Research Article

Validation of the Monocyte Activation Test with Three Therapeutic Monoclonal Antibodies

Ruth Daniels[1], Wim Van der Elst[2], Nele Dieltjens[1], Tinne Appels[1], Chi K. So[4], Thomas Nys[3], Liesbeth Voeten[1], Philip Breugelmans[1], Marijke W. A. Molenaar-de Backer[5], Eelo Gitz[6], Stephen Poole[7] and Mehul Patel[4]

[1]Microbiology CoE, Janssen R&D BE, Beerse, Belgium; [2]Nonclinical Statistics, Janssen R&D BE, Beerse, Belgium; [3]Discovery Biology, Janssen R&D BE, Beerse, Belgium; [4]Large Molecule Analytical Development, Janssen R&D US, Malvern, PA, US; [5]Sanquin Diagnostics, Dept Virology and MAT services, Amsterdam, The Netherlands; [6]Sanquin Reagents, Amsterdam, The Netherlands; [7]Independent consultant, London, United Kingdom

Abstract

Pharmaceutical products intended for parenteral use must be free from pyrogenic (fever-inducing) contamination. Pyrogens comprise endotoxins from Gram-negative bacteria and non-endotoxin pyrogens from Gram-positive bacteria, viruses, and fungi. The longstanding compendial test for pyrogens is the rabbit pyrogen test, but in 2010 the monocyte activation test (MAT) for pyrogenic and pro-inflammatory contaminants was introduced into the European Pharmacopoeia (*Ph. Eur.*) as a non-animal replacement for the rabbit pyrogen test. The present study describes the first product-specific Good Manufacturing Practice validation of *Ph. Eur.* MAT, Quantitative Test, Method A for the testing of three therapeutic monoclonal antibodies. The study used the MAT version with cryo-preserved peripheral blood mononuclear cells and interleukin-6 as the readout. Much of the data presented here for one of the antibodies was included in a successful product license application to the European Medicines Agency.

1 Introduction

The monocyte activation test (MAT) for pyrogenic and pro-inflammatory contaminants was first used in medicines testing as far back as 1988 (Poole et al., 1988) but did not become a compendial method until 2010, when it was introduced into the European Pharmacopoeia (*Ph. Eur.*) as a non-animal replacement for the rabbit pyrogen test (RPT) (EP, 2017).

Although the bacterial endotoxin test (BET) (EP, 2012) has largely superseded the RPT (EP, 1986) for routine testing of pharmaceuticals for pyrogens, the BET is a test for only the most frequent pyrogenic contaminant of medicinal products, namely bacterial endotoxin (which comprises largely lipopolysaccharide, LPS). The BET is also called the Limulus amoebocyte lysate (LAL) test and relies on the property of the lysate of amoebocytes from the blood of horseshoe crabs, *Limulus polyphemus*, to clot in the presence of endotoxin from the cell walls of Gram-negative bacteria. Consequently, it is a regulatory requirement that new products are shown to be free from any non-endotoxin pyrogens (NEPs) before the BET can be adopted as the sole pyrogen/endotoxin test for routine product testing. In Europe and other countries, but not the USA, this means applying the MAT to at least the initial 3 process performance qualification (PPQ) batches, while in the USA this means applying the RPT to these batches.

The basis of the MAT is the activation of human monocytes or monocytic cells to release endogenous mediators such as pro-inflammatory cytokines, e.g., tumor necrosis factor alpha, interleukin-1 beta (IL-1β), and IL-6 upon exposure to pro-inflammatory and pyrogenic contaminants. The released cytokines have a role in fever pathogenesis and other inflammatory responses. Consequently, the MAT detects the presence of pyrogenic/pro-inflammatory contaminants in the test sample by measurement of one of these pro-inflammatory cytokines.

The MAT differs from both the RPT and BET in a number of important ways: The RPT is a limit test with no standard and is intended to detect endotoxin and NEPs. In contrast, the BET detects only endotoxins and can be carried out as a limit test,

Received November 30, 2021; Accepted March 18, 2022;
Epub April 14, 2022; © The Authors, 2022.

ALTEX 39(4), 621-635. doi:10.14573/altex.2111301

Correspondence: Ruth Daniels, PhD
Discovery, Product Development & Supply
Analytical Development, Microbiology CoE
Turnhoutseweg 30 (P.O. 309), 2340 Beerse, Belgium
(rdanie22@its.jnj.com)

quantitative test, or semi-quantitative test, and there is an internationally agreed standard endotoxin that is used to measure the pyrogenic contamination in the product in endotoxin units (EU), with an endotoxin limit concentration (ELC), given in EU/mL, specified for each product. The MAT described in *Ph. Eur*. 2.6.30 can be carried out as a quantitative test (Method A), semi-quantitative test (Method B), or as a reference lot comparison test (Method C); there is no limit test variant. In Methods A and B, a standard curve of the international standard for endotoxin is used to quantify the (pyrogenic) contamination in the product in endotoxin equivalents (EE), with a contaminant limit concentration (CLC), given in EE/mL, specified for each product. Thus, the MAT "re-badges" the ELC as the CLC with the units changed from EU/mL to EE/mL. This is necessary because the BET detects only endotoxin whereas the MAT detects endotoxin and monocyte-activating contaminants other than endotoxin. The latter are a physico-chemically diverse class of molecules. Here, for convenience, we use the term NEP to describe non-endotoxin ligands for toll-like receptors (TLRs) and any other non-endotoxin monocyte-activating contaminant. However, it should be noted that NEP, while stimulating the release of pyrogenic/pro-inflammatory cytokines, may not always cause fever, e.g., the peptidoglycan (PGN) that contaminated the dialysis product Extraneal caused cytokine release *in vivo* leading to adverse reactions that did not include fever (Martis et al., 2005).

Until recent years, the wider application of the MAT had been hampered by a lack of commercial sources of human monocytes/monocytic cells, but this has now changed with a number of vendors offering cryo-preserved cells and even complete MAT kits.

Following an evaluation of various cell sources and reagents, it was decided to carry out product-specific GMP validations of the *Ph. Eur*. MAT with three therapeutic mAbs using human cryo-preserved peripheral blood mononuclear cells (cryo-PBMC), each PBMC lot comprising cells pooled from 4 donors. Experiments prior to the study had established that there was no averaging effect of pooling cells, i.e., the cells of individual donors (to a pool of 4 donors) gave responses that were similar to the responses of the pooled cells.

At the time that this study was initiated, *Ph. Eur*. 2.6.30 prescribed Method A Quantitative Test as the first of the three MAT methods with which to begin a product-specific GMP validation. In addition to the experiments required to demonstrate compliance with *Ph. Eur*. 2.6.30 MAT Method A, i.e., a method verification since pharmacopeial methods are considered validated, additional experiments were carried out to evaluate the overall robustness of the chosen MAT and to carry out the first product-specific GMP validation of Method A.

2 Materials and methods

There are two pharmacopeial acceptance criteria for the standard curve (EP, 2017):

(i) the regression of responses (appropriately transformed if necessary) on $\log_{10}$ dose shall be statistically significant ($p < 0.01$);
(ii) the regression of responses on $\log_{10}$ dose must not deviate significantly from linearity ($p > 0.05$).

Also, there are pharmacopeial requirements for the limit of detection (LOD; mean of blank + 3x standard deviation) and the basal release of the readout (< 0.1 optical density, OD units), which was IL-6 in this MAT (EP, 2017). LOD was defined as the concentration (EE/mL) that corresponded to the mean OD of four blanks (assay medium + cells) plus 3x their SD, i.e., mean OD of blank + 3x SD of OD of blanks; this is called the cut-off value.

In addition to verifying that these pharmacopeial requirements were satisfied by the MAT described here, several assay parameters based on the ICH Q2 R1 guideline were monitored to provide additional information about the MAT. For standard endotoxin and four NEPs (three TLR ligands and one nucleotide-oligomerization domain 2 (NOD2) ligand), these were: specificity, accuracy, precision, linearity, range, limit of quantification (LOQ), robustness, standard curve properties/parameters. The following parameters of the endotoxin standard curves were monitored: Squared correlation coefficient (R^2), lower asymptote, upper asymptote, slope, EC_{50}, and the ranges of the confidence intervals (CI) for lower asymptote, upper asymptote, and EC_{50} estimated using a 4-parameter logistic curve (4PL). In addition, the tolerance intervals (TI) for EC_{50}, range of CI of EC_{50}, and range of CI of upper asymptote were also determined. Data from the 39 standard curves obtained in this study and data from an additional 21 standard curves from other (contemporaneous) MATs that also complied with both acceptance criteria of the *Ph. Eur*. were combined and analyzed (60 standard curves in total) to determine which parameters, in addition to the two stated in *Ph. Eur*., might be used to monitor the endotoxin standard curves in future experiments. The robustness of the method was challenged by varying one of: PBMC lot, incubator temperature, CO_2 concentration, cell number/well, or incubation period.

The MAT assays were carried out using international standard endotoxin (LPS) and four different NEPs. Each MAT plate contained a 7-concentration curve of standard endotoxin, a blank, and a PGN positive control. Four replicates of each sample were tested. Dixon's test for outliers was used to identify outliers within the four replicates. It should be noted that all concentrations referred to in this study are nominal test concentrations in each well.

Abbreviations
4PL, 4-parameter logistic curve; BET, bacterial endotoxin test; CI, confidence interval; CLC, contaminant limit concentration; EE, endotoxin equivalents; ELC, endotoxin limit concentration; EMA, European Medicines Agency; EU, endotoxin units; FLA-ST, flagellin from *Salmonella typhimurium*; GCV, geometric coefficient of variation; GMP, Good Manufacturing Practice; HKSA, heat-killed *Staphylococcus aureus*; IL, interleukin; LAL, Limulus amoebocyte lysate; LOD, limit of detection; LOQ, limit of quantification; LPS, lipopolysaccharide; mAbs, monoclonal antibodies; MAT, monocyte activation test; NEP, non-endotoxin pyrogens; NOD, nucleotide-oligomerization domain 2; OD, optical density; PAM, synthetic triacylated lipoprotein (PAM_3CSK_4); PBMC, peripheral blood mononuclear cells; PGN, peptidoglycan; *Ph. Eur*., European Pharmacopoeia; PPQ, process performance qualification; RPT, rabbit pyrogen test; TI, tolerance intervals; TLR, toll-like receptor

Critical reagents

MAT Cell Set (Sanquin, REF: M2016), using the cryo-PBMC from the kit (pMAT Cells) and human AB serum as culture media supplement; Iscove's Modified Dulbecco's Media (IMDM; 40 mL from a new, unopened bottle) from Lonza (REF: BE12 722F); endotoxin reference standard that had been calibrated against the international standard, endotoxin standard biological reference preparation (BRP) from the *Ph. Eur*. (EDQM, REF: E0150000); sterile, non-pyrogenic distilled or deionized water to reconstitute endotoxin standard (Charles River REF: W120); PGN from *Staphylococcus aureus* (PGN-Sandi Ultrapure) from Invivogen (REF: tlrl-sipgn); flagellin from *Salmonella typhimurium* (FLA-STA ultrapure) from Invivogen (REF: tlrl-epstfla-5); synthetic triacylated lipoprotein (PAM_3CSK_4) from Invivogen (REF: tlrlp-pms), heat-killed *Staphylococcus aureus* from Invivogen (HKSA); PeliKine compact human IL-6 kit (Sanquin, REF: M1916); PeliKine tool set 1 (additional reagents for application in PeliKine compact ELISA kits) (Sanquin, REF: M1980); distilled water for ELISA buffers (WFI from Gibco, REF: A1287301 (500 mL)).

Cell culture and IL-6 ELISA

The MAT procedure comprised two main steps: (i) culture of cells with endotoxin standards and samples, and (ii) quantification of the cytokine concentration in the cell-conditioned supernatant. The cells were cryopreserved-PBMC (pMAT cells, each lot consisting of a PBMC pool from 4 different donors, Sanquin, The Netherlands), and the cytokine readout was IL-6 concentration (Sanquin PeliKine IL-6 ELISA). PBMCs were cultured with 4 replicates of each dilution of standard and test solution (added to the plate in columns), and plate layouts and cell additions (by rows) were as described previously (Gaines Das et al., 2004).

One vial of pMAT cells (sufficient for one 96-well plate) was taken from a -80°C freezer and immediately thawed in a water bath at 37ºC until only a small piece of ice/clump of cells remained (< 5 min). The entire content of the vial (1 mL in total) was transferred to a 50 mL tube and to this 5 mL of complete medium (at room temperature) was immediately added slowly while gently swirling the tube (1 min). Cells were not vortexed or pipetted vigorously, and care was taken to avoid forming air bubbles. The entire cell suspension was transferred to a multichannel reservoir and, beginning immediately after the transfer, 50 µL aliquots of the pMAT cells were added to each well of the plate by row using a 12-channel pipette. The order of addition of cells to stimuli or test samples was row A, E, B, F, C, G, D, H, with the cells mixed once in the reservoir before adding to each row. PBMC incubations were at 37 ± 1°C, 5% CO_2 in a humidified incubator for 18-24 h, after which time cell supernatants were assayed for IL-6.

Selection of PBMC lots

Six lots of PBMCs were screened in terms of their responses to standard endotoxin and to NEPs: the standard curve for endotoxin was required to meet both *Ph. Eur*. acceptance criteria and to be S-shaped. Three of these lots were used for the initial qualification, and 3 additional lots of PBMC were qualified later (NEP reactivity not shown). Where different lots of PBMC gave similar (good) responses to standard endotoxin and NEPs, the number of (frozen) vials of each lot remaining for future work was the deciding factor.

Determination of parameters to discriminate S-shaped curves for standard endotoxin: tolerance intervals for the curve for standard endotoxin

Standard curves were obtained for the 3 lots of PBMC to permit the calculation of the 95% confidence interval (CI) and 95% coverage tolerance intervals (95%/95% TIs) for the standard curve parameters EC_{50}, range of CI of EC_{50}, and range of CI of upper asymptote. The 95%/95% TIs give the ranges of the monitored standard curve parameters within which (at least) 95% of the future standard curve parameters will fall with 95% probability. The 95%/95% TIs were computed based on valid curves (i.e., S-shaped standard curves fulfilling both acceptance criteria for the standard curve prescribed in *Ph. Eur*. 2.6.30 (EP, 2017)), and so may be used to evaluate the validity of standard curves that will be obtained in future experiments. For example, if the EC_{50} of a future standard curve is outside the 95%/95% TI of the EC_{50}, it is implausible that this standard curve is valid (as it falls outside the range of the 95%/95% TI determined based on valid S-shaped standard curves). In the current study, the EC_{50}, range of the CI of the EC_{50}, and range of the CI of the upper asymptote were monitored because these parameters were shown to accurately discriminate S-shaped standard curves from partial S-shaped standard curves based on the 95%/95% TI. The use of additional parameters of the standard curve (such as the slope parameter or the range of the CI of the slope parameter) could be considered to further refine this approach.

Determination of LOQ and range for endotoxin: relative β-expectation tolerance limits

Six solutions of standard endotoxin were prepared for the determination of the LOQ, range for endotoxin, and relative β-expectation tolerance limits. Each of the solutions contained a known but different concentration of standard endotoxin (LPS1, LPS2, LPS3, LPS4, LPS5 and LPS6 corresponding to 0.02, 0.04, 0.08, 0.16, 0.24 and 0.32 EU/mL, respectively). For each 96-well plate, each concentration was prepared independently three times, and each solution was tested in four replicates. A total of 12 MATs were performed with the above endotoxin solutions by 2 operators, with 3 lots of PBMCs on different days. Total error profile (USP, 2018) was calculated for 6 concentrations of standard endotoxin: 0.02, 0.04, 0.08, 0.16, 0.24 and 0.32 EU/mL. Total error profile was not calculated for very low (0.01 EU/mL) and high (0.64 EU/mL) concentrations of LPS solutions/spiked samples since these gave values that were close to the lower and upper asymptote, respectively, i.e., low precision was obtained in the non-linear portions of the standard curve.

Determination of the method accuracy (relative bias), precision (includes repeatability and intermediate precision), linearity, and specificity of endotoxin

Six solutions of standard endotoxin were prepared essentially as described above. Relative bias was calculated with the following formula

$$Relative\ bias\ (\%) = \left[\frac{GM\ (Measured\ potency)}{Target\ potency} - 1 \right] \times 100\%,$$

where GM is the geometric mean. The 95% β-expectation TI (USP, 2018) was computed for each of the six concentrations of endotoxin: -50 to 100% relative error was equivalent to a spike recovery of 50-200%. The repeatability (variability within experiment) and intermediate precision (variability between different days and technicians) were based on the geometric coefficient of variation (GCV). Values for the GCV are obtained for all variance components (sources of variability) using a formula that assumes log-normally distributed responses:

$$GCV\,(\%) = 100 * \sqrt{\exp(\sigma^2) - 1}$$

In this equation, σ^2 represents the variance component considered. In this study, several variance components are considered:

- One variance component describes the batch-to-batch variability. This component is process-related and is provided as information only.
- One variance component describes the day-to-day variability, which is expected to be small/zero.
- The residual variability represents the repeatability.

The different variance components were estimated using a linear mixed-effects model that included log(concentration) as fixed effect and random effects for batch and day (in addition to the residual). The intermediate precision is the sum of the day-to-day variability and the repeatability.

The limits of the CIs for the GCV values are obtained by applying the transformation previously defined in the equation of GCV to the limits of the CI for the variance component provided by the JMP software package (JMP version 12.2. SAS Institute Inc., Cary, NC, 1989-2021).

Determination of accuracy, precision, linearity, LOQ, range and specificity for NEPs

Four different NEPs were tested, namely FLA, PAM, heat-killed HKSA, and PGN. Four solutions of each NEP were prepared. Each of the solutions contained a known but different concentration of NEP. For each plate, each concentration was prepared three times and each solution was tested in four replicates. Each NEP was tested 3 times (in 3 independent MATs). A total of 12 MATs with NEPs were performed by 2 operators, with 3 lots of PBMCs on different days.

The 95% β-expectation TIs (USP, 2018) were computed for each of the 4 concentrations of each NEP: -50 to 100% relative error was equivalent to a spike recovery of 50-200%.

The repeatability (variability within experiment) and intermediate precision (variability between different days and technicians) were based on GCV, which was calculated in the same way as described in the previous section.

Determination of robustness

In some experiments, different cell lots were used and 4 parameters (temperature of incubator, CO_2 concentration, cell number/well, and incubation period) were increased or decreased to challenge the robustness of the MAT. The cytokine responses to standard endotoxin and positive controls (NEPs) were compared. The positive controls (NEPs) were flagellin from *Salmonella typhimurium* (FLA-ST), synthetic triacylated lipoprotein (Pam_3CKS_4), heat-killed *Staphylococcus aureus* (HKSA), and PGN from *Staphylococcus aureus* (PGN). The negative control was cell culture medium.

Product-specific validation

Product-specific validations were carried out with three PPQ batches of each of three different mAbs. This comprised the test for interfering factors (50-200% recovery of standard endotoxin spiked into the mAb), validation with NEPs (including recovery of a NEP spiked into the mAb), and lack of interference by the mAbs in the ELISA at the dilution, f, the minimum dilution that gave 50-200% recovery of spiked standard endotoxin (*Ph. Eur.*). Once these requirements had been met, the three batches of each of the three mAbs were tested for compliance with their CLCs.

Data analysis

The IL-6 content of supernatants from wells containing dilutions of standard endotoxin, negative and positive controls, and the mAbs was quantified in EE/mL using values read off a 4PL fitted to the (7-point) endotoxin standard curve following the application of Dixon's test for outliers (USP, 2018). This allowed the determination of whether the standard curve satisfied the acceptance criteria and the evaluation of the additional analytical characteristics given above. The EE/mL values for dilutions of the mAbs were used to determine whether the mAb batches complied with their CLCs.

The statistical analyses described below were conducted in Softmax Pro v7 or v6.4.1 with the exception of the computation of the regression and non-linearity *p*-values, which were carried out in CombiStats v5.0 (EDQM). To compute the 2 *p*-values, the data of the endotoxin standard curve were exported from SoftMax Pro into CombiStats. After computation of the two *p*-values, these were entered into SoftMax Pro software. Data transfer was verified by a second person.

When the 4PL model was fitted, the lower asymptote parameter was fixed to the mean of the four replicates for the blank. In any instances where a "hook effect" was observed, which was defined as a mean OD value for the highest concentration of standard endotoxin that was less than the mean OD value for the second-highest concentration of standard endotoxin, data for the highest concentration of standard endotoxin was excluded from the standard curve. Consequently, in such cases, 6 rather than 7 concentrations of standard endotoxin were used to construct the standard curve. It should be noted that the exclu-

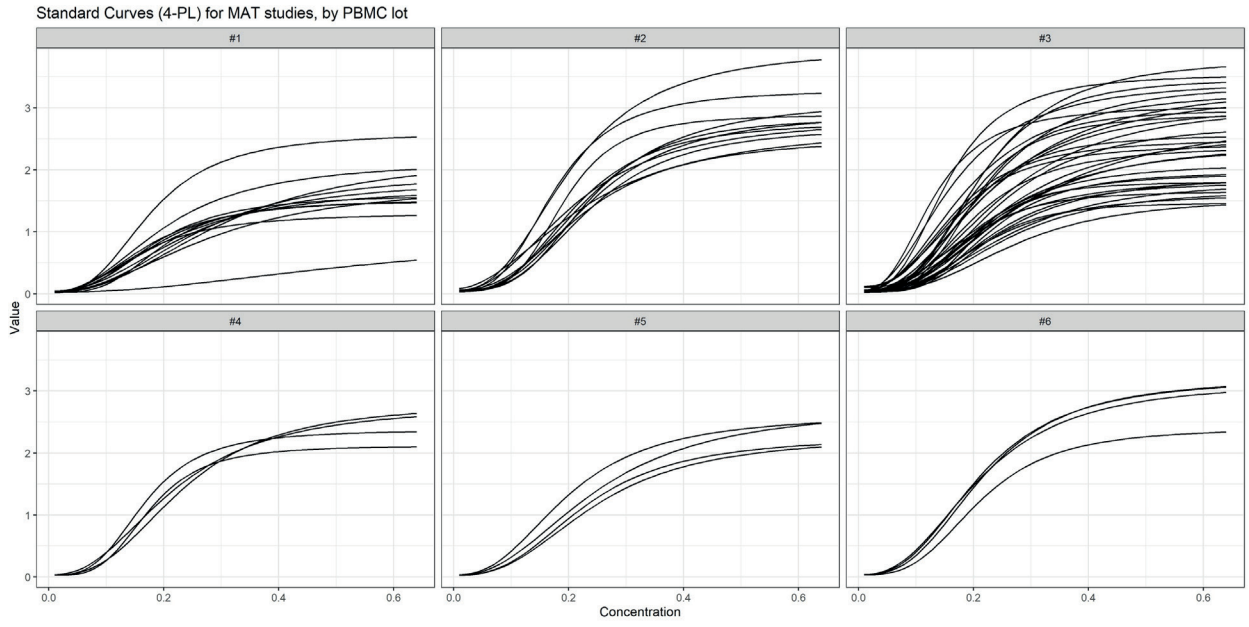

Fig. 1: Fitted standard curves from 60 MAT experiments
PBMC batch number is in the header of the figure. Absorbance units (OD) are shown on the Y-axis.

sion of the highest standard endotoxin from the standard curve still gave a 6-point standard curve that exceeded the minimum requirement of *Ph. Eur*., which is a 4-point standard curve for standard endotoxin.

When the standard curve fulfilled the two acceptance criteria specified above, it was used to appropriately transform OD values for samples, i.e., dilutions of products, into EE by back-transformation. Note that OD values that were above the upper asymptote of the standard curve cannot be converted into EE. Also, when the back-transformed EE concentration of a sample was below the LOD, it was assigned the LOD (instead of the calculated EE value). The latter approach is conservative, i.e., it uses a worst-case scenario by assigning the LOD to values for samples where the calculated EE value is below the LOD.

Spike recovery was calculated by comparing the EE detected in a solution spiked with endotoxin standard (1 EE = 1 EU for endotoxin standard) after the subtraction of any endotoxin equivalents detected in the (unspiked) solution prior to addition of the (added) endotoxin standard. Where the concentration of EE in the unspiked solution was below the LOD, the LOD (in EE) was subtracted from the values calculated for the spiked and unspiked solutions. Spike recovery is given as percentage recovery of the added standard endotoxin, i.e., EE detected in the well/endotoxin standard added to the well x 100.

Spike recovery was calculated using the following formula:

$$\text{Spike recovery [\%]} = (\text{spiked EE/mL} - \text{unspiked EE/mL}) / (\text{added spike EE/mL} - \text{LOD EE/mL}) * 100$$

3 Results

3.1 Method qualification

Aliquots from 6 different lots of PBMCs were used to generate curves that were fitted to the seven-point dose-response data for standard endotoxin shown in Figure 1. More data is presented for batches that were selected for subsequent use than for batches that were not selected. One standard curve, using PBMC from lot #1, was noticeably shallower than the other curves but still complied with the two acceptance criteria for the standard curve (EP, 2017), these being:

(i) the regression of responses (appropriately transformed if necessary) on $\log_{10}$ dose shall be statistically significant ($p < 0.01$);
(ii) the regression of responses on $\log_{10}$ dose must not deviate significantly from linearity ($p > 0.05$).

These acceptance criteria were met for all of the standard curves for which data is presented below. Also, the *Ph. Eur*. requirement for the basal release of the readout (blank) in the absence of added standard endotoxin to be below an OD of 0.100 was satisfied for all plates. The mean LOD (back-calculated EE/mL) and mean cut-off (OD) values, calculated from the data for all 39 plates from the method validation, were 0.03 EE/mL per well and 0.076 absorbance units, respectively.

Theoretical S-shaped standard curve

Parameters were determined that allowed dose-response data for standard endotoxin to be identified that did not fit the theoretical (S-shaped) curve for the 4PL. These parameters were addi-

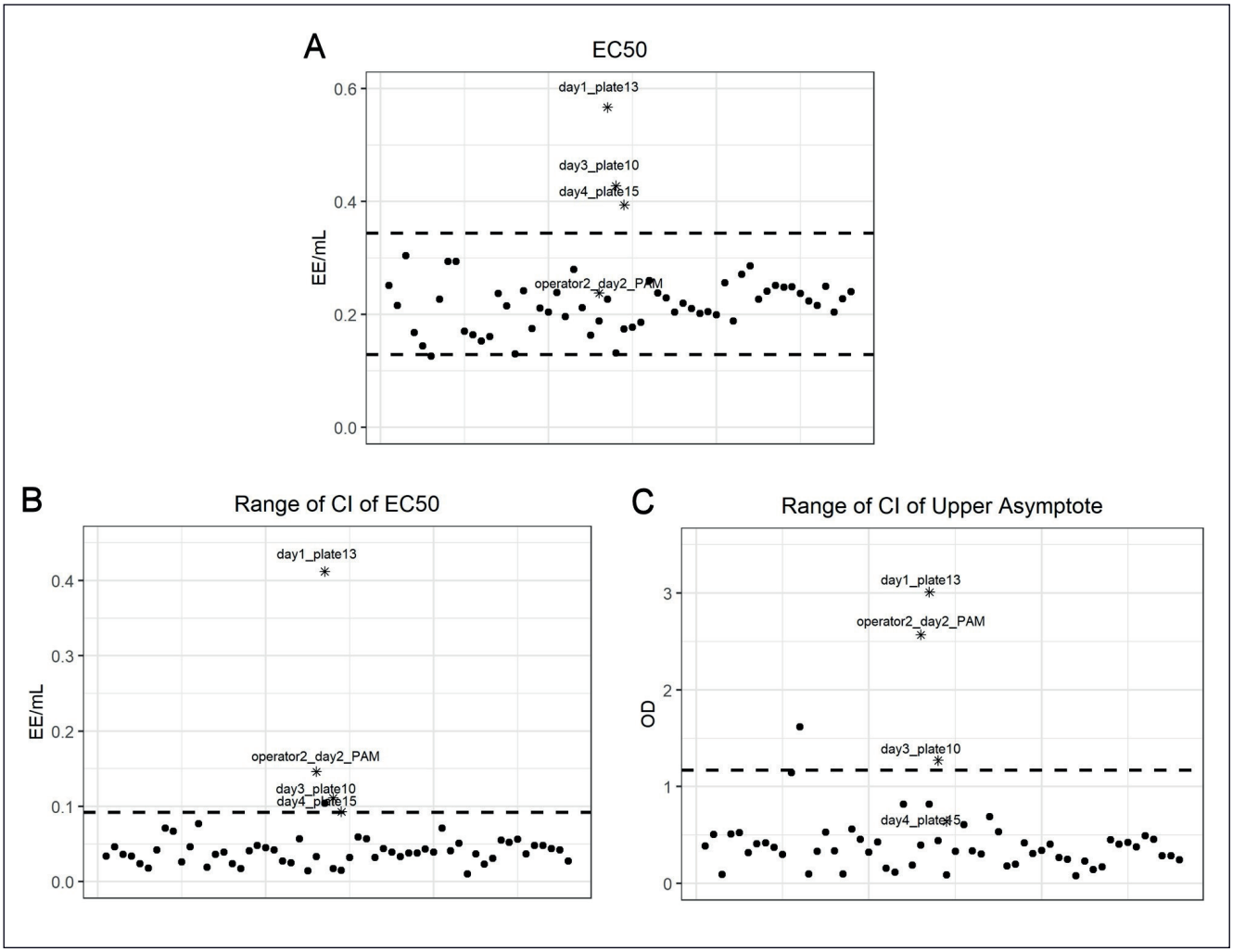

Fig. 2: 95%95% TI for the standard curve parameters EC_{50}, range of CI of EC_{50}, and range of CI of upper asymptote (dashed lines)

The data presented is the raw data for the three parameters. The filled circles correspond to curves that fit the theoretical (S-shaped) curve for 4PL, while the stars correspond to curves that do not fit the theoretical S-shaped curve.

tional to the two acceptance criteria for the standard curve prescribed in *Ph. Eur*. 2.6.30 and comprised: the (estimated) EC_{50}, the range of the CI for the EC_{50}, and the range of the CI for the upper asymptote (the range of an estimated parameter reflects the uncertainty in the estimated value). Data from the 39 standard curves obtained in this study together with an additional 21 standard curves from other (contemporaneous) MATs that complied with the 2 acceptance criteria of the *Ph. Eur*. were combined and analyzed (60 standard curves in total).

Summary statistics are shown in Table 1 for the following parameters of the standard curves: R^2, lower asymptote, upper asymptote, slope, EC_{50}, and the ranges of the CIs for the upper asymptote and EC_{50}. The 95%/95% TI for EC_{50} (2-sided interval), range of CI of EC_{50} (1-sided interval), and range of CI of upper asymptote (1-sided interval) are shown in Figure 2. The 95%/95% TIs were sensitive to dose-response data that did not fit the theoretical (S-shaped) curve. As can be seen for all the curves that do not fit the theoretical (S-shaped) curve, at least two of the 95%/95% TI for EC_{50}, range of CI of EC_{50}, and range of CI of upper asymptote will identify them as aberrant. The 95%/95% TI for the three parameters could be used to screen for aberrant standard curves in future MATs. With these, a rule that may be used to identify curves that do not fit the theoretical (S-shaped) curve for 4PL in the MAT assay is: "At least two of the estimates of EC_{50}, range of CI of EC_{50}, and range of CI of upper asymptote being outside the corresponding interval of each parameter". This would be a more objective way than an analyst judging by eye whether a standard curve fits the theoretical S-shape. It can also be seen from Table 1 that for R^2, the median was computed instead of the mean since the distribution of the estimated R^2 was heavily skewed. The median R^2 was close to 1, which indicates that the fitted standard curves adequately fitted the data.

Tab. 1: Summary statistics of all endotoxin standard curve parameters in the combined qualification study (39 standard curves) and an additional 21 standard curves from other (contemporaneous) MATs

Descriptive statistics	n	Lower asymptote (OD)	Slope	EC_{50} (EE/mL)	Upper asymptote (OD)	Range of 95% CI of EC_{50} (EE/mL)	Range of 95% CI of upper asymptote (OD)	R^2
Min.	60	0.017	1.769	0.126	1.096	0.010	0.078	0.994
Mean	60	0.044	2.716	0.228	2.825	0.050	0.491	0.999[a]
Max.	60	0.105	4.317	0.567	5.204	0.412	3.011	1.000
Std. Dev.	60	0.020	0.560	0.070	1.036	0.053	0.513	0.001

[a] Median

β-expectation tolerance limits, LOQ and range for endotoxin

The relative β-expectation tolerance limits of spike recovery (USP, 2018), with the 95% level of confidence for concentrations of standard endotoxin between 0.02 EU/mL and 0.32 EU/mL, are shown in Figure 3. For 0.08 EU/mL to 0.32 EU/mL, the upper β-expectation tolerance limits for the 95% level of confidence were within the -50% to 100% interval, corresponding to the 50%-200% spike recovery described in *Ph. Eur*. 2.6.30. In contrast, the lower β-expectation tolerance limits for the 95% level of confidence were a little below the lower limit over the whole concentration range, except for 0.24 EU/mL. The same plots were generated with lower levels of confidence of 80% and 75%, respectively (plots not shown). For both lower levels of confidence, the β-expectation tolerance limits were within the lower limit over the linear concentration range (-50%), except for 0.32 EU/mL, which was a little below the lower limit (-55.68% and -54.26% for the confidence levels of 80% and 75%, respectively). The overall probabilities of having subsequent measurements within the -50% to 100% interval (corresponding to the 50%-200% spike recovery specified in the *Ph. Eur*.) were calculated for concentrations of standard endotoxin of 0.08, 0.16 and 0.24 EU/mL as 94.52%, 89.61% and 99.99%, respectively. For 0.32 EU/mL the probability of having measurements within the -50% and 100% interval still reached a probability of 77.76%, which was a good performance bearing in mind that the upper asymptote of the standard curves had almost been reached at this concentration, with responses essentially flat between 0.32 and 0.64 EU/mL. In summary, the portion of the standard curve between 0.08 and 0.32 EE/mL provided the most reliable estimates of EE/mL in a sample (i.e., range) and permitted the LOQ of the MAT to be calculated as 0.08 EE/mL. This value was confirmed by the data from the accuracy (total error) and robustness study. (see Sections 3.1.3-3.1.5).

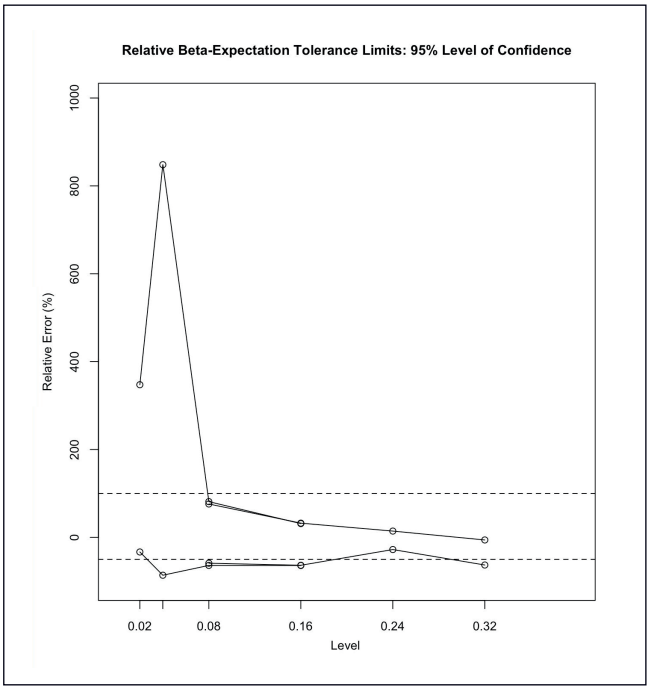

Fig. 3: Total error profile for 6 concentrations of standard endotoxin:0.02, 0.04, 0.08, 0.16, 0,24 and 0.32 EU/mL

The lines are the 95% upper and lower β-expectation TI derived from the LPS method qualification. The horizontal dashed lines represent the (-50% to 100%) interval (corresponding to 50%-200% spike recovery as described in *Ph. Eur.*).

Accuracy, precision (includes repeatability and intermediate precision), linearity, and specificity of endotoxin

The method accuracy, precision (includes repeatability and intermediate precision), linearity, and specificity of endotoxin in MAT were assessed at 6 concentrations of endotoxin (LPS1-6, i.e., 0.02, 0.04, 0.08, 0.16, 0.24 and 0.32 EU/mL) tested in MATs on multiple days by two operators with three PBMC lots (#1, #2 and #3). Endotoxin spike recovery of 0.16 EU/mL (EC_{50}) was chosen as the measure of the quality of the MAT assay. The target recovery was set at 100% recovery, meaning that the measured concentration of endotoxin spike would be exactly the expected concentration of the spike. Accuracy and relative bias of the assay were assessed from the data generated for the accuracy study as part of the method validation. The total error and the relative bias at each endotoxin concentration were calculated from values generated for LPS-1 to LPS-6 (with each of the 6 concentrations tested in at least 6 MATs).

Tab. 2: Relative bias, repeatability, and precision of measured endotoxin recovery in MAT at different endotoxin concentrations with 3 lots of PBMCs

Spike conc. (EU/mL)	Plate	Measured recovery (%)			GM (n = 18 or 36)	Relative bias (%)	Lower and upper 90% confidence limit of relative bias	Repeat-ability (GCV%)	Intermediate precision (IP) (GCV%)	Upper 90% confidence limit of IP (GCV%)
		Repl 1	Repl 2	Repl 3						
0.02	1	237.618	87.258	193.712	172.8	72.81	27.22; 134.7	47.40	51.89	113.0
	4	243.933	168.967	256.542						
	7	234.871	145.216	270.741						
	10	209.087	116.551	260.449						
	13	[a]	[a]	[a]						
	16	114.552	0.000	85.107						
0.04	1	85.684	89.280	122.834	114.57	14.57	-38.89; 114.8	16.11	116.8	395.8
	4	132.073	112.975	111.564						
	7	112.832	111.092	110.800						
	10	118.191	103.100	128.594						
	13	438.495	412.653	373.604						
	16	49.742	32.362	31.226						
0.08	1	57.286	74.078	75.788	83.03	-16.97	-25.12; -7.94	34.56	24.80	49.02
	4	86.380	80.761	85.742						
	7	106.793	104.821	102.722						
	10	99.105	86.429	87.826						
	13	101.628	113.027	99.317						
	16	48.895	52.138	46.049						
	19	74.948	46.834	86.802						
	20-2	96.799	72.616	102.795						
	21-2	67.705	55.501	86.792						
	22	99.493	97.885	114.864						
	23	75.839	70.611	77.132						
	24	104.743	76.774	207.973						
0.16	1	53.98	63.39	54.31	69.12	-30.88	-39.73; 20.73	20.46	26.86	47.61
	4	73.72	78.22	77.21						
	7	77.89	84.97	89.08						
	10	58.8	74.22	72.02						
	13	83.18	98.34	84.45						
	16	46.05	50.54	47.7						
	19	51.87	43.37	54.6						
	20-2	66.39	81.74	93.93						
	21-2	62.29	57.14	68.18						
	22	94.11	99.62	99.46						
	23	59.21	58.07	71.55						
	24	73.33	76.03	73.59						
0.24	1	97.35	92.19	93.8	90.93	-9.07	-14.97; -2.76	6.27	10.02	19.49
	4	74.17	79.29	87.57						
	7	96.652	101.361	96.286						

[a] means no value calculated because the spike – LOD was 0, resulting in a denominator of 0 for calculation of recovery.

Spike conc. (EU/mL)	Plate	Measured recovery (%)			GM (n = 18 or 36)	Relative bias (%)	Lower and upper 90% confidence limit of relative bias	Repeat-ability (GCV%)	Intermediate precision (IP) (GCV%)	Upper 90% confidence limit of IP (GCV%)
		Repl 1	Repl 2	Repl 3						
	10	90.237	92.822	108.16						
	13	82.46	88.72	82.04						
	16	94.48	97.08	88.59						
0.32	1	51.21	48.28	36.81	59.06	-40.94	-48.33; -32.49	15	21.96	43.38
	4	61.38	70.47	75.43						
	7	62.61	78.05	62.77						
	10	53.71	63.05	78.91						
	13	46.69	58.83	56.2						
	16	53.67	63.01	60.15						

As summarized in Table 2, the relative bias for endotoxin concentrations 0.08 EU/mL to 0.32 EU/mL was between -40.94% and -9.07% when calculated from the measured recoveries. The geometric means of the endotoxin recovery for the endotoxin concentrations 0.08 EU/mL to 0.32 EU/mL were between 59.06% and 90.93%. Further, the relative bias of endotoxin recovery from the robustness study for 0.16 EU/mL was 1.3%, which was in agreement with the data from the accuracy study (For details of the robustness study see Section 3.1.5).

The precision of the method was evaluated at two levels – repeatability and intermediate precision – from the data generated for the accuracy study. Repeatability ranged from 6.27% to 34.56% for the endotoxin concentrations above the LOQ of the MAT when calculated from recoveries (Tab. 2). The intermediate precision ranged from 10.02% to 26.86% for the concentrations 0.08, 0.16, 0.24 and 0.32 EU/mL based on the endotoxin recoveries (Tab. 2).

For linearity, the relationship between the assumed concentrations of endotoxin and the measured concentration was determined by linear regression. Linearity of the assay was generated from the accuracy data set. The linear regression slopes, y-intercepts, as well as the determination coefficient were calculated. For the standard curve, the linearity characteristics, i.e., determination coefficient, slope, and y-intercept were 0.99, 1.0273, and 0.0002, respectively. For the samples, the linearity characteristics, i.e., determination coefficient, slope, and y-intercept were 0.80, 0.54, and 0.036, respectively.

The method specificity (i.e., lack of matrix interference) was demonstrated by evaluating "samples" (= medium only) at multiple dilutions (with medium) with and without an endotoxin spike (spike concentration was EC_{50} = 0.16 EU/mL) to verify lack of matrix interference in the data generated for the robustness study. In addition, the negative samples (additional blanks different from the negative control of the standard curve) were also a measure of specificity: These were all negative. The EE values of the blanks for both endotoxin and NEPs were also evaluated from the data generated for the accuracy study. The overall endotoxin spike recovery from the robustness study was 102.67% (target 100%). The geometric mean was 101.3% (target 100%), showing that there was no matrix interference in the MAT assay when only medium, cells, and endotoxin were present in wells. When the additional blanks from the various accuracy studies were combined, their average value was 0.025 EU/mL,

Tab. 3: Concentrations of NEPs at each PBMC lot that have a prediction interval above LOQ (0.08 EU/mL)
A check symbol means that the 95% prediction interval was higher than LOQ, meaning that in > 95% of the future experiments the NEP will be detected above LOQ. A cross means that the 95% prediction interval was lower than LOQ (0.08) but the non-endotoxin pyrogen was still detected above LOD.

NEP	Concentration	#1	#2	#3
FLA	3.13 ng/mL	X	X	X
	6.25 ng/mL	X	X	X
	12.5 ng/mL	✓	✓	✓
	25 ng/mL	✓	✓	✓
PAM	0.625 ng/mL	X	✓	✓
	1.25 ng/mL	X	✓	✓
	2.5 ng/mL	✓	✓	✓
	5 ng/mL	✓	✓	✓
PGN	1.25 µg/mL	X	✓	✓
	2.5 µg/mL	X	✓	✓
	5 µg/mL	X[a]	✓	✓
	7.5 µg/mL	X	✓	✓
HKSA	0.06 x 10^6 cells/mL	✓	X	X
	0.125 x 10^6 cells/mL	✓	X	✓
	0.25 x 10^6 cells/mL	✓	X	✓
	0.5 x 10^6 cells/mL	✓	✓	✓

[a] Was detected above LOQ (0.08 EU/mL) during the robustness study.

Tab. 4: Relative bias for each NEP at each concentration and the 90% CI

NEP	Concentration	Relative bias (%)	Lower and upper 90% confidence limit of relative bias
FLA	3.13 ng/mL	96.38	52.19; 153.4
	6.25 ng/mL	59.21	44.42; 75.52
	12.5 ng/mL	20.05	0.3189; 43.65
	25 ng/mL	-0.8611	-19.48; 22.07
PAM	0.625 ng/mL	74.00	-48.64; 489.5
	1.25 ng/mL	15.74	66.34; 298
	2.5 ng/mL	-12.13	-77.94; 250.1
	5 ng/mL	16.65	-60.26; 74.80
PGN	1.25 µg/mL	230.4	19.54; 813.2
	2.5 µg/mL	113.8	-23.21; 495.5
	5 µg/mL	18.81	-55.65; 218.3
	7.5 µg/mL	-13.21	73.51; 184.3
HKSA	0.06 x 10^6 cells/mL	76.43	-58.31; 646.7
	0.125 x 10^6 cells/mL	43.16	-63.25; 457.7
	0.25 x 10^6 cells/mL	20.41	-49.82; 188.9
	0.5 x 10^6 cells/mL	5.30	-65.77; 223.9

Tab. 5: Characteristics of linear regression of FLA, PAM, PGN and HKSA

NEP	Parameter	Value	Lower and upper 90% CI
FLA	Slope	0.6634	0.6187; 0.7081
	y-intercept	-0.2628	-0.3187; -0.2069
	R^2	0.9488	NA
PAM	Slope	0.63	0.3814; 0.8786
	y-intercept	-0.1598	-0.3221; 0.002462
	R^2	0.3654	NA
PGN	Slope	0.2444	0.02218; 0.4666
	y-intercept	-0.5811	-0.8245; -0.3377
	R^2	0.09233	NA
HKSA	Slope	0.6802	0.4465; 0.9138
	y-intercept	-0.2430	-0.5169; 0.03081
	R^2	0.4318	NA

which was below the LOD of the assay (0.03 EU/mL) and well below the LOQ (0.08 EU/mL). Thus, negative samples resulted in values below the LOD.

Accuracy, precision, linearity, LOQ, range, and specificity of NEPs (TLR ligands, NOD2 ligand)

The method accuracy, precision, linearity, LOQ, range, and specificity of NEPs in the MAT assay were assessed with four NEPs, namely FLA, PAM, HKSA, and PGN. Each NEP was tested at 4 concentrations in MATs on multiple days by two operators with three PBMC lots (#1, #2 and #3). In the absence of any standards for NEPs, responses to them were back-transformed to EE/mL by comparison with responses to the endotoxin standard curve. The measured concentrations (EE/mL) for each dose of each NEP with each of the 3 PBMC lots are shown in Tables S1 (FLA), S2 (PAM), S3 (PGN) and S4 (HKSA) in the supplementary file[1]. The data is also summarized in Table 3, which shows the concentrations of NEPs that had a 95% prediction interval above the LOQ (0.08 EE/mL) for each of the 3 different PBMC lots, based upon pooled data for the 3 lots since that provided the most precise estimates despite the differences between lots. A check (tick) symbol indicates that the 95% prediction interval was above the LOQ. Overall, PBMC lot #3 was the most sensitive across the range of NEPs. The reactivity to FLA was similar for the 3 PBMC lots, whereas PBMC lot #1 was less sensitive to PAM and, especially, PGN. PBMC lot #1 detected PGN (1.25 µg/mL) above the LOD (0.03 EE/mL) but lower than LOQ (0.08 EE/mL). PAM concentrations of 2.5 ng/mL and 5 ng/mL showed larger inter-assay variability than other concentrations of the NEPs. The variability among the PBMC lots for responses to PGN was similar. PBMC lot #2 was less sensitive to HKSA. The relative bias for PAM, PGN and HKSA had broad CIs, whereas the relative bias for FLA had smaller CIs (Table 4). As a result, each new PBMC lot selected for use in the MAT will be qualified with the 4 NEPs at 4 different concentrations. Further, at least one NEP will be used as a positive control on each plate in each MAT.

The precision of the assay was assessed from the data generated for the accuracy study. The repeatability (GCV%) values were FLA: 10.38%, PAM: 42.90%, PGN: 13.70%, and HKSA: 17.75%. The high repeatability value for PAM was due to two atypically high values from 2 plates. The intermediate precision (GCV%) values were FLA: 14.20%, PAM: 106.10%, PGN: 88.77%, and HKSA: 104.90%. The intermediate precision values were high compared with the repeatability except for FLA. The values for the upper confidence limit at 90% of intermediate precision (GCV%) values were FLA: 52.53%, PAM: 1607%, PGN: 1472%, and HKSA: 2111%.

[1] doi:10.14573/altex.2111301s

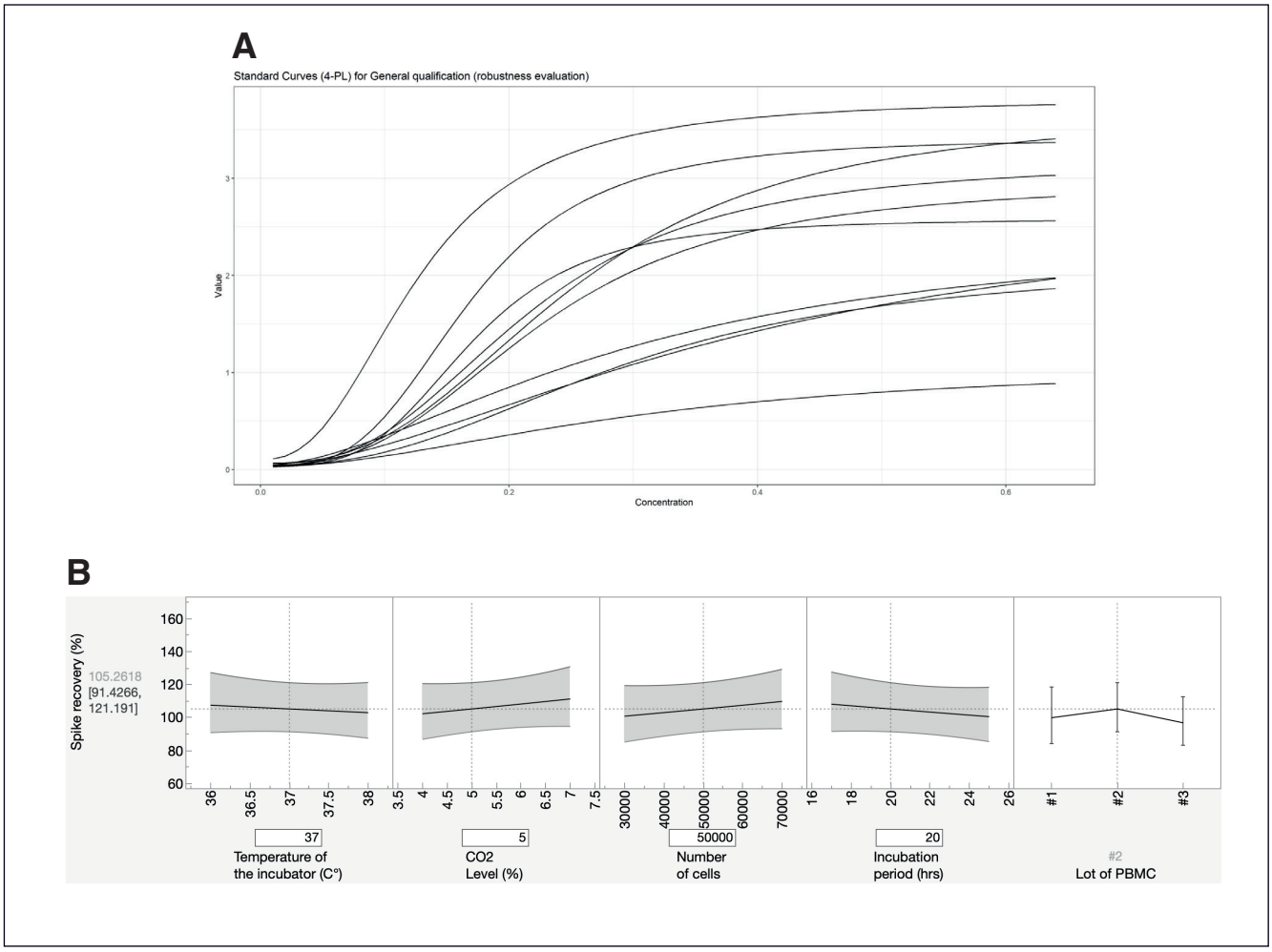

Fig. 4: General robustness of the MAT

(A) Fitted standard curves of the robustness study; (B) Prediction profile of spike recovery plotted against each of the five parameters investigated in the robustness design of experiments (DoE). The lines show how the predicted values change when changing one of the variables. The 95% CI for the predicted values is shown by the grey-shaded area surrounding the prediction trace. Note that the lot of PBMC was not a continuous variable, so discrete confidence intervals (rather than confidence bands) are shown.

Linearity was determined from the accuracy data set. The linear regression slopes, y-intercepts, and their corresponding 90% CIs were calculated as well as the determination coefficient as shown in Table 5.

LOQ and range was determined from the data generated for the accuracy and robustness study. For FLA, the LOQ was 6.25 ng/mL and the range was 6.25-25 ng/mL, with 6.25 ng/mL FLA being equivalent to 0.0476 EU/mL, i.e., below the LOQ of 0.08 EU/mL for standard endotoxin. The LOQ and range could not be defined for the other NEPs due to the fact that lower and upper β-expectation tolerance limits of the relative error were outside the acceptance limits.

The method specificity (i.e., lack of matrix interference) was demonstrated by evaluating "samples" (= medium only) at multiple dilutions (with medium) with four different spikes: FLA 12.5 ng/mL, PAM 2.5 ng/mL, PGN 5 µg/mL, HKSA 0.25 x 10^6 cells/mL (these spike concentrations being approximately the respective EC_{50} concentrations) to verify lack of matrix interference in the data generated for the robustness study. The FLA, PAM, PGN, and HKSA spikes gave EE values greater than the overall LOD (0.03 EE/mL) in the accuracy and robustness studies, except for HKSA 0.06 x 10^6 cells/mL and 0.125 x 10^6 cells/mL for PBMC lot #2 in the accuracy study. In the robustness study, for all NEPs and PBMC lot combinations, the back-calculated EE/mL value was above the LOQ of 0.08 EE/mL, except for FLA on one plate, although this value was above the LOD (0.03 EU/mL).

General robustness of the MAT

Additional data from experiments that were carried out to challenge the general robustness of the MAT are summarized in the prediction profile shown in Figure 4A and B.

Tab. 6: *p*-Value regression, *p*-value non-linearity, basal IL-6 release, and LPS, PGN, and FLA spike recovery for DP1 obtained with PBMC batch #3

		LPS				PGN				FLA			
No. batch DP1	Dilution in medium	*p*-value regress.	*p*-value non-lin.	Basal IL-6	LPS spike rec. (%)	*p*-value regress.	*p*-value non-lin.	Basal IL-6	PGN spike rec. (%)	*p*-value regress.	*p*-value non-lin.	Basal IL-6	FLA spike rec. (%)
	Accept. criteria	< 0.01	> 0.05	< 0.100	50-200	< 0.01	> 0.05	< 0.100	50-200	< 0.01	> 0.05	< 0.100	50-200
#1	1:4[a]	0.00	0.90	0.083	88	0.00	0.99	0.050	164	0.00	0.34	0.055	238
#1	1:8[b]	0.00	0.90	0.083	95	0.00	0.99	0.050	284	0.00	0.34	0.055	208
#1	1:16	0.00	0.90	0.083	72	0.00	0.99	0.050	136	0.00	0.34	0.055	131
#1	1:32	0.00	0.90	0.083	85	NA	NA	NA	NA	NA	NA	NA	NA
#2	1:4	0.00	0.27	0.080	114	0.00	0.99	0.047	174	0.00	0.23	0.029	174
#2	1:8[b]	0.00	0.27	0.080	117	0.00	0.99	0.047	180	0.00	0.23	0.029	158
#2	1:16	0.00	0.27	0.080	112	0.00	0.99	0.047	147	0.00	0.23	0.029	123
#2	1:32	0.00	0.27	0.080	120	NA	NA	NA	NA	NA	NA	NA	NA
#3	1:4	0.00	1.00	0.084	78	0.00	1.00	0.058	157	0.00	0.85	0.050	204
#3	1:8[b]	0.00	1.00	0.084	94	0.00	1.00	0.058	144	0.00	0.85	0.050	158
#3	1:16	0.00	1.00	0.084	91	0.00	1.00	0.058	128	0.00	0.85	0.050	115
#3	1:32	0.00	1.00	0.084	91	NA	NA	NA	NA	NA	NA	NA	NA

[a] Medium: Iscove's Modified Dulbecco's Medium supplemented with human AB serum. [b] Final selected dilution of the DP (f). PGN and FLA are spiked only for dilutions f, fx2, fx4, so no data are presented for other dilutions, which are labeled NA for not applicable.

Tab. 7: *p*-Value regression, *p*-value non-linearity, basal IL-6 release, and LPS, PGN, and FLA spike recovery for DP2 obtained with PBMC batch #3

		LPS				PGN				FLA			
No. batch DP2	Dilution in medium	*p*-value regress.	*p*-value non-lin.	Basal IL-6	LPS spike rec. (%)	*p*-value regress.	*p*-value non-lin.	Basal IL-6	PGN spike rec. (%)	*p*-value regress.	*p*-value non-lin.	Basal IL-6	FLA spike rec. (%)
	Accept. criteria	< 0.01	> 0.05	< 0.100	50-200	< 0.01	> 0.05	< 0.100	50-200	< 0.01	> 0.05	< 0.100	50-200
#1	1:4[a]	0.00	1.00	0.068	100	NA	NA	NA	NA	NA	NA	NA	NA
#1	1:8[b]	0.00	1.00	0.068	187	NA	NA	NA	NA	NA	NA	NA	NA
#1	1:16	0.00	1.00	0.068	131	NA	NA	NA	NA	NA	NA	NA	NA
#1	1:32	0.00	1.00	0.068	145	0.00	0.98	0.037	172	0.00	0.98	0.037	210
#1	1:64	0.00	1.00	0.068	124	0.00	0.98	0.037	126	0.00	0.98	0.037	167
#1	1:128	0.00	1.00	0.068	110	0.00	0.98	0.037	129	0.00	0.98	0.037	109
#2	1:4[a]	0.00	1.00	0.040	114	NA	NA	NA	NA	NA	NA	NA	NA
#2	1:8[b]	0.00	1.00	0.040	373	NA	NA	NA	NA	NA	NA	NA	NA
#2	1:16	0.00	1.00	0.040	191	NA	NA	NA	NA	NA	NA	NA	NA
#2	1:32	0.00	1.00	0.040	168	0.00	1.00	0.042	238	0.00	1.00	0.042	208
#2	1:64	0.00	1.00	0.040	144	0.00	1.00	0.042	133	0.00	1.00	0.042	157
#2	1:128	0.00	1.00	0.040	144	0.00	1.00	0.042	124	0.00	1.00	0.042	109
#3	1:4[a]	0.00	1.00	0.034	64	NA	NA	NA	NA	NA	NA	NA	NA
#3	1:8[b]	0.00	1.00	0.034	149	NA	NA	NA	NA	NA	NA	NA	NA
#3	1:16	0.00	1.00	0.034	126	NA	NA	NA	NA	NA	NA	NA	NA
#3	1:32	0.00	1.00	0.034	127	0.00	0.91	0.034	138	0.00	0.91	0.034	197
#3	1:64	0.00	1.00	0.034	103	0.00	0.91	0.034	129	0.00	0.91	0.034	129
#3	1:128	0.00	1.00	0.034	101	0.00	0.91	0.034	117	0.00	0.91	0.034	103

[a] The wells with 1:4 dilution of DP2 showed a yellow-colored supernatant instead of orange/pink color in other wells. This different color indicates a pH change, probably due to the large amount of product present in the wells.[b] Final selected dilution of the DP (f). PGN and FLA are spiked only for dilutions f, fx2, fx4, so no data are presented for other dilutions, which are labeled NA for not applicable.

Tab. 8: *p*-Value regression, *p*-value non-linearity, basal IL-6 release, and LPS, PGN, and FLA spike recovery for DP3 obtained with PBMC batch#3

		LPS				PGN				FLA			
No. batch DP3	Dilution in medium	*p*-value regress.	*p*-value non-lin.	Basal IL-6	LPS spike rec. (%)	*p*-value regress.	*p*-value non-lin.	Basal IL-6	PGN spike rec. (%)	*p*-value regress.	*p*-value non-lin.	Basal IL-6	FLA spike rec. (%)
	Accept. criteria	< 0.01	> 0.05	< 0.100	50-200	< 0.01	> 0.05	< 0.100	50-200	< 0.01	> 0.05	< 0.100	50-200
#1	1:4[a]	0.00	0.49	0.036	22	NA	NA	NA	NA	NA	NA	NA	NA
#1	1:8[b]	0.00	0.49	0.036	100	0.00	0.70	0.038	262	0.00	0.70	0.038	188
#1	1:16	0.00	0.49	0.036	101	0.00	0.70	0.038	117	0.00	0.70	0.038	173
#1	1:32	0.00	0.49	0.036	96	0.00	0.70	0.038	101	0.00	0.70	0.038	145
#1	1:64	0.00	0.49	0.036	124	NA	NA	NA	NA	NA	NA	NA	NA
#2	1:4	0.00	0.61	0.040	15	NA	NA	NA	NA	NA	NA	NA	NA
#2	1:8[b]	0.00	0.61	0.040	70	0.00	0.95	0.075	125	0.00	0.53	0.052	194
#2	1:16	0.00	0.61	0.040	82	0.00	0.95	0.075	95	0.00	0.53	0.052	184
#2	1:32	0.00	0.61	0.040	102	0.00	0.95	0.075	60	0.00	0.53	0.052	138
#2	1:64	0.00	0.61	0.040	110	NA	NA	NA	NA	NA	NA	NA	NA
#3	1:4	0.00	0.12	0.055	18	NA	NA	NA	NA	NA	NA	NA	NA
#3	1:8[b]	0.00	0.12	0.055	77	0.00	0.95	0.075	258	0.00	1.00	0.049	182
#3	1:16	0.00	0.12	0.055	82	0.00	0.95	0.075	62	0.00	1.00	0.049	164
#3	1:32	0.00	0.12	0.055	120	0.00	0.95	0.075	62	0.00	1.00	0.049	144
#3	1:64	0.00	0.12	0.055	118	NA	NA	NA	NA	NA	NA	NA	NA

[a] Medium: Iscove's Modified Dulbecco's Medium supplemented with human AB serum. [b] Final selected dilution of the DP (f). PGN and FLA are spiked only for dilutions f, fx2, fx4, so no data are presented for other dilutions, which are labeled NA for not applicable

The deviations from the optimized MAT cell culture conditions were as follows:

(i) increasing the incubation temperature from 37°C to 38°C or decreasing it to 36°C.
(ii) increasing the CO_2 concentration from 5% to 7% or decreasing it to 4%.
(iii) increasing the cell concentration/well from 50,000 to 70,000 or decreasing it to 30,000.
(iv) increasing the incubation time from 20 h to 25 h or decreasing it to 17 h.
(v) changing the lot of cryo-PBMC (each lot comprised PBMC from 4 different donors, i.e., 12 donors in total).

Not surprisingly, deliberate deviations from the optimized MAT cell culture conditions increased the variability of dose-response curves for standard endotoxin: This can be seen by comparing the curves in Figure 4A with the curves in Figure 1, which were all obtained under the usual optimized MAT cell culture conditions. Nonetheless, all of the curves in Figure 4A still complied with both acceptance criteria for the standard curve, and as can be seen from Figure 4B, the various changes made to assay conditions had little effect on the MAT as assessed by the recovery of an endotoxin spike of 0.16 EU/mL (EC_{50}), with the recoveries for all 90 endotoxin spikes being between 50% and 200%, with a mean recovery of 101.3%. Further, none of the above 5 changes to MAT cell culture conditions prevented the (deliberately sub-optimum) MATs from detecting responses above the LOQ for FLA 12.5 ng/mL, PAM 2.5 ng/mL, PGN 5 µg/mL, HKSA 0.25 x 10^6 cells/mL (approx. EC_{50} concentrations, data not shown).

3.2 Product-specific validation of 3 mAbs

The product-specific validation demonstrated freedom from factors in the product or its matrix that interfered in either responses of the PBMCs to endotoxin or to NEPs or the IL-6 ELISA.

(I.) Responses of the PBMCs to endotoxin and to NEPs

The test for interference factors was performed on 3 batches of each of 3 drug products (DP1, DP2 and DP3, 3 mAbs) with the following determined:

(i) The minimum dilution of each DP that reliably gave 50-200% spike recovery of standard endotoxin = f.
(ii) Standard endotoxin: p-value for regression (should be $p < 0.01$), *p*-value for non-linearity (should be $p > 0.05$), basal IL-6 release (should be < 0.100), standard endotoxin (0.16 EU/mL, EC_{50}) spike recovery (should be 50-200%).
(iii) PGN: *p*-value for regression (should be $p < 0.01$), p-value for non-linearity (should be $p > 0.05$), basal IL-6 release (should be < 0.100), PGN (5 µg/mL, EC_{50}) spike recovery (should be 50-200%) (values for PGN calculated from the endotoxin standard curve).
(iv) FLA: *p*-value for regression (should be $p < 0.01$), p-value for non-linearity (should be $p > 0.05$), basal IL-6 release (should

be < 0.100), FLA (12.5 ng/mL, EC_{50}) spike recovery (should be 50-200%) (values for FLA calculated from the endotoxin standard curve).

For DP1, DP2, and DP3, the minimum dilution of each DP that reliably gave 50-200% endotoxin spike recovery, f, was 1 in 4 for DP1, 1 in 32 for DP2, and 1 in 8 for DP3. These minimum product dilutions are based on consistent endotoxin spike recoveries for all 3 batches. The 3 dilutions, f, fx2 and fx4, were all well within the specified MVD (= CLC/LOQ) of the products.

The *Ph. Eur*. requires that at least one dilution from f, fx2 and fx4 gives 50-200% endotoxin spike recovery and also at least one non-endotoxin TLR ligand, NOD2 ligand (NEP), spiked into the product is detected. This requirement was met for all 3 batches of each of the 3 DPs (Tab. 6, 7, 8).

(II.) IL-6 ELISA:

Doubling dilutions of standard IL-6 from 1,000 pg/mL to 15.6 pg/mL were tested in the IL-6 ELISA and gave responses (ODs) in the absence and presence of product dilution f (1 in 4 for DP1, 1 in 32 for DP2, and 1 in 8 for DP3) that were as follows:

- DP1 batch 1, 98-109%; DP1 batch 2, 98-105%; DP1 batch 3, 98-108%.
- DP2 batch 1, 98-108%; DP2 batch 2, 99-113%; DP2 batch 3, 97-115%.
- DP3 batch 1, 99-113%; DP3 batch 2, 98-117%; DP3 batch 3, 98-112%.

All values were within the *Ph. Eur*. requirement of 80-120%.

4 Discussion

A product license application for a new product, e.g., a parenteral monoclonal antibody or a vaccine that is not of bacterial or viral origin (CFR, 610.13), is required to include data that shows that the product is free from pyrogenic contaminants. For the European Medicines Agency (EMA) and other regulatory agencies, this means data from the MAT is required, whereas the US Food and Drug Administration (FDA) requires data from the RPT. This situation requires pharmaceutical manufacturers to carry out both the MAT and the RPT (as well as the BET) on at least the PPQ production batches of new parenteral products. Although, strictly speaking, pharmacopeial methods are validated and so require only a method verification, a full GMP method validation was carried out in this study to validate *Ph. Eur*. 2.6.30 MAT Method A with 3 new mAbs. In addition, experiments were performed to evaluate the overall robustness of the chosen MAT to help facilitate its wider acceptance in the regulatory community, which should lead to a consequent reduction in RPTs. The additional experiments permitted the evaluation of accuracy, precision, linearity, LOQ, range, and specificity, as well as robustness, and any impact of using different lots of cryopreserved PBMC and/or different NEPs (TLR ligands, NOD2 ligand). There were differences in the responses of the different PBMC lots to the NEPs PAM, PGN and HKSA. The reactivity to FLA was more consistent between the different PBMC lots. All 4 tested NEPs were detected in MATs with all 3 PBMC lots. This testing confirmed that the chosen test system detects, in addition to bacterial endotoxin, non-endotoxin proinflammatory or pyrogenic contaminants. In future MATs, it is recommended that each new PBMC lot is qualified with the 4 NEPs, each at 4 different concentrations. Further, it is recommended that at least one NEP is used as positive control on each plate in each MAT.

The current guidance notes state that: "*Method A is not appropriate if the results of different dilutions* [of the contaminated product] *(endotoxin equivalents per millilitre) show that the dose-response curve* [for the contaminated product] *is not parallel to the standard endotoxin curve*" (EP, 2017). Since batches of pyrogen-contaminated new products are almost never available and unlikely to ever be available for most products, such appropriateness or otherwise can almost never be demonstrated for a product. Further, such a parallel line assay would only likely be valid when the pyrogenic contamination comprised exclusively endotoxin or largely endotoxin, making Method A effectively another endotoxin test, albeit a more labor-intensive, time-consuming and costly one than the BET, which is already carried out. Moreover, a parallel line assay could actually underestimate the level of pyrogenic contamination where the contamination comprises endotoxin (a TLR4 ligand) together with one or more ligands for other TLR receptors, resulting in a response to the contaminated product that may be parallel, at least over part of its length, to the curve for standard endotoxin, but which stimulates a larger maximum response due to synergy between different ligands stimulating cytokine release via different receptors. To ensure that the pyrogenic contaminants content is not underestimated, it is necessary to thoroughly validate the MAT with a broad range of NEPs, as described here, and to read off the curve for standard endotoxin the strongest response to the product that can be obtained from the minimum dilution that gives 50-200% endotoxin spike recovery, f, and also fx2 and fx4. This ensures that the maximum obtainable response to the product is captured and quantified in EE since, to date, NEPs in products have been best detected at minimum dilution of the product (Poole and Patel, 2010).

The *Ph. Eur*. states that: "*where contaminated/adverse drug reaction-positive/rabbit-positive batches of a product are not available, validation of the test system is to include at least 2 non-endotoxin ligands for TLR receptors, at least one of which is to be spiked into the product*." The *Ph. Eur*. does not state or assume that responses to all or to any NEP, non-endotoxin ligand for TLR receptors (TLRs) or any other non-endotoxin monocyte-activating contaminant, will dilute parallel to the standard endotoxin curve since ligands acting via different receptors rarely give parallel curves, and when they do it is likely a matter of coincidence. There is no standard or reference preparation for any TLR ligand or NOD2 ligand, and these molecules can be labile with variable/unstable intrinsic activities; also, these molecules may be contaminated with endotoxin. The present study included not 2 but 4 TLR/NOD2 receptor ligands, with not 1 but 2 of them spiked into the products. This was done to provide a large body of information about the MAT and to better decide which TLR/NOD2 receptor ligands were the most appropriate to use in subsequent product-specific validations. Currently, PGN is preferred as a positive control because, alone of the NEPs/TLR

ligands/NOD2 ligands suggested in the *Ph. Eur*., PGN has been found to contaminate a medical product and to cause adverse reactions (Martis et al., 2005). However, based on the availability of an international standard or reference preparation(s), the positive control(s) may change in future.

The MAT described here employed a 7-point standard endotoxin curve in order to define the upper and lower asymptotes. The *Ph. Eur*. specifies a minimum of 4 concentrations of standard endotoxin for the standard curve, but such a 4-point curve, or even a standard curve comprising 5 concentrations of standard endotoxin, was insufficient to properly define the asymptotes. The use of a 7-point standard curve permitted parameters to be determined that allowed dose-response data for standard endotoxin to be identified that did not fit the theoretical (S-shaped) curve for the 4PL. These parameters were additional to the two acceptance criteria for the standard curve prescribed in *Ph. Eur*. 2.6.30 and comprised: the (estimated) EC_{50}, the range of the CI for the EC_{50}, and the range of the CI for the upper asymptote. The use of a 7-point standard curve and the additional parameters provided information about how best to address the outliers and hook effects that can occur in this type of assay and about the useable, i.e., linear, part of the standard endotoxin curve, which was 0.08-0.32 EE/mL.

The reasoning behind the single 95% β-expectation TI for the total error (USP, 2018) was that the accuracy or total error simplifies the validation of an analytical procedure since the accuracy or total error evaluates the related risk associated with the future use of the MAT, which is regarded as more relevant than investigating all of the usual validation criteria independently, i.e., relative bias, linearity, intermediate precision, repeatability (ICH Q2 R1, 1994, 1996). However, the analysis of the large body of good-quality data that was obtained, with values for relative bias, linearity, intermediate precision, repeatability, robustness, etc. that were typical for a bioassay using primary cells, revealed that the 95% β-expectation TI for the total error was not realistic, likely at least in part due to the intrinsic variability of primary cells. Consequently, it was decided to use the data for relative bias, linearity, intermediate precision, and repeatability as the basis for assay validation.

The MAT described above was deemed to be validated for its intended use with three mAbs having met all of the acceptance criteria, indicating no product interference, detection of endotoxin and NEPs (TLR ligands, NOD2 ligand), and no interference of the products in the detection system. This validation study is the first GMP validation of *Ph. Eur*. 2.6.30 Method A with products, with much of the data presented here for one of the antibodies included in a successful product license application to the EMA.

References

EP – European Pharmacopoeia (1986). Chapter 2.6.8. Rabbit Pyrogen Test.

EP (2012). Chapter 2.6.14. Bacterial Endotoxin Test.

EP (2017). Chapter 2.6.30. Monocyte Activation Test.

Gaines Das, R. E., Brügger, P., Patel, M. et al. (2004). Monocyte activation test for pro-inflammatory and pyrogenic contaminants of parenteral drugs: Test design and data analysis. *J Immunol Methods 288*, 165-177. doi:10.1016/j.jim.2004.03.002

ICH Q2 R1 (1994, 1996). Validation of Analytical Procedures.

Martis, L., Patel, M., Giertych, J. et al. (2005). Aseptic peritonitis due to peptidoglycan contamination of pharmacopoeia standard dialysis solution. *Lancet 365*, 588-594. doi:10.1016/S0140-6736(05)70800-X

Poole, S., Thorpe, R., Meager, A. et al. (1988). Detection of pyrogen by cytokine release. *Lancet 8577*, 130. doi:10.1016/S0140-6736(88)90338-8

Poole, S. and Patel, M. (2010). Monocyte activation test better able to detect non-endotoxin pyrogenic contaminants in medical products. US Patent No 7,736,863 B2 June 15.

USP – US Pharmacopeia General Chapter <1210> (2018). Statistical Tools for Procedure Validation. USP 41/NF36.

Conflict of interest

Sanquin Diagnostics (MMB) performs MAT testing for customers. Sanquin Reagents (EG) produces and distributes MAT kits. All other authors have no conflicts of interest.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

The authors thank NIBSC (UK) scientists Drs Lucy Studholme, Karin Nordgren, and Janet Sutherland for helpful advice and the sharing of NIBSC MAT protocols.