Technical white paper

Design and validation of Olink™ Sample Index

Background

Pre-analytical variation encompasses variability introduced during sample collection and handling prior to analysis. In plasma proteomics, differences in time to centrifugation and sample preparation temperature can alter measured protein levels, introducing structured variability that may overlap with biological signals and influence downstream statistical interpretation.

The Olink Sample Index (OSI) is a machine learning-based metric designed to estimate sample handling conditions in EDTA plasma samples analyzed with Olink™ Explore HT. OSI models are tailored to estimate time between sample collection and centrifugation, and the temperature at which the sample was handled during preparation. OSI models generate relative, data-driven metrics that show how closely a sample's protein profile matches patterns seen under controlled handling conditions in a purpose-designed study. These relative metrics assist users with quality assessment, identifying potential causes of unwanted variation, and testing the robustness of downstream analyses, while keeping true biological signals intact and avoiding unnecessary exclusion of samples.

This white paper describes the design of the OSI models and presents validation studies that assess their accuracy, robustness, and reproducibility, providing technical transparency on how the OSI framework was developed and evaluated.

RELATED WHITE PAPER

The practical application, interpretation, and impact of OSI in real-world study settings are described in the related white paper *Olink™ Sample Index: A tool to optimize data quality by addressing pre-analytical variation* (1), available on the Olink website.

Overview of the OSI models

This section provides an overview of the OSI modeling framework, including the structure of the two component models, the underlying training and testing strategy, and their performance.

OSI consists of two independent but complementary machine-learning models, each designed to capture variation from different sources of sample handling. The models detect handling effects through consistent protein patterns. Each model is trained and validated separately, and produces its own parameter-specific value, allowing results to remain interpretable and flexible for analysis. Model training, including feature selection and tuning, was carried out using cross-validation within the training data, and reported performance metrics are based on data subsets that were not used during model development.

Controlled dataset for OSI model training and evaluation

To train and evaluate the OSI models under well-defined conditions, a controlled dataset was generated using samples from 100 study participants and analyzed on Olink Explore HT. For each participant, multiple aliquots were processed under different combinations of time to centrifugation and sample preparation temperature. This resulted in a total of 1,111 complete samples covering both short delays before centrifugation (0.5, 1, 4, and 7 hours) and long delays before centrifugation (24, 48, 72, and 96 hours). The temperature conditions examined were room temperature (RT) and on ice.

A total of 100 samples were excluded from model development and used specifically for testing, serving as the test set and ensuring a clear separation between training and evaluation data.

Feature selection and robustness considerations

The OSI models were trained on a precisely selected set of assays that showed consistent, biologically meaningful responses to pre-analytical variation. Assays were selected using both data-driven methods (using the Boruta algorithm with Random Forest-based feature importance) and biological knowledge, with a focus on features that perform reliably across reagent lots, analytical sites, and validation datasets. This approach reduces the impact of batch-specific effects and helps the models generalize well for real-world data.

The models are also designed to handle realistic levels of missing data, such as assay dropout during quality control. Sensitivity analyses show that OSIs remain stable even when random subsets of predictive assays are missing, supporting their use in routine analytical workflows.

OSI model for time to centrifugation

The Time to Centrifugation model measures how delays between blood draw and plasma separation affect the protein profile. When blood cells remain in contact with plasma for longer periods, consistent changes in protein levels build up over time. The model uses regularized linear regression with custom project-dependent normalization to reduce overfitting by preventing it from capturing random noise in the training data instead of true underlying proteomic patterns. It is trained on controlled data spanning both short and long delays and produces a relative metric between 0 and 1 for each dataset. This scaling is intended for comparing shorter and longer delays to centrifugation within a study, rather than for estimating exact processing times.

When evaluated on the test set, the model showed strong agreement with the lab-recorded centrifugation delays (R^2 = 0.869, Pearson's R = 0.976, Spearman's R = 0.947, Root Mean Square Error = 0.1). The high correlations and low errors across all delay conditions indicate that the metrics reliably capture time-dependent proteomic effects.

OSI model for sample preparation temperature

The Preparation Temperature model estimates whether a sample was handled on ice or at RT before centrifugation, based on its protein profile. Because the controlled training data includes two clear temperature conditions, samples kept on ice and samples kept at RT, the model uses logistic regression to estimate how similar a given sample's protein profile is to either condition. It outputs a number between 0 and 1 where values near 0 indicate on-ice handling, values near 1 indicate RT handling, and intermediate values reflect overlap or uncertainty.

When evaluated on the test set, the model clearly distinguished between on-ice and RT samples (accuracy = 1.0, receiver operating characteristic – area under the curve (ROC AUC) = 1.0, F1 score = 1.0), accurately capturing known temperature effects and demonstrating strong performance on data not used for training.

Summary metric for overall pre-analytical sample handling

While the individual OSI variables provide detailed insight into specific handling effects, overview analyses might benefit from a single summary metric that reflects overall consistency of pre-analytical sample handling.

The combined metric OSI Summary uses a non-linear formula to capture how time and temperature interact, acknowledging that temperature-related protein changes become stronger with longer delays. The combined metric is calculated directly from the independently validated OSIs, rather than from a separate machine-learning model, which maintains transparency and traceability. A categorical version of the combined metric (OSI Category) is also provided to support simpler communication and interpretation by non-expert users.

Validation of OSI

Validation of a model that estimates pre-analytical sample handling conditions against independent controlled studies with predefined handling conditions provides strong confidence in its performance. In addition, such a model must be reproducible across laboratories, robust to assay differences, and applicable to real-world datasets where metadata may be incomplete.

To address these requirements, we performed several complementary validation studies. Two focused on reproducibilty and two focused on accuracy. These included analyses of consistency across sites, robustness to different probe setups, and controlled experiments with lab-recorded handling data.

Together, they provide a comprehensive validation of the OSI models and support their use as a statistically sound and biologically informed approach to capturing pre-analytical effects in proteomic data. The analyses focus on performance, reproducibility, and interpretability, providing confidence in the utility of OSI for multiplex proteomic studies.

Reproducibility validation 1: OSI concordance across multiple laboratories

This analysis evaluates whether OSIs are reproducible when the same samples are processed at different analytical sites under unknown handling conditions.

Study design

A total of 48 EDTA plasma samples were collected, including 30 from donors with various disease states, 10 from healthy individuals, and 8 pooled healthy control samples. Replicate aliquots of the same samples were shipped to five sites and processed independently using the standardized Olink Explore HT protocol. The data were centrally collected, and OSIs were calculated separately for each dataset.

Results

Both Time to Centrifugation and Preparation Temperature OSIs showed strong agreement across the five sites, demonstrating that these proteomic signatures are reproducible and robust to site-specific differences. This strong cross-site agreement is illustrated by the high correlations observed between OSIs derived from data obtained from different sites (Fig. 1).

Reproducibility validation 2: OSI performance across PEA™ probe batches

This analysis evaluates whether OSIs are reproducible when the same samples are analyzed using different Olink Explore HT PEA probe batches. Read more about the PEA probes in the technical note *Consistency with excellence: Lot-to-lot monitoring of Olink PEA probes* (2).

Study design

Fifty EDTA plasma sample replicates were analyzed using five different PEA probe batches. OSIs were calculated separately for each batch's dataset.

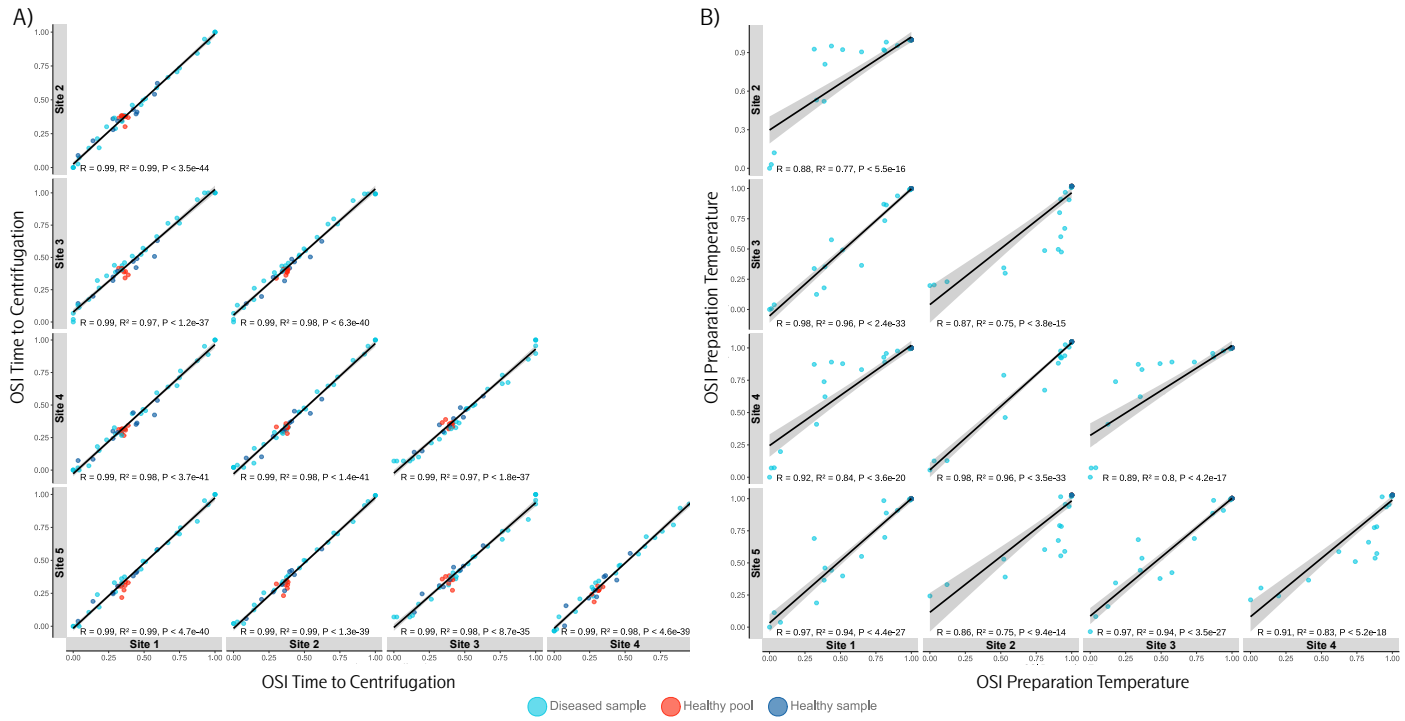

Figure 1. Correlation of OSIs across different analysis sites. Scatter plots show OSIs for A) Time to Centrifugation and B) Preparation Temperature, for the same samples analyzed at five independent laboratories. The linear regression line in each plot represents the agreement of OSI values for pairs of sites, with the shaded area indicating the 95% confidence interval. Pearson's correlation coefficient (*R*), coefficient of determination (R^2), and corresponding *P*-values are reported for each comparison.

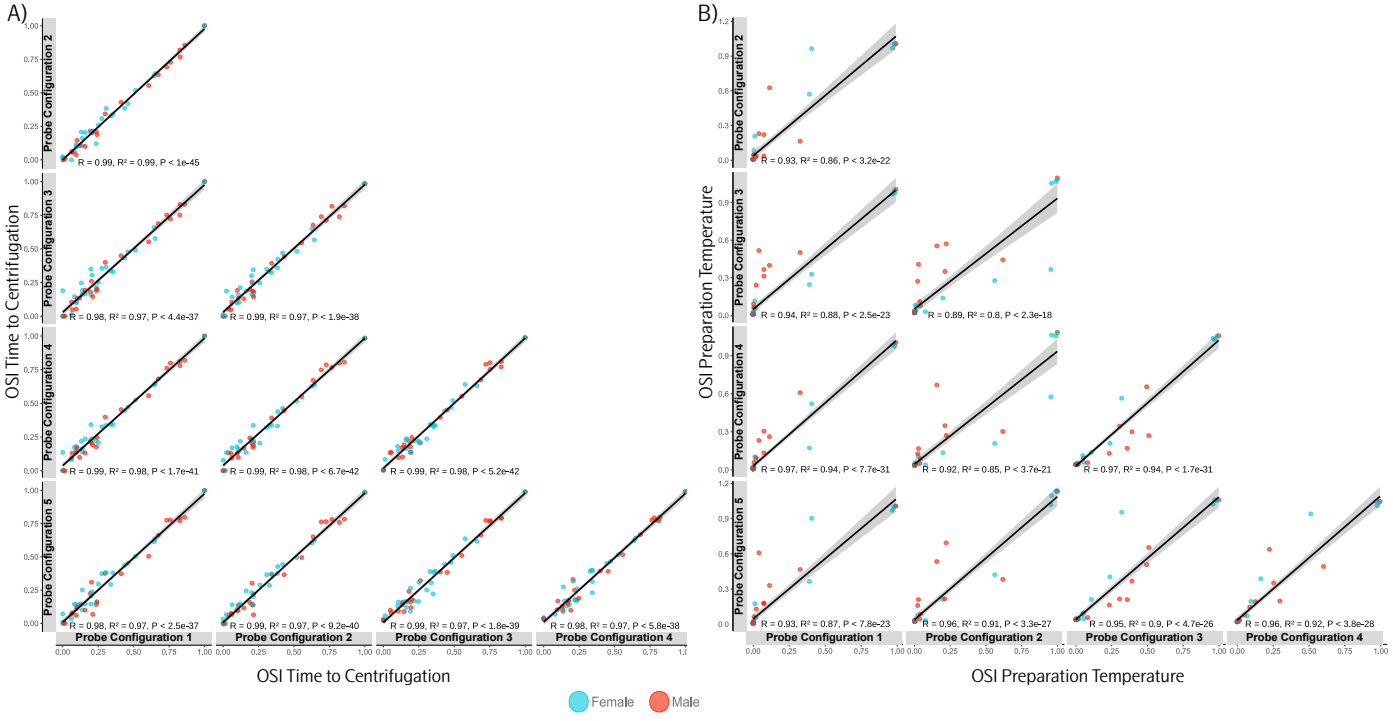

Figure 2. Correlation of OSIs across five different PEA probe batches. Scatter plots show OSIs for A) Time to Centrifugation and B) Preparation Temperature for different probe batches. The linear regression line in each plot represents the agreement of OSI values for pairs of probe batches, with the shaded area indicating the 95% confidence interval. Pearson's correlation coefficient (*R*), coefficient of determination (R^2), and corresponding *P*-values are reported for each comparison.

Results

OSIs for Time to Centrifugation and Preparation Temperature were highly consistent when the same set of samples was analyzed using different PEA probe batches, as illustrated by the strong correlations observed between batches (Fig. 2). OSIs for Time to Centrifugation showed near-perfect agreement across probe batches, and OSIs for Preparation Temperature showed strong concordance.

These results demonstrate that OSI captures proteomic signatures while demonstrating PEA probe reproducibility.

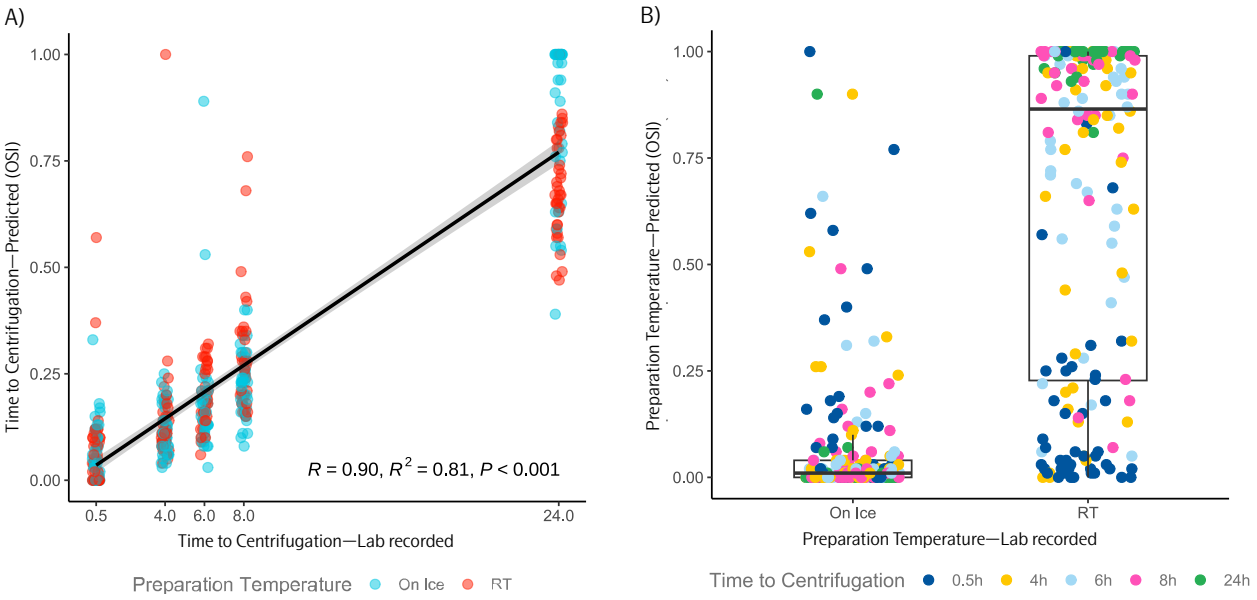

Figure 3. A) Correlation between lab-recorded centrifugation delay and the calculated Time to Centrifugation OSI. Each dot represents a sample, with its color representing true preparation temperature. The linear regression line represents the agreement of OSI values with lab-recorded delays of centrifugation, with the shaded area indicating the 95% confidence interval. Pearson's correlation coefficient (R), coefficient of determination (R^2), and corresponding P-values are reported. B) Distribution of lab-recorded preparation temperatures and the calculated Preparation Temperature OSI. Each dot represents a sample, with its color representing true time to centrifugation. Boxplots are overlaid on the data points to summarize the distribution of OSIs for each preparation temperature group.

Accuracy validation 1: OSI compared to lab-recorded handling conditions

This study evaluates OSI accuracy by comparing OSIs with lab-recorded handling conditions.

Study design

A controlled pre-analytical variation study was carried out in collaboration with deCODE™ genetics using blood samples from 47 healthy donors. Each sample was split into aliquots and processed under predefined handling conditions that varied by two factors: Time to Centrifugation (0.5, 4, 6, 8, and 24 hours) and Preparation Temperature (on ice or at RT).

All samples were collected at a single site and immediately placed either at RT or on ice. Subsequently, different aliquots of the samples underwent pre-determined delays of centrifugation.

Although this represents consistent handling, it differs from typical clinical workflows and from the conditions used to train the OSI Preparation Temperature model, which assumes approximately 30 minutes at RT after collection regardless of subsequent preparation temperature. In total, 391 samples were used in this evaluation.

The aim of this study was to validate that OSIs accurately reflect true laboratory handling conditions and capture known proteomic effects related to delayed centrifugation and temperature differences.

Results

In this controlled setting, OSI accurately captured the major sources of pre-analytical variation examined.

For Time to Centrifugation, OSIs showed strong correlation with the recorded delays (Pearson's $R = 0.9$; $R^2 = 0.81$) and followed the expected trends across all time points (Fig. 3A).

For Preparation Temperature, OSIs reliably distinguished samples handled on ice from those handled at RT and reflected the increasing proteomic impact of temperature exposure with longer delays (Fig. 3B). Minor deviations were observed for short-delay RT samples. ROC AUC for this model was 0.91, highlighting the excellent performance of the model.

Accuracy validation 2: OSI performance under different time and temperature conditions

This independent controlled study further validates OSI performance across a broader range of centrifugation delays in a matched donor cohort.

Study design

A controlled pre-analytical variation study was performed in collaboration with Protavio™ Ltd using blood samples from 37 donors. Participants were well matched for key anthropometric and clinical variables, including age, sex, body-mass index, and health status.

Each sample was aliquoted and processed under predefined handling conditions designed to isolate the effects of two laboratory factors: Time to Centrifugation (0.5, 1, 4, 24, and 72 hours) and Preparation Temperature (on ice or at RT). All samples were collected at a single site, handled under controlled conditions, and analyzed using the Olink Explore HT platform.

For the purpose of validating OSI against true pre-centrifugation effects, 218 samples were included in the analysis. This controlled setup was specifically designed to test whether OSI accurately reflects known laboratory handling conditions.

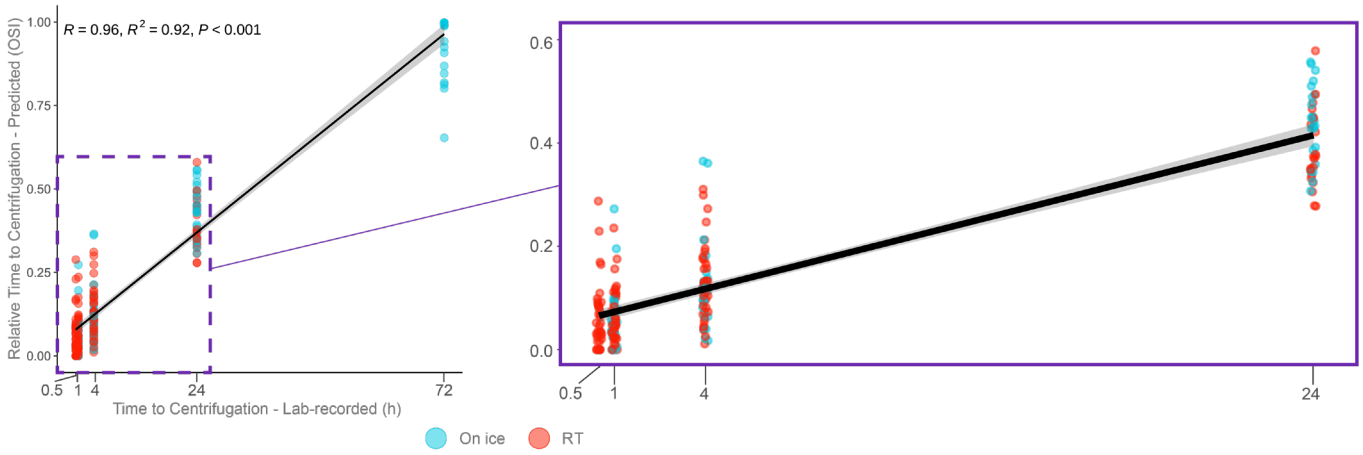

Figure 4. Correlation between lab-recorded centriugation delay and the calculated Time to Centrifugation OSI. Each dot represents a sample, with its color representing true preparation temperature. The linear regression line represents the agreement of OSI values with lab-recorded delays of centrifugation, with the shaded area indicating the 95% confidence interval. Pearson's correlation coefficient (*R*), coefficient of determination (R^2), and corresponding *P*-values are reported. The zoomed-in view allows clearer visualization of the correlation at shorter centrifugation delays, where data points overlap in the full-scale plot.

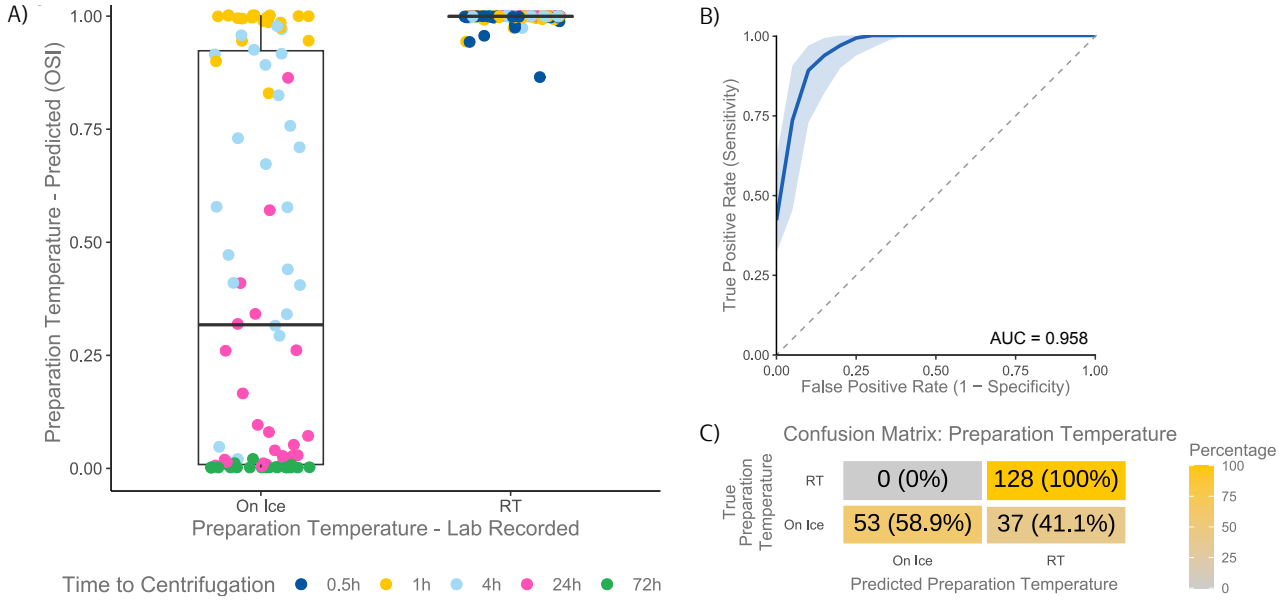

Figure 5. A) Distribution of lab-recorded preparation temperatures and the calculated Preparation Temperature OSI. Each dot represents a sample, with its color representing true time to centrifugation. Boxplots are overlaid on the data points to summarize the distribution of OSIs for each preparation temperature group. B) Receiver operating characteristic (ROC) curve evaluating the ability of Preparation Temperature OSIs to distinguish between lab-recorded preparation temperatures (on ice vs. RT). The area under the curve (AUC) value is shown in the bottom-right corner. C) Confusion matrix showing the performance of Preparation Temperature OSIs in classifying samples as on ice or RT based on lab-recorded conditions. A threshold of 0.5 in the OSI predictions is used for classification. Samples with OSI values below 0.5 are classified as handled on ice, and samples with OSI values of 0.5 or higher are classified as handled at RT. Each cell shows the number and percentage of samples for each predicted and laboratory-recorded preparation temperature.

Results

In this controlled evaluation, OSI accurately captured both major sources of pre-analytical variation.

Time to Centrifugation OSIs showed a very strong correlation with recorded laboratory delays (Pearson's $R = 0.96$; $R^2 = 0.92$) and followed the expected consistent increase across all time points, confirming that OSI reliably reflects the impact of delayed centrifugation (Fig. 4).

Preparation temperature effects were well captured. OSI clearly distinguished samples handled at RT from those handled on ice, with RT samples showing consistently higher OSIs and excellent classification (Fig. 5A and 5C). Samples handled on ice were generally classified correctly, with those with the shortest centrifugation delays being misclassified. Nevertheless, the ROC AUC for the OSI Preparation Temperature model was excellent (AUC = 0.958) (Fig. 5B).

Taken together, these results suggest that OSI effectively captures the impact of preparation temperature on sample quality.

Validation of OSI Summary and OSI Category

In the controlled deCODE™ genetics validation dataset, samples were intentionally exposed to defined delays before centrifugation and different preparation temperatures. In this

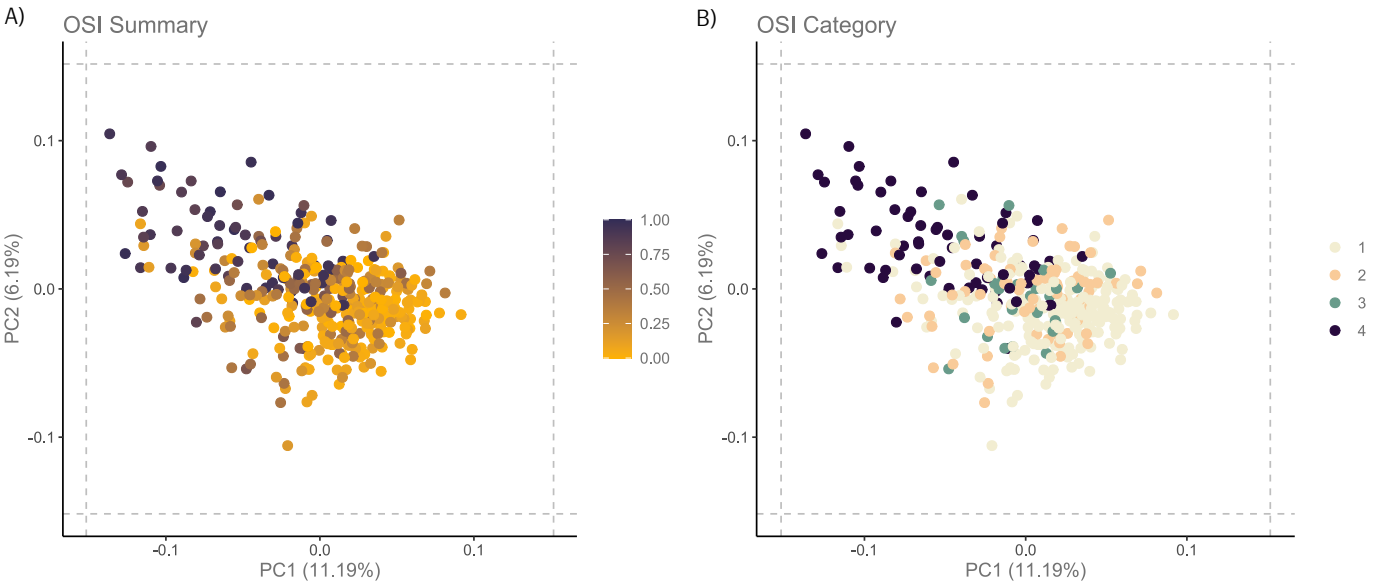

Figure 6. Projection along the first two principal components from a principal component analysis (PCA) performed on the NPX values of the dataset. Each dot represents a sample, with its position based on the measured protein values and its color based on the value for A) OSI Summary, and B) OSI Category. Dashed lines indicate ±3 standard deviations from the mean value for each component.

setting, OSI Summary shows a clear gradient across the principal component analysis (PCA) space, with the lowest OSI values in the bottom right and the highest in the top left (Fig. 6A).

This pattern directly reflects the known study design. The continuous increase in OSI across the PCA projection demonstrates that the summary OSI variables capture systematic deviations introduced during sample handling, rather than random variation. Samples handled under similar controlled conditions cluster together, confirming that OSI identifies meaningful handling-related structure in the proteomic data.

The categorical variable, OSI Category, simplifies this gradient into interpretable groups, making the results easier to communicate to non-expert audiences (Fig. 6B).

Similar examples showing the effects of consistent versus inconsistent handling on data structure are presented in the related educational white paper (1). In this dataset, the observed gradient provides direct validation that OSI reflects known laboratory conditions in real data. Together, these results demonstrate the value of the summary OSI variables. They reliably identify and quantify handling-related deviations where such structure is expected to occur.

Summary and conclusions

In this white paper, we evaluated OSI across a series of controlled scenarios designed to test both its reproducibility and accuracy. Across multiple analyses, OSI shows consistent and biologically coherent behavior in measuring the proteomic effects of pre-analytical variation in EDTA plasma samples analyzed with Olink Explore HT. Controlled studies with recorded laboratory handling data show that OSIs follow the expected patterns for Time to Centrifugation and Preparation Temperature, confirming that OSI is sensitive to the main sources of pre-analytical variation it was designed to capture.

A notable observation is that the preparation-temperature model shows most uncertainty at short centrifugation delays. This reflects the limited biological impact of temperature differences when delays are brief, making such samples inherently harder to distinguish. Importantly, this reduced discrimination ability does not compromise downstream interpretation, as proteomic changes between refrigerated and RT handling are modest at short delays. This is consistent with prior literature showing that short pre-centrifugation delays produce fewer significant protein changes and smaller effect sizes compared to prolonged delays (3).

Beyond controlled validation, OSI also replicates well across different settings that may introduce batch-related variations. Multi-site analyses show that replicate samples processed at different laboratories receive highly similar OSIs, demonstrating robustness across sites. Similarly, analyses across different Explore HT PEA probe batches show stable OSIs, indicating that OSI reflects true proteomic handling signatures rather than probe-specific effects.

Overall, these results demonstrate that OSI provides a reliable and biologically grounded framework for assessing pre-analytical variation in Explore HT data. When utilized and interpreted appropriately, OSI can support outlier detection, investigation of handling effects, and statistical adjustment in multiplex proteomic studies as exemplified in the educational white paper (1).

Limitations

As with any model-based approach, OSI carries limitations that should be considered when interpreting results and applying to downstream analyses.

OSI estimates pre-analytical variation from protein patterns and does not replace laboratory-recorded handling metadata, which should be used when available. The models are trained on a specific range of handling conditions, and applying them outside this range may limit OSI interpretation. Additionally, OSI is validated only for EDTA plasma and should not be used (or should be interpreted with caution) for other sample matrices.

Reproducibility analyses cannot detect biases that are shared across datasets, and OSIs may partly reflect biological effects in studies where sample handling is linked to the biological variable of interest. For these reasons, OSI should be used as a supportive tool, not as a standalone measure for data quality or sample exclusion.

References

1. Olink white paper Olink™ Sample Index: A tool to optimize data quality by addressing pre-analytical variation (2026).
2. Olink technical note Consistency with excellence: Lot-to-lot monitoring of Olink PEA probes (2021).
3. Shen *et al.* Strong impact on plasma protein profiles by precentrifugation delay but not by repeated freeze-thaw cycles, as analyzed using multiplex proximity extension assays. *Clin Chem Lab Med.* (2018)

www.olink.com

1692, v1.0, 2026-03-05