SUPPLEMENTAL MATERIALS

This Supplement contains additional information concerning the following aspects:

1. Further information from the statistics obtained in the study
2. Summary of the analytical performance and usability claims of the devices used in the study: Evernoa® from Eversens® and Niox Vero® from Circassia
3. Parameters definition according to Clinical and Laboratory Standards Institute

Supplement Abbreviations

FENO. Fractional concentration of exhaled Nitric Oxide

NO. Nitric oxide

Ppb. Parts per billion

ATS. American Thoracic Society

ERS. European Respiratory Society

R^2. Squared correlation coefficient

SD. Standard deviation

CV. Coefficient of variation

MEN. Mean of values obtained with evernoa

MNX: Mean of values obtained with Niox Vero

ICs. Inhaled corticosteroids

Statistical analysis

Deming regression is an appropriate method to compare the performance of two measurement procedures when both are subject to error. According to the specifications provided by the manufacturers (see section Analytical performance and usability) evernoa and Niox introduce error not necessarily constant, so the weighted Deming regression was selected for analyzing the obtained results.

As part of the process, the measurement error ratio(λ)was estimated from two studies of the reproducibility of data obtained with both devices. The results obtained are shown in the following figure:
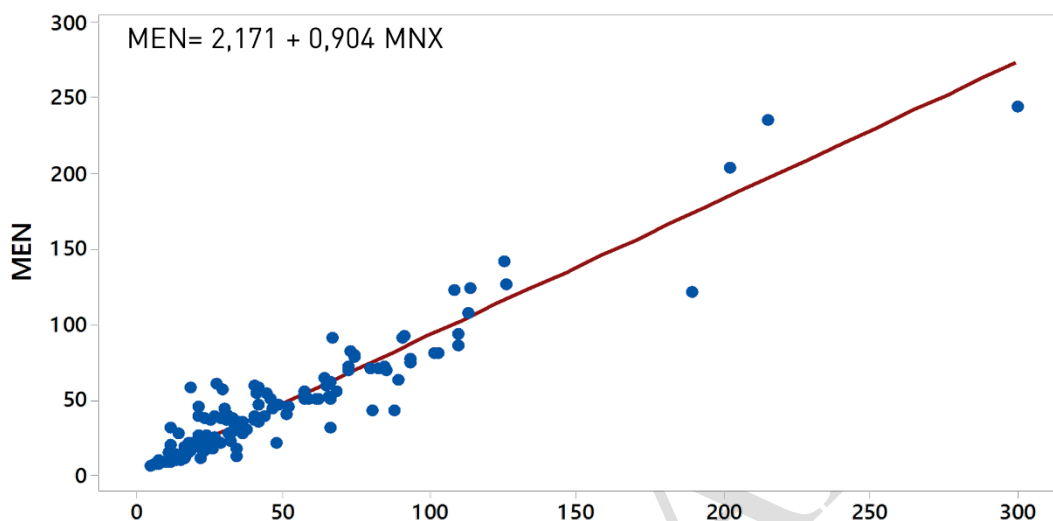

Figure 1. Weighted Deming Regression evernoa vs Niox Vero

The Deming Regression gives a slope of 0.904 I.C. (0.851; 0.956) and an intercept of 2.71 I.C. (-1,182; 5,523). Results obtained with evernoa areslightly lower than those of NIOX, although there is a high degree of concordance between the two devices. Additionally, as it can be seen in the fig.1 differences of paired measurements are higher as the FENO concentration increases. This result is consistent with the results obtained in other studies when electrochemical based devices are used to measure FENO.

According to the international guides, the level of greatest clinical relevance when accomplishing the measurement of FENO in adults is 50 ppb, as it is considered the cut-off point to determine the likelihood of the presence of airway inflammation and of the response to ICs. The bias obtained at this level is -2,634 ppb revealing the similarity of measures obtained with evernoa and Niox Vero.

Analytical performance and usability claims

Analytical performance

A summary of the analytical performance parameters is presented in the following table:

	Measuring range	Precision	Trueness	Limit of detection	Linearity
evernoa	5-300 ppb	<2ppb FENO <50 ppb ≤3% FENO ≥ 50 ppb	±4 ppb FENO ≤ 50 ppb ±8% FENO >50 ppb	5 ppb	R² ≥ 0.998 Slope 0.978 Intercept -1.408
Niox Vero	5-300 ppb	<3ppb FENO <30 ppb <10% FENO ≥ 30 ppb	±5ppb FENO ≤50 ppb ±10% FENO > 50 ppb	5 ppb	R² ≥ 0.998 Slope 0.95-1.05 Intercept ± 3 ppb

Table 1. Analytical performance parameters stated by manufacturers

Measuring range:

Both devices state the same measuring range (5-300 ppb). According to the levels of FENO obtained in the study, evernoa and Niox Vero are suitable for the study purposes.

Trueness

For both devices., trueness is expressed as the upper 95% confidence limit, based on absolute mean of differences from certified gas concentration of NO

The bias for FENO ≤ 50 ppb is ± 4 ppb for evernoa and ± 5 ppb for Niox Vero, while in the range over 50 ppb, they are ± 8% and ± 10 % respectively.

Different guidelines have defined a cut-off point (50 ppb for adults and 35 ppb for children) as the decision point to assess the likelihood of airway inflammation when using FENO as a biomarker. Because of that, trueness has a higher impact in the diagnosis phase.

Precision

Precision is expressed as one standard deviation for replicate measurements with the same instrument, using a certified gas concentration of NO reference standard.

The imprecision for evernoa is less than 2 ppb below 50 ppb and less or equal to 3% for FENO values over 50 ppb. Niox Vero exhibits an imprecision of 3ppb of measured value for values < 30 ppb, less than 10% of measured value for values ≥ 30 ppb.

For the monitoring over the time, variations of FENO over the time are considered a useful tool in asthma management. In this sense "minimally important decrease of the FENO value is defined as a difference larger than 20% for values over 50ppb or more than10ppb for values lower than 50ppb from one visit to the next". Devices with higher precision would provide a more accurate information in the monitoring phase.

Linearity was similar in evernoa and Niox Vero obtaining a linearity of Squared correlation coefficient $r^2 \geq 0.998$.

Lowest Detection limit

Determined by analyzing gas concentrations around and below the detection limit., 5 ppb was the lowest detectable level for both devices

Usability claims

In the following table the most important features concerning the usability of the devices are included.

	Inhalation through the device	Exhalation time	Response time
evernoa	Not needed	Adults: 10 s Children: 10 s	5 s
Niox Vero	Required	Adults: 10 s Children: 10 s	≈ 60 s

Table 2. Usability features stated by the manufacturers

Evernoa and Niox Vero follow the recommendations provided by the ATS and ERS for the measurement of FENO and thus, the exhalation time needed to successfully complete themanoeuvreis 10 seconds for adults and 6 seconds for children.

The major difference is that Niox requires an inhalation through the device to reach the total lung capacity while with evernoa the patient inhales directly ambient air.

Another difference is that evernoa has a lower response time (5 seconds) whilst Niox requires around 60 seconds to accomplish the measurement process.

Parameters definitions

Measuring range*:* A set of values of measurands for which the error of measuring instrument is intended to lie within specified limits.

Accuracy*:* Closeness of agreement between a measured quantity value and a true quantity value of a measurand.
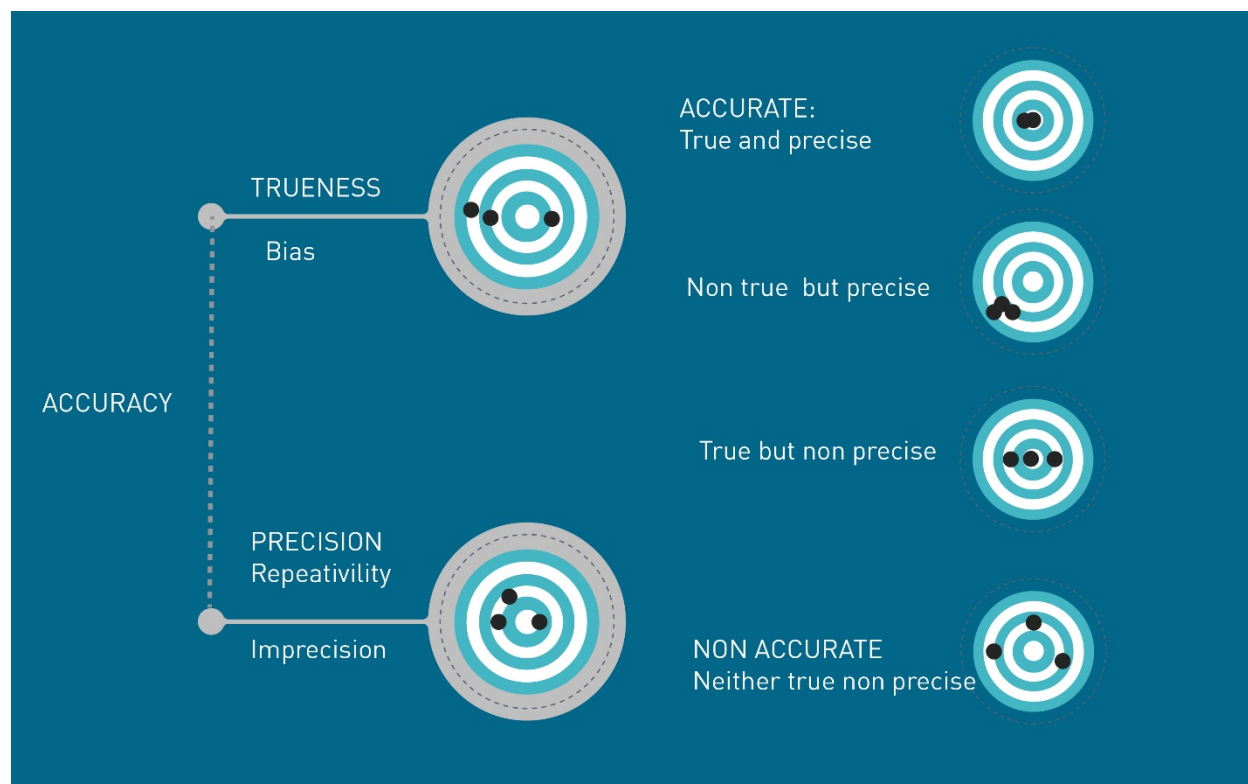
Figure 2. Application of accuracy terms

Trueness*:* The closeness of agreement between the average valued obtained from a large series of test results and an accepted reference value.

Bias*:* The difference between the expectation of the test results and an accepted reference value.

Precision- Repeatability*.* The ability of a measuring instrument to provide closely similar indications for repeated applications of the same measurand under the same conditions of measurement.

Imprecision*:* The random dispersion of a set of replicate measurements and/or values expressed quantitatively by a statistic, such as SD or CV.

Linearity: The ability (within a given range) to provide results that are directly proportional to the concentration of the analyte in the test sample.
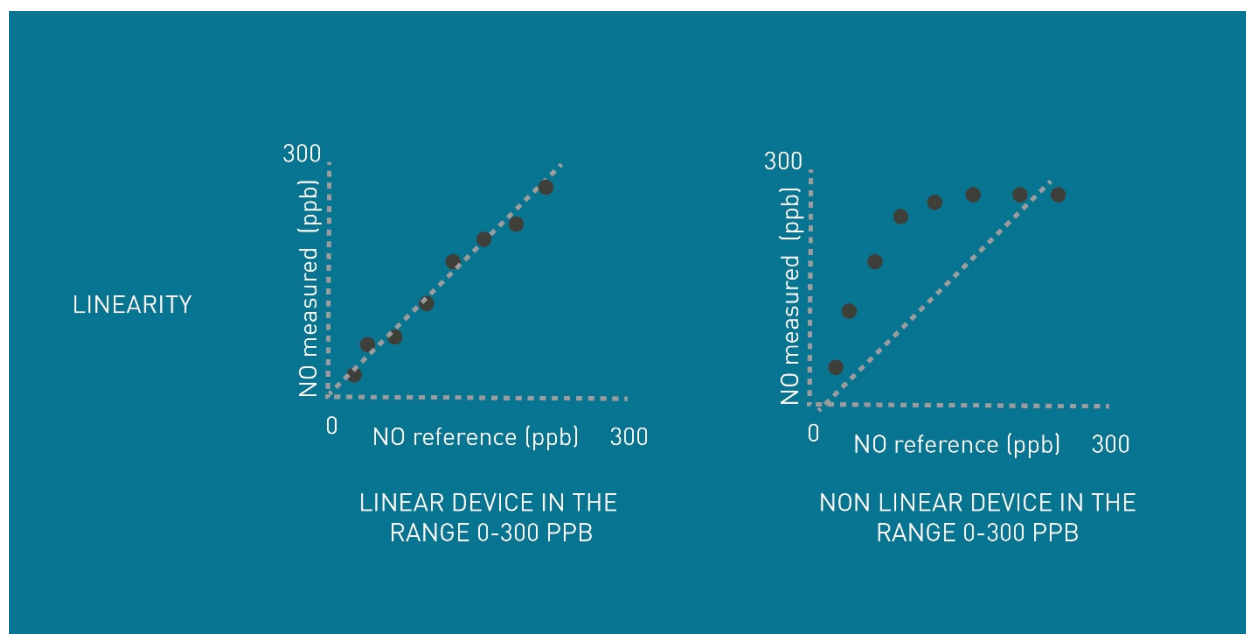
Figure 3. Example of a linear and a nonlinear system

Lowest detection limit: The lowest concentration of analyte that can be consistently detected (typically, in ≥ 95% of samples tested under routine clinical laboratory conditions and in a defined type of sample).