in response. Such observations are consistent with a gradual occupation of the recognition layer as well as the limited sensitivity of the modal interference to an extra increase in refractive index at a crowded interface. The absence of such a consideration in the process of calibration would lead to overestimation of accuracy.

Interpretation that is aware of saturation helps increase the accuracy of

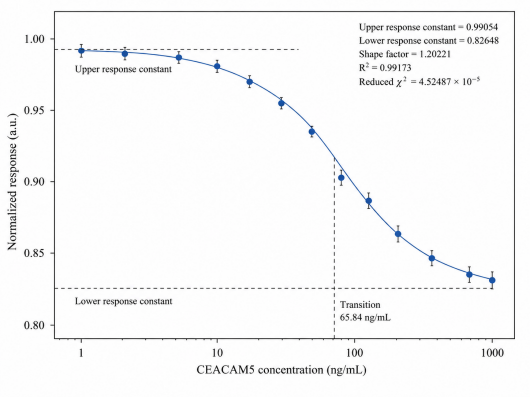


Figure 7: Full-range saturation.

analysis. In the lower branch, the instrument can express differences in concentration with the help of the steeper local gradient. Near the boundary between the branches, one needs to think about possible uncertainty since a tiny difference in response can mean a much larger interval in concentration. In the upper branch, the output must take into account the fact that extra amounts of CEACAM5 affect the descriptor less noticeably. Thus, a seemingly unfavorable effect is transformed into an advantageous outcome.

The large span also explains why the frame partitioning is needed. Without considering validation and test sets, one can get an artificially good picture since the calibration process benefits from using replicates under the same conditions. The latter can be achieved by using independent frames in the evaluation stage.

3.4. Evidence on descriptors from residual specklegrams

At the descriptor level, there is a physical connection made between the changes seen in the speckles and the concentration of CEACAM5. The entropy measure captures whether the residual image is more disorganized; the radial Fourier energy captures whether the speckle pattern exhibits more fine or coarse structures; the co-occurrence contrast and homogeneity capture small-scale changes in the distribution of gray levels; the angular anisotropy captures directional information; the lacunarity measure captures irregular clusters; and the wavelet energy separates scale-specific variations. This list of descriptors comes from well-established methodology in texture and signal analysis [20, 21, 28, 29]. However, their significance lies beyond mere computation.

The set of descriptors shown in Figure 8 shows that the assay is robust against relying on a single susceptible image statistic. The residual specklegram is analyzed from several angles, such as disorder, spatial frequencies, local textures, directionality, clustering, and multiscale energy. The multiplicity in the set of descriptors is advantageous in terms of analysis because different types of perturbations affect different kinds of descriptors. For example, a uniform shift in brightness could impact intensity-based statistics but not directionality; a minor shift in orientation could alter directionality, and noise can impact higher frequencies.

Interpretable feature sets also help in facilitating transfer of the assay. The laboratory may examine the extent to which its specklegrams exhibit consistent frequency and texture patterns rather than relying solely on whether or not a trained classifier yields the desired classification result. This is relevant to biosensor instruments, since different surface treatments, different optical coupling mechanisms, and different camera setups may be used. The inspection of features allows the analyst to pinpoint the source of response modification.

Further, residual descriptors contribute to uncertainty quantification. For an image that falls between two doses but exhibits consistent patterns in its descriptors, reporting continuous doses in a medium range would be appropriate. In a situation where the features in entropy, radial energy, and co-occurrence indicate various locations in the dose path, the range becomes wider. This phenomenon is challenging to replicate using just one similarity measure and more difficult for a simple image categorizer system.

3.5. Concentration estimates with awareness of uncertainty

This report does not provide just the concentration estimate, but rather the concentration estimate along with the branch assignment and uncertainty estimate. The point cannot be understated since for speckle-based sensors, which have high physical sensitivity, such sensitivity also comes with a certain vulnerability to changes in the environment and misalignment. An interval-free report of a deterministic value would incorrectly suggest that all images are equally robust.

This sequence in Figure 9 transforms the position along the dose-path into a quantitation choice. The first one involves categorizing the frame in terms of its association with either the low, transition, or high response group. The second one tests whether the confidence band is sufficiently narrow for meaningful quantitation. Finally, the third choice decides if this particular outcome can be used, re-measured, or considered to be saturation-controlled. This quantitation approach is more relevant for analytical purposes than choosing the nearest known concentration.

The idea of quality control naturally flows from the logic of reporting above. The replicates of the image taken from the same concentration should lie on the same part of the curve. The consecutive concentrations must maintain their order. There should also be no need to impose an artificial precision of discrimination when there is no response in the concentration range. This is thus not indicative of failure in the assay. Instead, it tells us that we cannot accurately measure this optical state in the present acquisition setup.

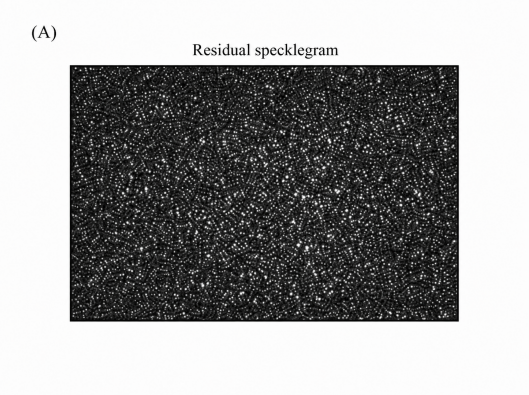
The reporting methodology likewise provides a straightforward path for validating the assay moving forward. Reproducibility might be determined using separate microfiber samples to see if the branching and uncertainty handling remain the same. Matrix effects may be verified through a comparison between responses in PBS versus serum or plasma matrices. Thermal robustness would involve determining if the uncertainty interval expands due to descriptor drifting prior to an erroneous concentration determination. These are all tests that are directly tied to how the assay outputs results.

3.6. Comparison of continuous calibration with image classification

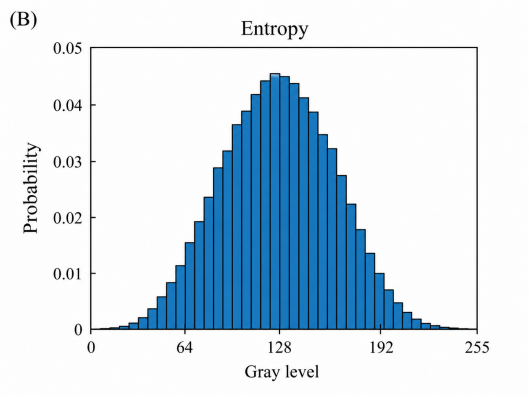
Image classification is appealing in that it yields a high degree of accuracy if the image falls into any of the preset concentration categories. However, this is only possible since the amount of CEACAM5 is a continuous variable. It is possible that the sample does not contain exactly 1, 5, 10, 20, 30, 50, 100, 500, or 1000 ng mL^{-1} of the substance in question. If the model was trained using such concentrations, then it could misidentify any other values. Continuous calibration appears more appropriate for the application.

The comparison presented in Figure 10 highlights the distinction between the two techniques. Classification utilizes fixed class labels that subdivide the output into boxes of concentrations, while the calibration technique utilizes the ordered dose residuals for estimating the concentration level with uncertainty. In the case of CEACAM5 assessment, the calibration technique is better since it is the gradual change in concentration rather than its classification that matters. The dose residual path allows for a physical interpretation of the concentration via descriptor changes.

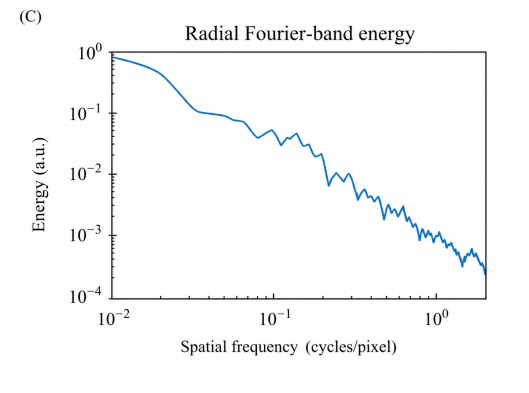
The computational cost of the descriptors method is low. All of Fourier, co-occurrence, lacunarity, and wavelet descriptors can be computed



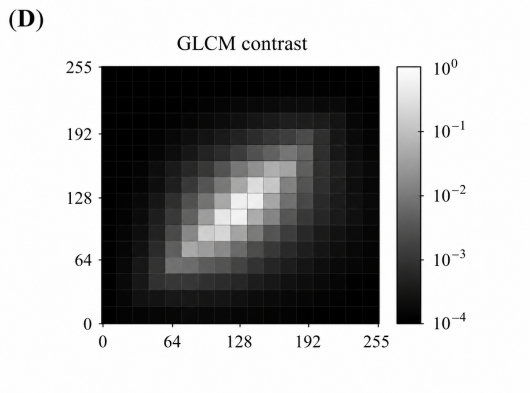
(a) Residual field



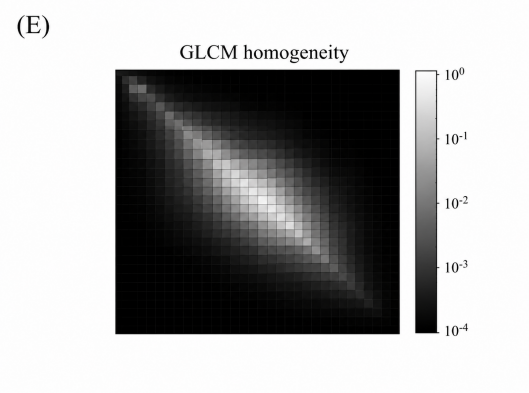
(b) Entropy



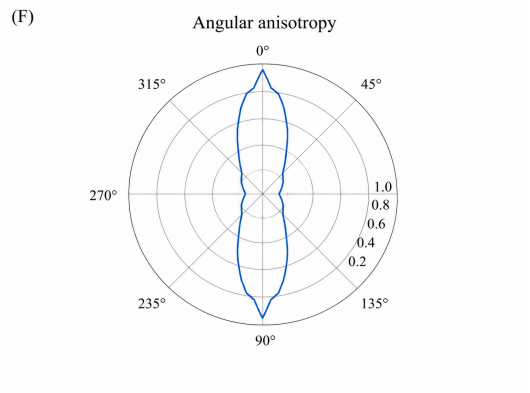
(c) Radial energy



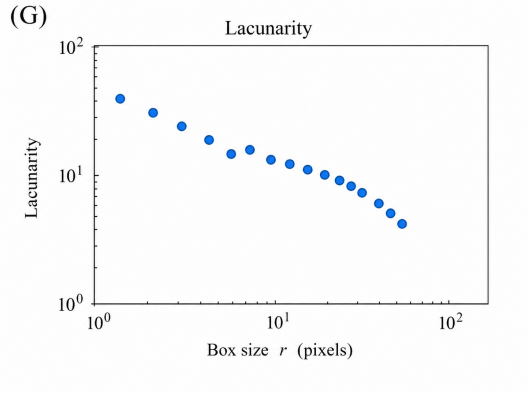
(d) Contrast



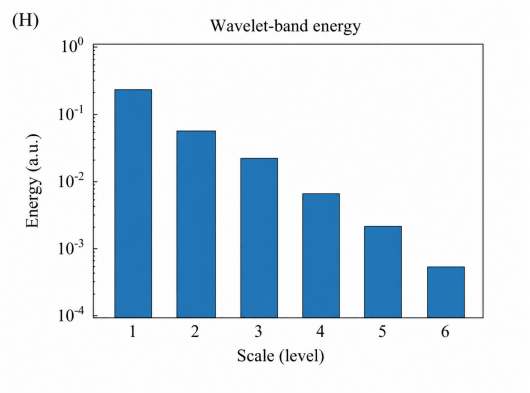
(e) Homogeneity



(f) Anisotropy



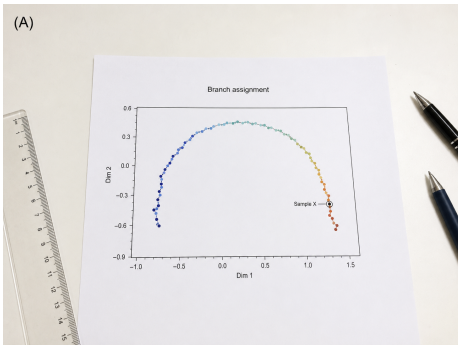
(g) Lacunarity



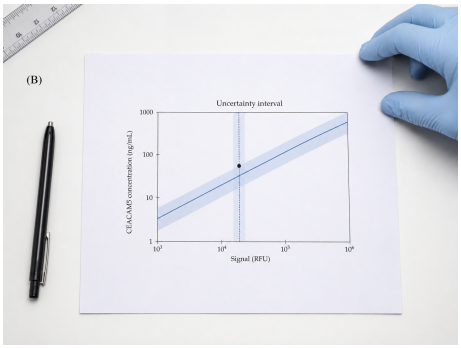
(h) Wavelet energy

Figure 8: Residual descriptors.

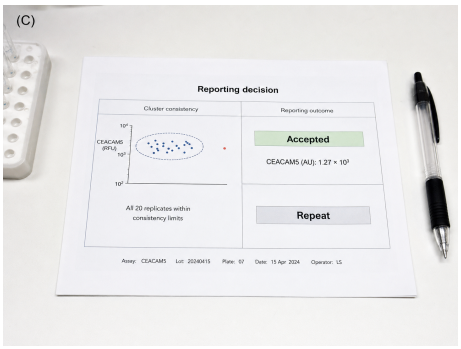
on regular desktop PCs, and the Gaussian process calibration does not cause difficulties when working with the 900-frame database applied here. As a result, the proposed assay may be easily employed in small bench-top devices. Another advantage is that there is no need for the



(a) Branch call

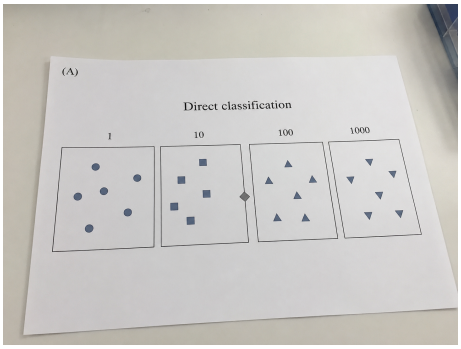


(b) Interval width

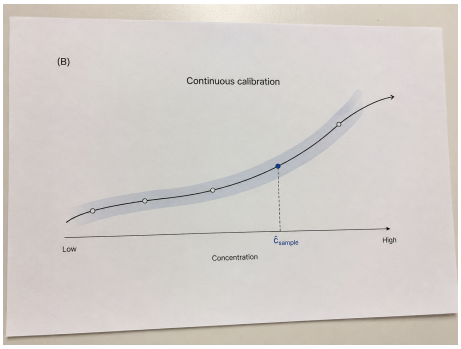


(c) Report decision

Figure 9: Reporting logic.



(a) Fixed labels



(b) Continuous estimate

Figure 10: Classification versus calibration.

extremely large collection of images needed to train an image classifier model.

One such benefit is that the process is auditable at each stage: surface pre-conditioning, residual image creation, descriptor acquisition, dose-path sequence selection, branch selection, and uncertainty quantification. This auditing process becomes important for analytical chemistry since method efficacy cannot be proven solely based on prediction accuracy alone. The testing method can therefore be improved through a process of elimination that finds the source of mistakes made.

3.7. Implementation requirements

The findings validate residual speckle dose-path calibration as means of determining CEACAM5 level, however, some constraints persist. The set of concentration levels covers nine values, which are widely spaced at 100, 500, and 1000 ng mL^{-1}. Some additional concentrations around 50 ng mL^{-1}-150 ng mL^{-1} will make transition from low to high more accurate around 65.84 ng mL^{-1}. Separate fibers manufactured on separate days will have to be tested to establish how much of the descriptor path will be universal and how much specific to a particular taper or coating.

Matrix validation is also required. The calibration performed using PBS confirms the immuno-optic principle, whereas serum or plasma matrixes are expected to provide additional nonspecific adsorption, variability of refractive index, protein crowding, and fouling. These will influence residual descriptors despite constant CEACAM5 level and need to be investigated by proper analytical validation. The reporting rules mitigate this problem by increasing uncertainty at lower descriptor agreement, however, comprehensive analytical validation will have to test recovery, precision, interference, carry-over and stability [30, 31].

The temperature and mechanical stability of the device need also to be carefully controlled. The multimodal nature of the specklegram makes it very sensitive to phase perturbation, so the minor variations of the fiber tension, bending, and imaging setup alignment will result in image alteration. Consequently, practical assay should use fixed mounting, referent monitoring, replica acquisition and automated descriptor quality check. This is an implementation condition of converting sensitive optical phenomenon to analytical measurement.

4. Conclusion

In relation to the research question, we posed whether the concentration of CEACAM5 can be found through antibody-coated tapered multimode microfiber speckles utilizing understandable residual dose-path calibration rather than actual image classification. Yes, this is possible within the range of concentrations examined and with the set of specklegrams used. Residual texture-spectral descriptors retained the concentration ordering within 900 frames between 1 ng mL^{-1} and 1000 ng mL^{-1}, and also had sensitivity to low concentrations along with saturation characteristics.

In particular, the low branch between 1 ng mL^{-1} and 50 ng mL^{-1} generated a slope of -0.00123 response units per ng mL^{-1} and an R^2 of 0.99221, which is a strong evidence for significant optical redistribution caused by early CEACAM5 binding. This complete range gave a saturating curve with an asymptotic value close to 0.82648, a transition concentration around 65.84 ng mL^{-1}, and an R^2 value of 0.99173. All these observations imply that the assay must not be analyzed using

one straight global line. These results need to be presented using two parallel regimes: one that is highly sensitive but operates in the low regime, and another that is saturation aware and operates over a wide range.

The technical innovation consists in transforming specklegrams into an auditable path through residual image construction, spectraltexure feature extraction, concentration ordering, Gaussian process calibration, branching, and uncertainty reporting. The chosen pathway retains the optical relevance of the method while allowing the analyst to accept, repeat, or qualify the measurement. Further research will involve testing independent preparations of the fiber, adding concentrations at the transition region, analyzing serum or plasma, and determining inter-day reproducibility of uncertainty intervals.

References

[1] Gold, P., & Freedman, S. O. (1965). Specific carcinoembryonic antigens of the human digestive system. *The Journal of Experimental Medicine*, *122*(3), 467–481.

[2] Hammarström, S. (1999). The carcinoembryonic antigen (CEA) family: Structures, suggested functions and expression in normal and malignant tissues. *Seminars in Cancer Biology*, *9*(2), 67–81.

[3] Duffy, M. J. (2001). Carcinoembryonic antigen as a marker for colorectal cancer: Is it clinically useful? *Clinical Chemistry*, *47*(4), 624–630.

[4] Locker, G. Y., Hamilton, S., Harris, J., Jessup, J. M., Kemeny, N., Macdonald, J. S., … & Bast, R. C., Jr. (2006). ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *Journal of Clinical Oncology*, *24*(33), 5313–5327.

[5] Sturgeon, C. M., Duffy, M. J., Stenman, U. H., Lilja, H., Brunner, N., Chan, D. W., … & Diamandis, E. P. (2008). National Academy of Clinical Biochemistry laboratory medicine practice guidelines for use of tumor markers in testicular, prostate, colorectal, breast, and ovarian cancers.

[6] Fan, X., White, I. M., Shopova, S. I., Zhu, H., Suter, J. D., & Sun, Y. (2008). Sensitive optical biosensors for unlabeled targets: A review. *Analytica Chimica Acta*, *620*(1–2), 8–26.

[7] Homola, J. (2008). Surface plasmon resonance sensors for detection of chemical and biological species. *Chemical Reviews*, *108*(2), 462–493.

[8] Vollmer, F., & Arnold, S. (2008). Whispering-gallery-mode biosensing: Label-free detection down to single molecules. *Nature Methods*, *5*(7), 591–596.

[9] Estevez, M. C., Alvarez, M., & Lechuga, L. M. (2012). Integrated optical devices for lab-on-a-chip biosensing applications. *Laser & Photonics Reviews*, *6*(4), 463–487.

[10] Leung, A., Shankar, P. M., & Mutharasan, R. (2007). A review of fiber-optic biosensors. *Sensors and Actuators B: Chemical*, *125*(2), 688–703.

[11] Mustapha Kamil, Y., Abu Bakar, M. H., Zainuddin, N. H., Yaacob, M. H., & Mahdi, M. A. (2023). Progress and trends of optical microfiber-based biosensors. *Biosensors*, *13*(2), 270.

[12] Brambilla, G. (2010). Optical fibre nanowires and microwires: A review. *Journal of Optics*, *12*(4), 043001.

[13] Tong, L., Gattass, R. R., Ashcom, J. B., He, S., Lou, J., Shen, M., … & Mazur, E. (2003). Subwavelength-diameter silica wires for low-loss optical wave guiding. *Nature*, *426*(6968), 816–819.

[14] Snyder, A. W., & Love, J. D. (1983). *Optical waveguide theory*. Chapman and Hall.

[15] Teich, M. C., & Saleh, B. (2007). *Fundamentals of photonics* (Vol. 2). Wiley.

[16] Goodman, J. W., & Cox, M. E. (1969). *Introduction to Fourier optics*.

[17] Goodman, J. W. (2007). *Speckle phenomena in optics: Theory and applications*. Roberts and Company Publishers.

[18] Dainty, J. C. (Ed.). (2013). *Laser speckle and related phenomena*. Springer Science & Business Media.

[19] Redding, B., Popoff, S. M., & Cao, H. (2013). All-fiber spectrometer based on speckle pattern reconstruction. *Optics Express*, *21*(5), 6584–6600.

[20] Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-3*(6), 610–621.

[21] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(7), 674–693.

[22] Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, *290*(5500), 2323–2326.

[23] Belkin, M., & Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, *15*(6), 1373–1396.

[24] Hermanson, G. T. (2013). *Bioconjugate techniques*. Academic Press.

[25] Williams, C. K., & Rasmussen, C. E. (2006). *Gaussian processes for machine learning* (Vol. 2, No. 3). MIT Press.

[26] Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4). Springer.

[27] Hastie, T. (2009). *The elements of statistical learning: Data mining, inference, and prediction*.

[28] Unser, M. (1995). Texture classification and segmentation using wavelet frames. *IEEE Transactions on Image Processing*, *4*(11), 1549–1560.

[29] Löfstedt, T., Brynolfsson, P., Asklund, T., Nyholm, T., & Garpebring, A. (2019). Gray-level invariant Haralick texture features. *PLOS ONE*, *14*(2), e0212110.

[30] Armbruster, D. A., & Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. *The Clinical Biochemist Reviews*, *29*(Suppl. 1), S49.

[31] Thompson, M., Ellison, S. L., & Wood, R. (2002). Harmonized guidelines for single-laboratory validation of methods of analysis: IUPAC technical report. *Pure and Applied Chemistry*, *74*(5), 835–855.