This article was originally published in a journal published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues that you know, and providing a copy to your institution's administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

http://www.elsevier.com/locate/permissionusematerial

ELSEVIER

Journal of Immunological Methods 316 (2006) 42–51

JIM
Journal of Immunological Methods

www.elsevier.com/locate/jim

Research paper

International validation of pyrogen tests based on cryopreserved human primary blood cells

Stefanie Schindler [a], Ingo Spreitzer [b], Bettina Löschner [b], Sebastian Hoffmann [c], Kilian Hennes [d], Marlies Halder [c], Peter Brügger [e], Esther Frey [e], Thomas Hartung [c,*], Thomas Montag [b]

[a] *Biochemical Pharmacology, University of Konstanz, Universitätsstr.10, D-78457 Konstanz, Germany*
[b] *Paul-Ehrlich Institute, Paul-Ehrlich-Strasse 51-59, D-63225 Langen, Germany*
[c] *European Centre for the Validation of Alternative Methods, Institute for Health and Consumer Protection, European Commission, Joint Research Centre, Via Fermi 1, I-21020 Ispra, Italy*
[d] *Qualis Laboratories, Blarerstr. 56, D-78462 Konstanz, Germany*
[e] *Biological Analytics, Novartis Pharma AG, CH-4002 Basel, Switzerland*

Received 21 November 2005; received in revised form 22 June 2006; accepted 13 July 2006
Available online 7 September 2006

Abstract

Pyrogens as fever-inducing agents can be a major health hazard in parenterally applied drugs. For the control of these contaminants, pyrogen testing for batch release is required by pharmacopoeias. This has been done either by the in vivo rabbit pyrogen test (since 1942) or the limulus amoebocyte lysate test (LAL), since 1976. New approaches include cell-based assays employing in vitro culture of human immune cells which respond e.g. by cytokine production (IL-1β; IL-6) upon contact with pyrogens. Six variants of these assays have been validated in a collaborative international study. The recent successful development of cryopreservation methods promises to make standardized immunoreactive primary human blood cells available for widespread use. Furthermore, the pretesting of donors for infectious agents such as HIV or hepatitis has made it possible to develop a safe and standardised reagent for pyrogen testing. Using a total of 13 drugs, we have validated the pyrogen test based on fresh and cryopreserved human whole blood in four laboratories. The test reached >90% sensitivity and specificity. In contrast to the LAL, the test was capable of detecting non-endotoxin pyrogens derived from Gram-positive bacteria or fungi.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Pyrogen testing; Validation study; IL-1β; Cryopreservation

Abbreviations: CV, coefficient of variation; DL, developing laboratory; DMSO, dimethyl sulfoxide; ECVAM, European Centre for the Validation of Alternative Methods; ELC, endotoxin limit concentration; ELISA, enzyme-linked immunosorbent assay; IU, international unit; GLP, good laboratory practice; h, hours; IL, interleukin; IPT, in vitro pyrogen test; LoD, Limit of Detection; LPS, lipopolysaccharide; LTA, lipoteichoic acid; MVD, maximum valid dilution; PEI, Paul Ehrlich Institute; PM, prediction model; PPC, positive product control; NL, naive laboratory; NPC, negative product control; OD, optical density; SOP, standard operating procedure; WHO, World Health Organisation.

* Corresponding author. Tel.: +39 0332 786256; fax: +39 0332 786297.
E-mail address: Thomas.HARTUNG@cec.eu.int (T. Hartung).
URL: http://www.qualis-laboratorium.com (K. Hennes).

0022-1759/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jim.2006.07.023

1. Introduction

Testing for pyrogens has employed animals, either the rabbit (since 1942) or the horseshoe crab, *Limulus polyphemus* (since 1976). Pyrogens, especially lipopolysaccharide (LPS) as part of the cell walls of Gram-negative bacteria, are a ubiquitous threat to human health due to their stability. Other pyrogens are LTA (Morath et al., 2001), exotoxins (Harada et al., 1968; Brunson and Watson, 1974; Blackman and Woodland, 1995), muramyldipeptide (Nakatani et al., 2002) or peptidoglycan (Harada et al., 1968; Rotta, 1975; Schleifer, 1975). Alternative pyrogen tests, comprising 6 cell-based in vitro pyrogen tests using the production of inflammatory mediators, (i.e. IL-1β, TNF-α or IL-6 as well as neopterin) (detailed by Poole et al., 1988; Taktak et al., 1991; Eperon and Jungi, 1996; Hartung and Wendel, 1995; Peterbauer et al., 1999, 2000) of human blood cells have been developed and validated in an international collaborative study (Hoffmann et al., 2005a).

Subsequently, two variants, one using freshly drawn human whole blood and measuring the IL-1β response (Fennrich et al., 1999; Hoffmann et al., 2005a) and another using freshly drawn isolated PBMCs and measuring IL-6 (Poole et al., 1988), have undergone further development and refinement by the introduction of a standardized freezing procedure using DMSO as a cryoprotective agent. This allows both methods to become more standardized and more widely available. For the human whole blood test (IPT), two variations of freezing were developed and optimized. The newer one was optimized for storage at −80 °C, whereas the other one making use of storage in liquid nitrogen has been described previously (Schindler et al., 2004).

In a second study identical to the one previously performed, the refined human whole blood assay has been validated. This study focused again on LPS as a pyrogen, since it is the most frequent contamination, and the ability of the cell-based assays to detect non-endotoxin pyrogens has been demonstrated previously.

Table 1
Laboratories performing the assays

Test system	Developing lab (DL)	1st lab (NL 1)	2nd lab (NL 2)	3rd lab (NL 3)
IPT IL-1β	University of Konstanz*	Paul-Ehrlich Institute	Qualis laboratories	Novartis Pharma

Since the developing lab of the IPT (endpoint IL-1β) cannot work under GLP, the assay was performed by three naïve labs (PEI, NL 1, Qualis, NL 2 and Novartis, NL 3). The IPT was performed in three variant formats, one involving fresh blood and two involving frozen whole blood stored at −80 °C and liquid nitrogen, respectively. (*The results of this lab did not enter the formal evaluation since it is not operating under good laboratory practice (GLP)).

2. Materials and methods

The test system was validated in the developing lab and in three different naïve labs (Table 1) after a detailed SOP had been compiled by the developing lab (DL) and made available for the naïve labs (NL) by ECVAM. The successful technology transfer had been assessed in a prevalidation phase.

2.1. Endotoxin stimulus

The second international WHO standard for endotoxin 94/580 from *E. coli* O113: H10, which was used in the previous validation study, served again as the standard endotoxin (Poole et al., 1997). 100 pg/ml of this endotoxin are equivalent to 1 IU (International Unit)/ml.

3. Fresh and cryopreserved human whole blood test (IL-1β release assay)

3.1. Blood collection

Blood from healthy donors was collected into heparinized tubes (Li-Heparin, 15 IU/ml) (Sarstedt-monovette, 7.5 ml, Nümbrecht, Germany) using a multifly needle set and used within 4 h. Additionally, for the preparation of cryopreserved blood, a sample from each of the five donors was drawn into Serum- and EDTA tubes (Sarstedt) for differential blood counts and infection serology. For the fresh blood assay, one single donor was used and his blood was subjected to a cell count (Pentra 60, ABX Diagnostics, Montpellier, France), in order to exclude infections.

3.2. Testing for infectious agents in cryopreserved blood

The additionally drawn blood of each donor was tested by a qualified laboratory (Dr. U. Brunner, Konstanz, Germany) for Hepatitis A, B, C and HIV according to the standards for blood donations for transfusion purposes in Germany. In the meantime, the blood samples were frozen.

3.3. Freezing for storage at −80 °C (Method A)

Endotoxin-free Soerensen Buffer (Acila GMNmbh, Mörfelden-Walldorf, Germany) was mixed with 20% v/v endotoxin-free DMSO (Wak-Chemie Medical GmbH, Steinbach, Germany). 1.8 ml cryotubes (Nalge Nunc International, Denmark) were unscrewed under a laminar flow and 0.6 ml of the pooled blood of five donors was pipetted into each one. Using a multipette

with a 5 ml combitip, the cryoprotective solution was added in three aliquots of 200 μl each, gently swirling the blood in between. The tubes were closed and placed in a storage box (Nalge Nunc), leaving a space of about 1 cm between each vial in order to ensure a homogenous freezing process. The boxes were then placed in a freezer at −80 °C and left to freeze. They were stored at −80 °C.

3.4. Freezing for storage in liquid nitrogen (Method B)

An alternative freezing procedure involved a controlled freezing process using the vapour phase of liquid nitrogen and has been described earlier (Schindler et al., 2004). For this, the DMSO was added directly to the blood of the individual donors at 10% final concentration (v/v). The blood was then pooled and aliquoted at 1.2 ml/vial and frozen in a computer-controlled freezer (Nicool Plus PC, Air Liquide, Marne-la-Vallée Cedex 3, France) until the temperature reached −120 °C. The aliquots were then taken out of the machine and placed in the vapour phase of liquid nitrogen.

3.5. Shipment

The required number of aliquots of frozen blood was shipped to the other laboratories using a transporterer containing liquid nitrogen (Air Liquide Kryotechnik GmbH, Düsseldorf, Germany). The temperature was monitored using a computer-controlled temperature probe (Thermory Mobile, Air Liquide, France; software Logiciel Recwin, Marne la vallée Cedex, France). After arrival, the aliquots were kept either in the transportation vehicle itself or transferred to the vapour phase of liquid nitrogen, if available. Alternatively, the aliquots that had been frozen at −80 °C (Method A) could be retransferred to the freezer at −80 °C.

3.6. Thawing procedure

The vials were taken out of the transporterer/the nitrogen tank and placed immediately in an incubator at 37 °C. After 15 min, the blood was pooled in a centrifuge tube if more than one aliquot was used, and gently swirled in order to ensure complete mixing.

3.7. Pretesting of the aliquots

After all donors had clearly tested negative for the key infectious agents, the cryopreserved blood was pretested before sending it out to the other laboratories by carrying out a dose–response curve using the WHO standard or an endotoxin calibrated to it. The criteria were an absorbance for the saline control of 0.1 OD or lower, and a response to the 0.5 IU/ml of at least 1.6 times the OD of the saline control.

3.8. Incubation procedure

3.8.1. Cryopreserved blood (Methods A and B)

3.8.1.1. Method A. 180 μl of RPMI (Charles River Endosafe), 20 μl sample/control and 40 μl of thawed blood were added to a pyrogen-free microtiter plate (Falcon). After adding the blood, the contents of the wells were mixed by gently aspirating/dispensing 5 times using a multichannel pipette and sterile, pyrogen-free tips, changing the tips in-between the rows in order to avoid cross-contamination. The plates were then covered with a lid and placed in an incubator at 37 °C and 5% CO_2 for 10–24 h.

3.8.1.2. Method B. The alternative version using nitrogen-stored blood was handled as above, except that the incubation involved 200 μl of RPMI, 20 μl of samples/controls and 20 μl of thawed blood, since the blood had not been prediluted in the freezing process.

3.8.2. Fresh blood (Method C)

In order to allow direct comparisons, the method validated in 2005 (Hoffmann et al., 2005a) was adapted from 1 ml vial incubation to microtiter plates as used for the cryopreserved blood. This variant was also included in the validation to exclude an effect of this changed format.

200 μl of saline (Charles River Endosafe, Charleston, South Carolina, USA), 20 μl sample/control and 20 μl of blood were added to a pyrogen-free microtiter plate (Falcon 96-well flatbottom tissue culture plate, Becton Dickinson Labware, Meylan Cedex, France). After adding the blood, the contents of the wells were mixed by gently aspirating/dispensing 5 times using a multichannel pipette and sterile, pyrogen-free tips, changing the tips in between the rows in order to avoid cross-contamination. The plates were then covered with a lid and placed in an incubator at 37 °C for 10–24 h.

3.8.3. ELISA procedure

The IPT Kit was used (Charles River Endosafe). Aliquots of 100 μl of each well of the incubation plates were added to the wells of the ELISA plate. When transferring the supernatants, they were mixed by aspirating and dispensing them 2–3 times using a multichannel pipette. The ELISA was done according to the manufacturer's instructions.

3.8.4. Data analysis

Data analysis was the same as in the previous validation. The quality criterion for acceptable variability, i.e. allowing a maximum coefficient of variation (CV) of 0.45 has been empirically established in the previous study for both assays in order to ensure the interpretability of test results and was adopted here. The data of those samples and control exceeding this CV-value were tested for outliers by the Grubbs-test. If an outlying replicate caused the excessive variation it was excluded and further analysis was performed with the remaining three replicates. The samples and controls, whose large variation was not caused by an outlier were excluded from further analysis. In the cases when the positive product control (PPC) CV exceeded 45%, the corresponding 0.5 IU/ml in saline, which was part of the dose–response curve using the WHO standard endotoxin, were considered instead. If the CV of this standard also exceeded 45%, the whole set of data was not considered for analysis. Furthermore, the response of the 0.5 IU/ml had to be significantly higher, i.e. a p-value below 0.01, than the respective response of the negative saline control.

Accepted data were analysed by a so-called prediction model (PM): the data of a blinded sample were compared with the PPC data or, if the PPC did not fulfil the quality criterion, the 0.5 IU/ml control using a one-sided t-test with log-transformed data and a local significance level of 1%.

3.8.5. Blinding procedure

All test items are registered medicinal products and were obtained from a pharmaceutical supplier. For the validation, test items and endotoxin spiking samples were prepared by the University of Konstanz and blinded/coded under GLP by personnel (G. Bowe and J. de Lange) from ECVAM, Italy. These were then shipped by the University of Konstanz to each of the appropriate test facilities participating in the study.

3.8.6. Prevalidation

The drugs used were the same as in the previous validation study (Hoffmann et al., 2005a), that is Gelafundin, a volume-replacement therapy for transfusion with high protein content (B. Braun Melsungen AG, Melsungen, Germany), Jonosteril, an electrolyte infusion (Fresenius AG, Bad Homburg, Germany) and Haemate, a factor VIII preparation (Aventis Behring GmbH, Marburg, Germany). Additionally, a negative and a positive control (0.5 IU/ml) were included in each run.

This set was tested in the developing laboratory (DL) Konstanz as well as in two naive laboratories PEI (NL 1) and Qualis (NL 2) with the three approaches of the IPT (Methods A–C) in order to establish successful transfer.

Prior to preparing the spikes, an interference test was performed with all three substances by the DL. Although this had been done in the previous validation, a shift in the interference due to the DMSO/freezing process could not be excluded. Interferences differed for the IPT (data not shown), and the spikes were calculated according to the required dilution.

Sample preparation and blinding was done at the University of Konstanz using pyrogen-free clinical-grade saline and the WHO reference standard endotoxin.

3.8.7. Validation

The same ten drugs were employed for the validation phase, in order to maximise comparability with the previous validation study. The concentrations were based on a recent in-depth analysis of the fever response of a sensitive rabbit strain (Hoffmann et al., 2005b): Five blinded spikes, two of them defined as non-pyrogenic, that is below 0.5 IU/ml (0 and 0.25 IU/ml), and three as pyrogenic (2×0.5 and 1.0 IU/ml) were tested in the different laboratories. All drugs were tested at their MVD (maximum valid dilution), thus adopting the rationale of the pharmacopoeial LAL reference (limit) test. The MVD is calculated from the endotoxin limit concentration (ELC) in IU/ml, defined for a drug by the European Pharmacopoeia (Council of Europe, 2001), divided by the threshold of pyrogenicity as the limit of detection (LoD), in this case 0.5 IU/ml. Drugs, sources, ELCs and MVDs are summarized in Table 2.

4. Results

4.1. Pretesting of the cryopreserved blood

The blood was tested employing an *E. coli* O113: H10 dose–response curve (Fig. 1). The blood was considered suitable since the OD of the saline control was below 0.1 OD and the mean OD of the 0.5 IU/ml was 1.6 times the mean OD of the saline or higher.

4.2. Prevalidation

4.2.1. Method A — cryoblood, −80 °C version

The data produced with the method based on cryoblood frozen at −80 °C are summarized in Fig. 2 as an example for all three methods. It has to be noted that for the NL 1 the level approximated to the maximum response level of 4 OD. This might cause problems for discriminating pyrogenic spikes, especially when the positive control of 0.5 IU/ml produces such high-level

Table 2
Drugs employed in the validation

Drug	Code	Source	Agent	Indication	MVD
Glucose 5% (w/v)	GL	Eifel	Glucose	Nutrition	70
Ethanol 13% (w/w)	ET	B. Braun	Ethanol	Diluent	35
MCP®	ME	Hexal	Metoclopramid	Antiemetic	350
Syntocinon®	SY	Aventis	Oxytocin	Initiation of delivery	700
Binotal®	BI	Aventis	Ampicillin	Antibiotic	140
Fenistil®	FE	Novartis	Dimetinden-maleat	Antiallergic	175
Sostril®	SO	GlaxoSmith Kline	Ranitidine	Antiacidic	140
Beloc®	BE	Astra Zeneca	Metoprolol tartrate	Heart dysfunction	140
Drug A	LO	–	0.9% NaCl	–	35
Drug B	MO	–	0.9% NaCl	–	70

All ten clinical-grade drugs, that had been used in the previous validation using freshly drawn cells, were used again. All drugs were used at their respective MVD (maximum valid dilution).

responses. The background OD-levels were small in the NL 1 and in the DL. NL 2 provided the data with the background level subtracted. The three standard curves, consisting of the negative control (C−), 0.25, 0.5 and 1 IU, indicate a typical monotone increase in OD-response with increasing concentration.

As Fig. 2 only gives an indication about variability of replicates, the CVs were calculated for all samples and controls for all laboratories. While the major part of the total of 48 CV-values was smaller than 30%, four samples, all of which were unspiked, showed a CV larger than 45%. In all of these, one of the four replicate values was much larger than the others, and thus caused the high variability, and was excluded based on the outlier analysis.

Application of the prediction model (PM) to these data revealed that eleven out of the twelve spikes were classified in the same way in all laboratories. Comparing the laboratories pairwise showed that 34 of the total of 36 single comparisons, i.e. 94.4%, resulted in the same classification.

A preliminary assessment in the final step of the predictive capacity revealed that all negative samples were classified correctly and that one 0.5 IU spike (NL 1: H-0.5) at the rabbit classification threshold was classified as false-negative. In terms of performance parameters, this resulted in a specificity of $18/18=100\%$ and a sensitivity of $17/18=94.4\%$.

4.2.2. *Method B — cryopreserved blood (liquid nitrogen version)*

Background OD-levels were small in the DL and NL 1. NL 2 provided the data with the background level subtracted. The three endotoxin standard curves indicated a higher limit of detection than 0.25 IU/ml for this version of the IPT, with NL 2 achieving a low OD response for the 0.5 IU/ml (data not shown).

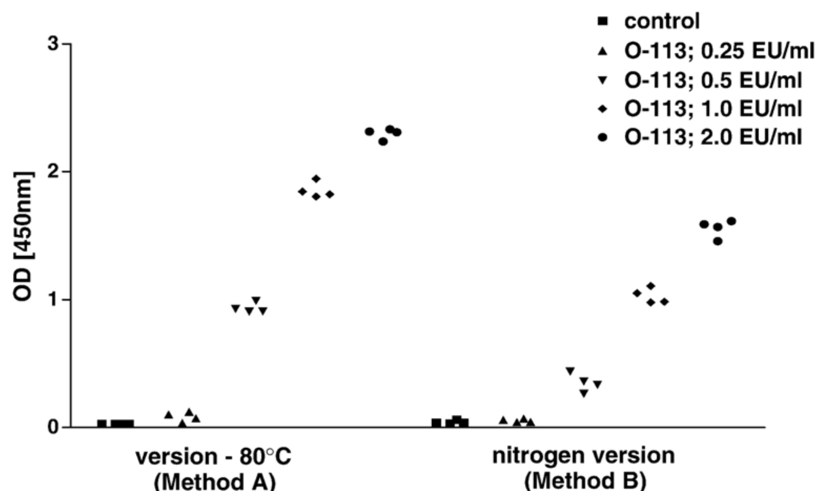

Fig. 1. Pretesting of both versions of the cryopreserved blood prior to shipping. Both versions of the cryopreserved blood (Methods A and B) were pretested measuring the release of IL-1β in response to a pyrogenic stimulus prior to shipping them to the participating labs. For this purpose, a dose–response curve (0.25–2 IU/ml) was prepared using the international WHO standard ($n=4$).

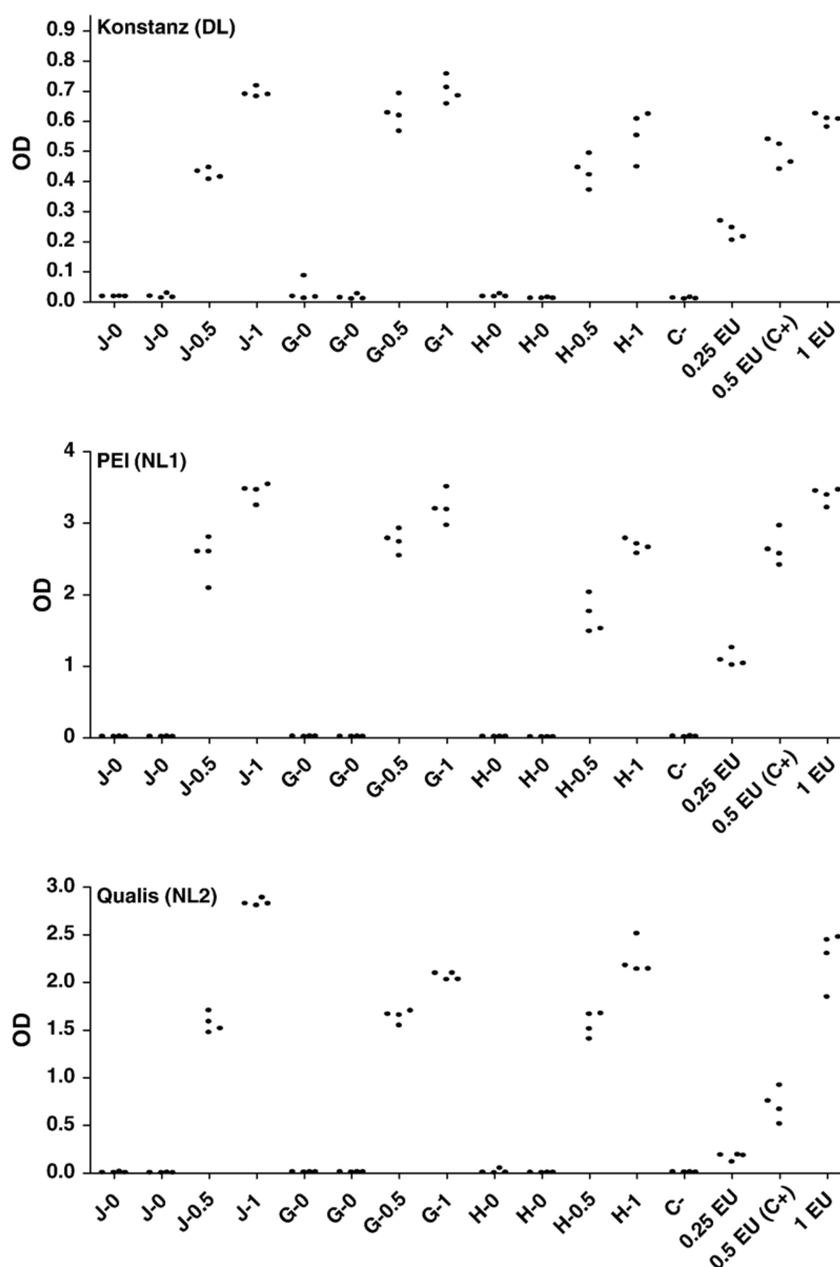

Fig. 2. Prevalidation data for Method A for each of the three involved laboratories. The treatments and controls are abbreviated (J = Jonosteril: G = Gelafundin; H = Haemate; C− = saline; C+ = positive control) indicating the endotoxin contamination in IU (0, 0.5 and 1 IU/ml).

The CVs were calculated for each treatment or control for all laboratories. While the major part of the CVs was smaller than 40%, six samples (mainly from DL) and one standard showed a CV larger than 45%. Nine out of the twelve spikes were classified in the same way in all laboratories. Comparing the laboratories pairwise showed that 30 of the total of 36 single comparisons, i.e. 83.3%, resulted in the same classification.

Preliminary assessment in the final step of the predictive capacity, revealed that one negative sample was classified wrongly (NL 2: J-0) due to one outlying value, and that a Haemate 0.5 IU/ml sample (DL and NL 1) at the rabbit classification threshold was twice classified as a false-negative. In terms of performance parameters, this resulted in a specificity of 17/18 = 94.4% and a sensitivity of 16/18 = 88.9%.

4.2.3. Method C — fresh whole blood

Background OD-levels were small in the DL and NL 1. NL 2 provided the data with the background level subtracted. The three standard curves, consisting of the negative control C− (0 IU/ml), 0.25, 0.5 and 1 IU/ml, showed a typical steady increase in OD-response with increasing concentration.

Table 3
Classification of samples by all methods and all laboratories in validation exercise

Drug	Spike	Truth	IPT Method A			IPT Method B			IPT Method C		
	(IU/ml)		NL 1	NL 2	NL 3	NL 1	NL 2	NL 3	NL 1	NL 2	NL 3
Beloc	0.0	0	0	0	CV	0	0	0	0	0	0
	0.25	0	0	**1**	CV	0	0	CV	**1**	0	0
	0.5	1	1	1	1	**0**	1	1	1	1	1
	0.5	1	1	1	1	1	1	1	1	1	1
	1.0	1	1	1	1	1	1	1	1	1	1
Binotal	0.0	0	0	0	0	0	0	0	0	0	0
	0.25	0	0	**1**	**1**	0	**1**	CV	0	**1**	0
	0.5	1	**0**	1	1	1	1	1	1	1	1
	0.5	1	1	1	1	1	1	1	1	1	1
	1.0	1	1	1	1	1	1	1	1	1	1
Ethanol	0.0	0	0	0	nq	0	0	0	0	0	CV
	0.25	0	CV	**1**	nq	0	0	CV	0	**1**	0
	0.5	1	1	1	nq	1	1	**0**	1	1	1
	0.5	1	1	1	nq	1	1	**0**	1	1	1
	1.0	1	1	1	nq	1	1	1	1	1	1
Fenistil	0.0	0	0	0	0	0	0	0	0	0	0
	0.25	0	0	**1**	**1**	0	CV	1	CV	**1**	**1**
	0.5	1	1	1	CV	CV	1	CV	CV	1	1
	0.5	1	1	1	1	1	1	1	CV	1	1
	1.0	1	1	1	1	1	1	1	1	1	1
Glucose	0.0	0	0	0	0	0	0	0	0	0	0
	0.25	0	CV	0	0	0	**1**	0	0	**1**	**1**
	0.5	1	1	1	1	**0**	1	CV	1	1	1
	0.5	1	1	1	1	**0**	1	1	1	1	1
	1.0	1	1	1	1	**0**	1	1	1	1	1
MCP	0.0	0	0	0	0	0	0	0	0	0	0
	0.25	0	CV	**1**	CV	0	**1**	0	0	0	**1**
	0.5	1	1	1	1	**0**	1	CV	**0**	1	1
	0.5	1	1	1	1	**0**	1	1	1	1	1
	1.0	1	1	1	1	1	1	1	CV	1	1
Sostril	0.0	0	0	0	nq	0	0	nq	0	0	0
	0.25	0	0	**1**	nq	0	1	nq	CV	0	**1**
	0.5	1	**0**	1	nq	**0**	1	nq	1	1	CV
	0.5	1	1	1	nq	1	1	nq	1	1	1
	1.0	1	1	1	nq	1	1	nq	1	1	1
Syntocinon	0.0	0	0	0	0	0	0	nq	0	0	0
	0.25	0	0	CV	CV	CV	CV	nq	0	0	0
	0.5	1	1	1	1	1	1	nq	1	1	1
	0.5	1	1	1	1	1	1	nq	1	1	1
	1.0	1	1	1	1	1	1	nq	1	1	1
A (saline)	0.0	0	CV	0	nq	0	0	nq	CV	0	0
	0.25	0	0	0	nq	0	0	nq	0	0	0
	0.5	1	1	1	nq	**0**	1	nq	CV	1	1
	0.5	1	1	1	nq	**0**	1	nq	1	1	1
	1.0	1	1	1	nq	1	1	nq	1	1	1
B (saline)	0.0	0	0	0	nq	0	0	nq	0	0	0
	0.25	0	0	0	nq	0	0	nq	0	0	CV
	0.5	1	1	1	nq	**0**	1	nq	CV	1	1
	0.5	1	1	1	nq	**0**	1	nq	1	1	1
	1.0	1	1	1	nq	1	1	nq	1	1	1
Sample size *n*			46	49	25	48	48	24	42	50	47
Specificity			100	68.4	75	100	77.8	88.9	94.1	80.0	76.5
Sensitivity			93.3	100	100	62.1	100	86.7	96.0	100	100

Grey shading indicates that for these drugs the PPCs did not qualify so that the PC was used in the PM. CV = sample showed a variability resulting in exclusion, i.e. CV>45% and no significant outlier present. nq = not qualified according to quality criteria, i.e. failure of PPCs and PCs. 0 = considered/classified negative, 1 = considered/classified positive. False classifications are in bold type.

Table 4
Inter-laboratory reproducibility and sensitivity/specificity with the respective sample sizes in validation

Test	Inter-laboratory reproducibility (%)	Sample size: sensitivity	Sensitivity (%)	Sample size: specificity	Specificity (%)
IPT Method A	DL–NL 1: 86.7 DL–NL 2: 87.5 NL 1–NL 2: 100	77	97.4	45	82.2
IPT Method B	DL–NL 1: 66.0 DL–NL 2: 63.3 NL 1–NL 2: 83.3	74	82.4	46	89.1
IPT Method C	DL–NL 1: 88.1 DL–NL 2: 89.7 NL 1–NL 2: 91.5	84	98.8	55	83.6

Inter-laboratory reproducibility was calculated by the proportion of samples classified identically for each pair-wise laboratory comparison.

The CVs were calculated for each treatment or control for all laboratories. In general, the CVs were smaller than 30%. Only two samples resulted in a CV larger than 45%. These two samples were an H-0 tested at NL 2, which was caused by an aberrant value, and a G-0 tested in DL with a CV of 48.8%. Furthermore, a tendency for larger CV of endotoxin-free samples/treatments was observed, as the background OD-level was lower compared to comparable assays, e.g. in the main validation study. Ten out of the twelve spikes were classified in the same way in all laboratories. Comparing the laboratories pairwise, showed that 32 of the total of 36 single comparisons, i.e. 89.9%, resulted in the same classification.

Preliminary assessment in the final step of the predictive capacity revealed that all negative samples were classified correctly and that two 0.5 IU/ml spikes (DL: J-0.5; NL 1: H-0.5), which are at the rabbit classification threshold, were classified as false-negative.

In terms of performance parameters, this resulted in a specificity of 18/18 = 100% and a sensitivity of 16/18 = 88.9%.

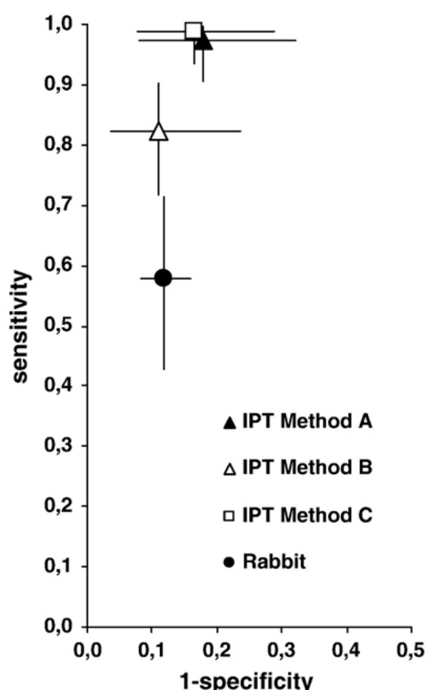

Fig. 3. Sensitivity and specificity of validation with 95% confidence intervals including the rabbit pyrogen test. The sensitivity and specificity of the in vitro methods were assessed after the validation phase and compared to the rabbit pyrogen test assuming a binomial distribution. The *y*-axis shows the sensitivity in percent, the *x*-axis the specificity for the rabbit pyrogen test and all four in vitro assays.

5. Validation

5.1. Inter-laboratory reproducibility

As within-laboratory reproducibility was generally successfully shown in prevalidation, only inter-laboratory reproducibility, based on three laboratories per test method, was assessed in the validation. The similarity of laboratories was based on the classification resulting from the PM (Table 3) to compare the laboratories with each other with respect to concordance, i.e. without taking the true classifications of the samples into account. The results are presented in Table 4. The overall inter-laboratory reproducibilities of IPT Methods A and C were consistently high. Regarding IPT Method B, only about 65% of the samples were

Table 5
Performance of the IPT in the developing laboratory

	IPT A	IPT B	IPT C
Sample size *n*	48	37	47
Specificity	100	100	94.4
Sensitivity	96.7	90.5	86.2

The developing laboratory in Konstanz performed all three methods of the IPT in parallel to the naïve labs. This table shows the sample size that qualified for evaluation and the achieved specificity and sensitivity for these samples.

classified the same way due to NL 3 where four drugs caused problems.

5.2. Predictive capacities

Table 4 summarises the sensitivity and specificity for each method together with the respective sample sizes. For IPT Method A, eight samples at the NL 1 and three at the NL 3 were excluded due to their high variability, i.e. CVs>45%. For IPT Method B and Method C, the sample sizes were reduced for both methods in addition to ten samples with high variability when using four drugs tested at the NL 3 that failed the quality criteria for both the PPC and the PC.

The overall performances of the IPT Methods A and C were very good: High sensitivities over 90% could be achieved, while specificities around 80% were established, reflecting the safety approach in the PM emphasizing sensitivity. In contrast to these methods, the IPT Method B performed differently with a higher specificity of 89% at the cost of a decreased sensitivity (82%). Misclassifications occurred with one exception only for samples with contaminations close to the pyrogenicity threshold, i.e. 0.25 and 0.5 IU/ml. The 95%-confidence intervals for the parameters were calculated, assuming a binomial distribution, and are presented in Fig. 3. The parameters for the rabbit test, calculated with the model of Hoffmann et al. (2005b), were also included. While the new test, except for IPT Method B, had a slightly lower specificity than the rabbit test, the sensitivity was substantially increased by 20% up to 40%.

6. Discussion

This follow-up study had the aim to refine and improve the already established, validated systems for pyrogen testing using freshly drawn human blood and make them available as a safe and standardized reagent in cryopreserved form. In comparison to the fresh blood assays, the cryopreserved cells showed a higher response with regard to IL-1β, but the variance also tended to be higher. Therefore, some sets of data of all of the IPT methods had to be excluded due to variance-controlling quality criteria. In the former validation (Hoffmann et al., 2005a), the whole blood test, at this time performed with fresh blood in reaction tubes, achieved a sensitivity of 72.7% (n=88) and a specificity of 93.2% (n=59). The inter-lab reproducibility had been 72.9 and 81.6% between the developing lab and the two naïve labs. In this study, the sensitivity of the fresh blood assay could be improved by transferring it to the microtiter plate (98.8%) with a minor reduction of specificity (83.6%). This approach proved to be more easily transferable to the naïve labs (inter-lab reproducibility 88–92%). The frozen cells performed well, although the −80°C method (A) was more readily transferred than the nitrogen method (86–100% vs. 63–83% for the interlaboratory reproducibility). However, one of the laboratories (NL 3) had particular problems with both cryopreserving methods as the data of four drugs failed quality criteria and thus were excluded from analysis.

We assume that the IPT assay at the NL 3 posed problems in performance because of insufficient transfer of the method. This problem could not be overcome because of the very tight time schedule of the validation. Nevertheless, IPT-data generated in parallel at Konstanz (Table 5) yielded good results, although they did not enter the formal evaluation according to the study protocol which foresaw only the participation of three GLP-concordant laboratories. The importance of successful assay transfer is stressed considering the fact that, had the results of the developing lab instead of those of NL 3 entered the evaluation, the sample size for Method A qualifying for evaluation had increased from 120 to 143.

The specificity of this method would have increased slightly to 89.5% and the sensitivity would hardly have changed at all, indicating that the method of evaluation used in this study was adequate. In general, the −80 °C version produced higher ODs and was more sensitive than the nitrogen version of the cryoblood. It was noticeable, though, that the linear range in the dose–response curve from 0.25 IU/ml to 1 IU/ml is much smaller in the −80 °C than in the nitrogen version. All in all, this validation has shown that the novel, recently validated pyrogen test based on human blood can be performed with cryopreserved cell preparations. The IPT could be improved with regard to performance making it at the same time easier to handle by transferring it to the microtiter plate. The optimization employing cryopreserved cells permits the assay to become more standardized and cells as test reagents to be more widely available. The fact that safety standards of blood transfusions can be implemented as shown for the IPT suggests that concerns regarding possibly infected donors can be ruled out.

References

Blackman, M.A., Woodland, D.L., 1995. In vivo effects of superantigens. Life Sci. 57, 1717.

Brunson, K.W., Watson, D.W., 1974. Pyrogenic specificity of streptococcal exotoxins, staphylococcal enterotoxin, and gram-negative endotoxin. Infect. Immun. 10, 347.

Council of Europe, 2001. Biological tests, 2.6.14. Bacterial endotoxins. European Pharmacopoeia. Council of Europe, Strasbourg, p. 140.

Eperon, S., Jungi, T.W., 1996. The use of human monocytoid lines as indicators of endotoxin. J. Immunol. Methods 194, 121.

Fennrich, S., Fischer, M., Hartung, T., Lexa, P., Montag-Lessing, T., Sonntag, H.G., Weigandt, M., Wendel, A., 1999. Detection of endotoxins and other pyrogens using human whole blood. Dev. Biol. Stand. 101, 131.

Harada, T., Misaki, A., Saito, H., 1968. Curdlan: a bacterial gel-forming beta-1,3-glucan. Arch. Biochem. Biophys. 12, 292.

Hartung, T., Wendel, A., 1995. Detection of pyrogens using human whole blood. ALTEX 12, 70.

Hoffmann, S., Peterbauer, A., Schindler, S., Fennrich, S., Poole, S., Mistry, Y., Montag-Lessing, T., Spreitzer, I., Löschner, B., van Aalderen, M., Bos, R., Gommer, M., Nibbeling, R., Werner-Felmayer, G., Loitzl, P., Jungi, T., Brcic, M., Brugger, P., Frey, E., Bowe, G., Casado, J., Coecke, S., de Lange, J., Mogster, B., Naess, L.M., Aaberge, I.S., Wendel, A., Hartung, T., 2005a. International validation of novel pyrogen tests based on human monocytoid cells. J. Immunol. Methods 298, 161.

Hoffmann, S., Luederitz-Puechel, U., Montag-Lessing, T., Hartung, T., 2005b. Optimisation of pyrogen testing in parenterals according to different pharmacopoeias by probabilistic modelling. J. Endotoxin Res. 11, 25.

Morath, S., Geyer, A., Hartung, T., 2001. Structure–function relationship of cytokine induction by lipoteichoic acid from *Staphylococcus aureus*. J. Exp. Med. 193, 393.

Nakatani, T., Tsuchida, K., Sugimura, K., Yoshimura, R., Takemoto, Y., 2002. Response of peripheral blood mononuclear cells in hemodialyzed patients against endotoxin and muramyldipeptide. Int. J. Mol. Med. 10, 469.

Peterbauer, A., Werner, E.R., Werner-Felmayer, G., 1999. Further development of a cell culture model for the detection of bacterial pyrogens. ALTEX 16, 3.

Peterbauer, A., Eperon, S., Jungi, T.W., Werner, E.R., Werner-Felmayer, G., 2000. Interferon-gamma-primed monocytoid cell lines: optimizing their use for in vitro detection of bacterial pyrogens. J. Immunol. Methods 233, 67.

Poole, S., Thorpe, R., Meager, A., Hubbard, A.R., Gearing, A.J., 1988. Detection of pyrogen by cytokine release. Lancet 8577, 130.

Poole, S., Dawson, P., Gaines Das, R.E., 1997. Second international standard for endotoxin: calibration in an international collaborative study. J. Endotoxin Res. 4, 221.

Rotta, J., 1975. Endotoxin-like properties of the peptidoglycan. Z. Immunitatsforsch., Exp. Klin. Immunol. 149, 230.

Schindler, S., Asmus, S., von Aulock, S., Wendel, A., Hartung, T., Fennrich, S., 2004. Cryopreservation of human whole blood for pyrogenicity testing. J. Immunol. Methods 294, 89.

Schleifer, K.H., 1975. Chemical structure of the peptidoglycan, its modifiability and relation to the biological activity. Z. Immunitatsforsch. Exp. Klin. Immunol. 149, 104.

Taktak, Y.S., Selkirk, S., Bristow, A.F., Carpenter, A., Ball, C., Rafferty, B., Poole, S., 1991. Assay of pyrogens by interleukin-6 release from monocytic cell lines. J. Pharm. Pharmacol. 43, 578.