Bevital´s Olink FAQ

1. **What is NPX?**

NPX, Normalized Protein eXpression, is Olink's arbitrary unit which is in Log2 scale. It is calculated from Ct values and data pre-processing (normalization) is performed to minimize both intra- and inter-assay variation. NPX data allows users to identify changes for individual protein levels across their sample set, and then use this data to establish protein signatures.

The NPX scale is inverted compared to that of Ct. This means that a high NPX value equals a high protein concentration. Because NPX is in a log2 scale, a 1 NPX difference means a doubling of protein concentration. If needed NPX values can be converted into linear scale: 2^{NPX}= linear NPX.

NPX is derived from the Ct values obtained from the qPCR using the following equations:

Extension Control:
CtAnalyte – CtExtension Control = dCtAnalyte

Inter-plate Control:
dCtAnalyte – dCtInter-plate Control = ddCtAnalyte

Adjustment against a correction factor:
Correction factor – ddCtAnalyte = NPXAnalyte

NPX can be used for ***relative quantification only.*** This means:

- NPX values can be compared only for the same protein across the samples analyzed in your project
- NPX values cannot be compared across projects run at separate occasions without the use of reference bridging samples

For more information about relative quantification use this link.

2. **How does the data file look like?**

• Concentrations are reported as NPX values for T96 or as absolute concentrations in pg/ml for T48 and Flex
• Values below maximum plate LOD have a red cell background
• Data for samples with a QC warning are indicated in red text. Data from samples that do not pass QC should be treated with caution. Data are manually checked using a PCA and outliers with QC warnings are removed from the data and marked as "missing". These samples are listed in the analysis report
• The maximum plate LOD value for each assay is presented on a separate row below the data for the samples and is indicated as LOD
• Missing data frequency is presented for each assay and indicates the percentage of samples with values below Maximum plate LOD

3. What is LOD?

Limit of detection (LOD) is calculated separately for each Olink assay and sample plate and presented in the output data file. The LOD is based on the background, estimated from Negative Controls included on every plate, plus three standard deviations. The standard deviation is assay specific and estimated during product validation for every panel.
LOD may be informative for technical evaluations, including CV calculations where it is recommended to base CV calculations on data > LOD. As data below LOD may be non-linear this data should be interpreted cautiously.

However, it is recommended to include data <LOD in downstream statistical analysis. Indeed, LOD is less important in downstream statistical analyses as values under LOD typically converge across groups. As such, including data below LOD is unlikely to increase the risk of false positive discoveries. Furthermore, data below LOD can be instrumental in downstream analyses such as biomarker discovery as a protein may be well expressed in one group and not measured in another group. In this case, this protein can be a strong biomarker candidate for specific groups, while filtering data >LOD for the statistical analysis would not allow for such findings. Especially in large multi-plate studies LOD is a conservative measurement. Using actual data may increase the statistical power and gives a more normal distribution of the data. Including data under LOD does commonly not increase false positives as there is generally no significant difference between groups under LOD (values tend to be condensed to a very small range).

4. What is detectability?

Detectability is defined as number of proteins detected in >75% of the samples. The *expected* detectability given in the analysis report is based on EDTA plasma from healthy donors. These values are intended as guidelines only and protein levels (as well as detectability) may vary depending on different pathological conditions, sample matrices, or sample preparation methods.

Some of the proteins addressed by our panels have a very broad range of expression levels in vivo, which makes it difficult to develop an assay in multiplex that can measure both healthy and diseased individuals.

For example, some of the assays on the CVD-panels are developed mainly to measure the elevated levels that can be seen after or during cardiovascular disease. If you would still like to use the results from an assay with low detectability you can use non-parametric statistics (i.e. detected vs. not detected in your groups).

5. What are QC warnings?

Each of the internal controls is spiked into all samples at a set concentration. The signals for these are therefore expected to be the same over the entire plate. Sample QC is performed using the Detection Control and Incubation Control 2. Within each run, the levels of these controls are monitored for each sample and compared against the plate median of all samples. If either one of the controls deviate more than the acceptance criteria allow (see below), the sample gets a QC warning and the sample is flagged in NPX Signature. The Extension Control is used in the normalization step and in generation of NPX, and hence is not included in the quality control of data. Acceptance criteria for passing a sample:

• Incubation Control 2 and Detection Control deviates < ±0.3 NPX from the plate median. Deviating values for the internal controls can be caused by, for example, errors in pipetting or pre-analytical factors in the samples that affect the performance of the controls.

6. How is precision calculated?

%CV is calculated using external control samples in duplicates on each sample plate. The reported %CV is the mean %CV over all assays, and this is only calculated using data over LOD. A high %CV does not fail a run automatically but should be a cause for further investigation.

• Reference value for Inter %CV: < 25%
• Reference value for Intra %CV: < 15%

7. How are results normalized between plates?

NPX Signature supports three methods for between-plate normalization. Between-plate normalization adjusts median NPX levels based either on customer samples or IPC controls. An important concept when deciding on the normalization procedure is randomization, which in this context applies to the sample placement across the plates. For randomized studies with more than one plate, sample intensity normalization is the default normalization. For other studies, IPC normalization is the default. (See Randomization FAQ on the Olink website.) Olink uses an arbitrary, relative quantification unit called Normalized Protein eXpression (NPX). In qPCR, the x-axis value of the point where the reaction curve intersects the threshold line is called the Cycle threshold (Ct). This indicates the number of cycles needed for the signal to surpass the fluorescent signal threshold line.

8. How is absolute quantification performed?

Repeated testing and validation have shown that 4PL curve fitting describes the standard curves well and can be used to correctly estimate the protein concentration in analyzed samples within the limits of quantification. The approximate lower and upper limits of quantification (ULOQ and LLOQ) are defined during the development of the panel, and specifications for LLOQ and ULOQ are established for each lot. This information is available in the individual Panel Validation Reports provided by Olink. Results are reported in pg/mL. The list below provides an overview of how data in pg/mL is reported in the results file:

• Data between LQL and ULOQ is reported as a pg/mL value. NOTE: Lowest Quantifiable Level (LQL) is defined as the value used as the lower limit, LLOQ (default) or plate LOD (when plate LOD > LLOQ)
• Data below the lowest fitting parameter in the 4PL curve fit model cannot be calculated and is indicated as NaN
• Data >ULOQ is indicated as >ULOQ and are not reported in pg/mL due to high risk of misinterpretation of hooking data
• Data below LQL is presented in pg/mL value and values below LQL should be treated with caution due to decreased precision and accuracy in the lower range and should not be used for individual comparison to reference values

More information regarding the use of relative NPX values and absolute quantification can be found here.

9. How can I analyze the data?

Information regarding data analysis can be found here:

- Olink Analyze which can be found on Olink´s main website > Products. This is the R-package that is freely available on CRAN. You can the COVID-19 data story on Insight to familiarize yourself with the R-package and how to best use Olink Data.

- The Stat Analyzer tool on Insight which also uses the R-package, but has limited options. It is a great start if you are not so familiar with R.

Contact us at olink@bevital.no or use Olink´s faq page for further question.

You can also try our new Olink chatbot for quick response: