ORIGINAL ARTICLE

The CIMT-monitoring panel: a two-step approach to harmonize the enumeration of antigen-specific CD8+ T lymphocytes by structural and functional assays

C. M. Britten · C. Gouttefangeas · M. J. P. Welters · G. Pawelec · S. Koch · C. Ottensmeier ·
A. Mander · S. Walter · A. Paschen · J. Müller-Berghaus · I. Haas · A. Mackensen · T. Køllgaard ·
P. thor Straten · M. Schmitt · K. Giannopoulos · R. Maier · H. Veelken · C. Bertinetti · A. Konur · C. Huber ·
S. Stevanović · T. Wölfel · S. H. van der Burg

Abstract The interpretation of the results obtained from immunomonitoring of clinical trials is a difficult task due to the variety of methods and protocols available to detect vaccine-specific T-cell responses. This heterogeneity as well as the lack of standards has led to significant scepticism towards published results. In February 2005, a working group was therefore founded under the aegis of the Association for Immunotherapy of Cancer ("CIMT") in order to compare techniques and protocols applied for the enumeration of antigen-specific T-cell responses. Here we present the results from two consecutive phases of an international inter-laboratory testing project referred to as the "CIMT monitoring panel". A total of 13 centers from six European countries participated in the study in which pretested PBMC samples, synthetic peptides and PE-conjugated HLA-tetramers were prepared centrally and distributed to participants. All were asked to determine the number of antigen-specific T-cells in each sample using tetramer staining and one functional assay. The results of the first testing round revealed that the total number of cells analyzed was the most important determinant for the sensitive detection of antigen-specific CD8+ T-cells by tetramer staining. Analysis by ELISPOT was influenced by a combination of cell number and a resting phase after thawing of peripheral blood mononuclear cells. Therefore, the experiments were repeated in a second phase but now the participants were asked to change their protocols according to the new guidelines distilled from the results of the first phase.

C. M. Britten (✉) · M. J. P. Welters
Department of Immunohaematology and Blood Transfusion,
Leiden University Medical Center, Leiden, The Netherlands
e-mail: cebritten@web.de

C. Gouttefangeas (✉) · S. Stevanović
Department of Immunology, University of Tuebingen,
Tuebingen, Germany
e-mail: cecile.gouttefangeas@uni-tuebingen.de

G. Pawelec · S. Koch
Centre for Medical Research, University of Tuebingen,
Tuebingen, Germany

C. Ottensmeier · A. Mander
Cancer Sciences Division,
Southampton University Hospitals, Southampton, UK

S. Walter
Immatics Biotechnologies, Tuebingen, Germany

A. Paschen
Skin Cancer Unit of the German Cancer Research Centre,
University Clinics of Mannheim, Mannheim, Germany

J. Müller-Berghaus
Paul-Ehrlich-Institute, Langen, Germany

I. Haas · A. Mackensen
Department of Haematology and Oncology,
University of Regensburg, Regensburg, Germany

T. Køllgaard · P. thor Straten
Department of Haematology,
Centre for Cancer Immune Therapy, Herlev, Denmark

The recommendations improved the number of antigen-specific T-cell responses that were detected and decreased the variability between the laboratories. We conclude that a two-step approach in inter-laboratory testing allows the identification of distinct variables that influence the sensitivity of different T-cell assays and to formally show that a defined correction to the protocols successfully increases the sensitivity and reduces the inter-center variability. Such "two-step" inter-laboratory projects could define rational bases for accepted international guidelines and thereby lead to the harmonization of the techniques used for immune monitoring.

Abbreviations

APC	Antigen presenting cell
ELISPOT	Enzyme linked immuno spot
PBMC	Peripheral blood mononuclear cells
IFNγ	Interferon gamma
ICS	Intracellular cytokine staining
ND	Not determined

Introduction

In the last two decades, more than 200 clinical trials of different anti-tumor vaccines aiming to induce tumor-specific immunity in cancer patients have been described [1]. Most of these trials primarily assessed safety and immunogenicity while reporting partial or complete clinical responses in a minority of patients [2, 3]. Despite the fact

M. Schmitt
Third Department of Internal Medicine,
University of Ulm, Ulm, Germany

K. Giannopoulos
Clinical Immunology Department,
Medical University of Lublin, Lublin, Poland

R. Maier
Research Department, Kantonal Hospital St Gallen,
St Gallen, Switzerland

H. Veelken · C. Bertinetti
Department of Haematology and Oncology,
Freiburg University Medical Centre, Freiburg, Germany

A. Konur · C. Huber · T. Wölfel
Third Medical Department, University Mainz, Mainz, Germany

S. H. van der Burg (✉)
Department of Clinical Oncology,
Leiden University Medical Center, Leiden, The Netherlands
e-mail: shvdburg@lumc.nl

that the low fraction of clinical responders still precludes the establishment of a direct correlation between clinical efficacy and T-cell reactivity, it has become clear from animal models and clinical observations that naturally-occurring or vaccine-induced CD8$^+$ or CD4$^+$ T-cells play an important role in the control and regression of tumors [4–9]. Therefore, the number of subjects that mount a vaccine-induced T-cell response as well as the strength of a detected T-cell response represent important surrogate markers for vaccine efficacy. The enzyme-linked immunospot (ELISPOT) assay [10, 11], staining with HLA-peptide multimers [12] and intracellular cytokine staining (ICS) [13, 14] are technologies used commonly for the monitoring of antigen-specific immune responses. For these three assays, a huge variety of different protocols are available worldwide. This heterogeneity, together with the fact that the sensitivity of the individual protocols can vary significantly, makes a comparison of the results obtained in different trials a difficult task. Moreover, an increasing number of new technologies are constantly being introduced to the field, which makes interpretations even more complex [15–25].

Current data and opinion support the use of a functional assay like the ELISPOT or ICS in combination with a phenotyping assay like HLA-multimers [26, 27], but recognized international standards for all these methodologies are still lacking.

The main aim of the "CIMT monitoring panel" is to harmonize and optimize the monitoring of antigen-specific T-cells among the participating laboratories, based on objective rationales with respect to the testing procedure, the analysis and the interpretation of results. Important requirements for an immunological test are sensitivity, applicability to large amounts of clinical material and feasibility at reasonable cost. The results generated by the tests should be reproducible and sensitive, independently of the place where they have been performed. After the first meeting of the working group, a series of inter-laboratory testing projects was initiated, in which individual laboratories could compare their performance, express their needs and exchange experience in order to improve their local assays. Here we report the results of the first two phases of the CIMT monitoring panel, with 13 participating centers from six European countries.

Materials and methods

Preparation and screening of PBMC samples

Buffy coats from HLA-typed healthy volunteers were kindly provided from the Blood Bank of the University Mainz. HCMV sero-status was known. PBMC were isolated by Ficoll density gradient separation (Pharmacia,

Uppsala, Sweden), washed two times in RPMI 1640 (GIBCO BRL, Grand Island, NY, USA) containing 10 mM Hepes buffer, L-arginine (116 mg/ml), L-glutamine (216 mg/ml), penicillin (10 IU/ml), streptomycin (100 mg/ml) and 10% FCS (GIBCO BRL), counted and frozen at 10 to 20×10^6 cells per cryovial in 1 ml of FCS 90% + DMSO 10% at −80°C in freezing-boxes filled with iso-propanol. After 20 h, all cryovials were transferred to liquid nitrogen and stored until distribution to the participating laboratories.

Pre-screening and selection of the PBMC donors for influenza- and CMV- T-cell reactivities were performed by a central lab using the IFNγ ELISPOT assay following a local protocol as described previously [38]. Five donors were selected for the first phase of the panel and eight for the second phase. One HLA-A*0201-negative donor was included in each phase (negative control), all other samples were HLA-A*0201-positive.

Synthetic peptides and HLA-tetramers

Peptides were synthesized using standard Fmoc chemistry, dissolved at 10 mg/ml in DMSO, aliquotted and stored at −80°C. The purity was checked by reverse-phase HPLC and was found to be >80%. Two known HLA-A*0201 T-cell epitopes were used: influenza MP 58–66 GILGFVFTL and HCMV pp65 495–503 NLVPMVATV (http://www.syfpeithi.de). Biotinylated recombinant HLA-A*0201 monomers folded with the influenza MP 58–66 or the HMCV pp65 peptides were produced essentially as described, purified by gel filtration and stored as aliquots at −80°C [12]. Fluorescent multimers were obtained by incubation with streptavidin-PE (Molecular Probes, Leiden, The Netherlands), then frozen as aliquots after addition of 0.5% BSA and 16% glycerol. HLA-concentrations of influenza-tetramer and HCMV-tetramers were 700 and 350 μg/ml, respectively. Both tetramers were checked by HPLC and/or validated by staining of a specific CD8$^+$ T-cell line (Influenza) or PBMC from HLA-A2-negative and HLA-A2-positive CMV seronegative donors (CMV). Such tetramers are stable at 4°C for at least 1 month (personal observation) and participants were asked to perform all tests within this time period.

Participating centers

Twelve centers from five European countries participated in the first phase of the monitoring panel. As one of the investigators moved to another institution during the study a 13th center from a 6th European country was added to the group in the second phase of the panel. Participation in the panel was open to all interested laboratories with a focus on T-cell monitoring, independently of membership in the Association for Immunotherapy of Cancer.

Reagent distribution and assay guidelines

Coded PBMC samples, synthetic peptides and HLA-A*0201 tetramers were shipped on dry ice to the participants. Additionally, guidelines for the two T-cell assays were distributed for each phase:

Phase I/2005. A protocol for tetramer staining was included. Briefly, 1×10^6 PBMC per test were transferred directly after thawing into one well of a 96 well u-bottom plate and washed in FACS buffer consisting of PBS, 2% FCS, 2 mM EDTA, 0.02% azide. Incubation with 5 μg/ml HLA-tetramer was then performed in FACS buffer with 50% FCS for 30 min at room temperature in the dark. After one wash in FACS buffer, mAb for T-cell staining were added for 20 min at 4°C. Finally, cells were washed twice before fixing in FACS buffer containing 1% formaldehyde solution. Three mAb combinations were proposed, CD8 alone, CD3 plus CD8, or CD4 plus CD8. Each lab could choose here the antibody clones, fluorescent dye and concentrations used. Stainings were performed in duplicate.

For the functional assays, synthetic peptides were diluted at 1 mg/ml in PBS as a stock solution. Concentrations in further tests were 1–10 μg/ml, left to the choice of the participants. There were no recommendation which functional test should be performed, so that each group could choose the test either routinely used, or to be implemented for its own needs. In this first phase, 11/12 laboratories chose the IFNγ ELISPOT assay, one lab (Z10) a FACS-based intracellular IFNγ staining and one lab performed both assays (Z7). Spot counting was performed locally.

Phase II/2006. Following the results obtained in the first testing phase, requirements were introduced and participants were asked to apply exactly these new criteria (two for the tetramer staining, and four for the ELISPOT, see "Results" section). The assay guidelines were modified accordingly. However, in order to reduce the variability in the FACS analysis of the 13 laboratories, a figure showing exemplary dot-plots, settings of gates and quadrants, and statistical analysis was provided. All laboratories were now required to perform an IFNγ ELISPOT as the functional test, with a fixed peptide concentration of 1 μg/ml. Participants were encouraged to use a distributed model protocol but were allowed to use their local protocol, provided that they applied the four new requirements introduced in the second phase.

Collection and analysis of results

After performing the required tests in each phase, participants returned a completed report form containing all relevant information. Number of cells recovered after thawing was included to assess viability after transport. For the tetramer staining experiments, mAb clone, manufacturer,

amount, cytometer type and number of lymphocytes and/or CD8$^+$ cells analyzed were noted. Results were expressed as percentage of tetramer-positive cells among CD8$^+$, CD3$^+$CD8$^{+,}$ or CD4$^-$ lymphocytes, depending on which mAb combination was used for the staining. Additionally, FACS dot-plots containing all gates, quadrants and deduced statistical analysis were collected. For the functional test, medium and thawing procedure (e.g. addition of DNAse, of a resting phase, etc.) had to be described, as well as the number of cells per test, the antibodies used (clone, manufacturer, concentration), the final peptide concentration and the incubation times. For the ELISPOT assay, the type of plate, the enzymatic visualization system and the spot reader were also noted. Absolute spot numbers were given by each participant, and filter plates were kept for possible second analysis.

All results from both phases were collected and centrally analyzed. For the tetramer stainings, the number of lymphocytes, number of CD8$^+$ T-cells and frequencies of tetramer-positive cells were calculated on the basis of the stainings and statistics provided by the participants. Apart from these calculated frequencies, a "visual evaluation" was necessary (see "Results"). For the ELISPOT, analysis was performed based on the spot numbers reported by the participants, followed by a student t test. Results were accepted as positive reaction only when the numbers of antigen-specific spots exceeded the number of spots in the background wells by atleast a factor two. The coefficients of variation (CV) were calculated for all results (CV = SD/mean × 100) and are shown in supplementary Tables S1a, b.

The raw data from both panel phases will be provided to interested readers upon request.

Results

Phase I/2005 of the interlaboratory testing project—general aspects

Coded PBMC samples from four HLA-A*0201-positive and one HLA-A*0201-negative healthy donor (D1–D5) were included in this first testing phase. The thawing procedure for PBMC samples in the test centers was not standardized and the recovery of viable cells varied greatly between 45 and 102% (mean 73%) in the 12 labs. However, the number of cells recovered was in all cases sufficient to perform the required analyses. When all the data from the tetramer staining and functional tests were combined it became clear that subjects D1 and D5 had responded to the HLA-A*0201 restricted CMV-derived peptide, consistent with their CMV seropositive-status, and that subjects D1, D2, D3, and D5 had responded to influenza. In total, each laboratory should in theory have been able to measure six positive (2× CMV and 4× influenza) responses.

Detection of antigen-specific T-cells by tetramer staining and IFNγ ELISPOT

The protocol required that all PBMC samples should be analyzed by the 12 participants for the presence of HLA-A*0201-restricted CMV-specific and influenza-specific CD8$^+$ T-cells using centrally-prepared tetramers. The indicated frequencies of antigen-specific CD8$^+$ cells generally represent the mean of two separate stainings with CD3 Ab/CD8 Ab/tetramer, except for centers Z1 (CD8/tetramer), Z7 (CD3/CD4/CD8/tetramer), Z5 and Z10 (one staining CD3/CD8/tetramer and one staining CD3/CD4/tetramer) and are based on the analysis and dot plots provided by each participant. As illustrated in Fig. 1, the absolute numbers of tetramer-positive T-cells were influenced by the individual decision of where to set the gates and quadrant markers for the analysis. For example, the inclusion of the subset of T lymphocytes expressing CD8 at a low density influenced the number of CD8$^+$ and consequently the frequency of tetramer$^+$ cells. Moreover, non-specific binding of the tetramer (as seen on the CD8-negative subset) also varied between the different laboratories. For these reasons, not only the frequencies, but also the appearance of the tetramer-positive populations was carefully examined. Two parameters were chosen for validation of "positive" results: (1) a clustered, but not diffuse, tetramer binding-population, and (2) strong intensity of tetramer staining, especially marked for the CMV-tetramer-binding population (Fig. 1). Table 1 shows: (I) the minimum, mean and maximum frequencies of antigen-specific CD8$^+$ T-cells, (II) the results obtained from the individual centers Z1–Z12, and (III) the number and percentage of centers that detected a response. The high frequencies of CMV-specific CD8$^+$ T-cells in donors D1 and D5 were readily detected by all participants (mean of 1 per 141 CD8 ± 113 in D1 and mean of 1 per 80 CD8 ± 24 in D5, respectively). For influenza-specific CD8$^+$ T-cells, the results were more variable. Influenza-tetramer$^+$ cells in donor D3 were detected by all participants with a mean frequency of one cell in 1014 CD8$^+$ T-cells ± 355. In Donor D5, 11 of 12 laboratories detected a mean of one tetramer binding cell per 1106 CD8$^+$ T-cells ± 508. Influenza-specific cells were less numerous in healthy subjects D1 and D2 and were only detected by five and eight laboratories, respectively. No false positive reactivity was reported by any of the participants.

Eleven laboratories analyzed the five PBMC samples for the presence of HLA-A*0201-restricted CMV-specific and influenza-specific IFNγ-producing T-cells by ELISPOT assay. Only one group (Z10) used an intracellular cytokine staining as a functional test (data not shown because no comparison with other groups possible). Table 2 shows (I) the minimum, mean and maximum frequencies of antigen-specific cells, (II) the results obtained from the individual
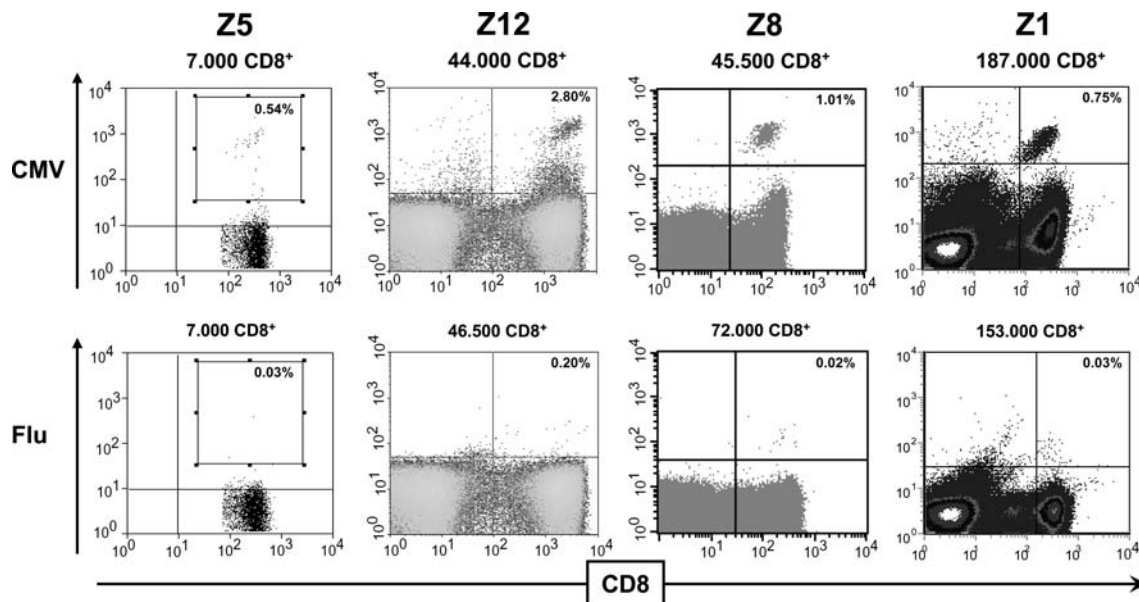
Fig. 1 Example of tetramer staining results as provided by four selected participating centers Z5, Z12, Z8 and Z1. All stainings were performed on donor D1 from phase I/2005 who showed reactivity with both of the tested tetramers. Cells were gated either on the lymphocyte population (Z1), or the subsets of CD3$^+$CD8$^+$ (Z5) or CD3$^+$ (Z8, Z12), according to the Ab combination used by each lab. The *upper panel* shows results for tests with the CMV-tetramer, the lower panel shows results for tests with the influenza-tetramer. In all dot-plots, the tetramer staining is displayed on the *y*-axis and anti-CD8-staining on the *x*-axis. Number of counted CD8$^+$ T-cells and percentage of tetramer-positive cells among the CD8 subset are indicated

Table 1 Overview of the tetramer results from phase I/2005 of the CIMT monitoring panel

	D1 CMV	D1 Flu	D2 Flu	D3 Flu	D5 CMV	D5 Flu	
Min	448	20,000	6,667	1,818	106	1,786	(I)
Mean	141	8,095	1,909	1,014	80	1,106	
Max	35	3,774	526	595	26	588	
Z1	129	3,774	635	870	106	758	(II)
Z2	112	–	–	1,818	70	1,770	
Z3	n.d	–	–	701	n.d	–	
Z4	97	–	–	854	106	1,342	
Z5	161	–	–	909	88	606	
Z6	139	6,896	1,316	1,000	91	651	
Z7	448	20,000	6,667	1,266	82	1,724	
Z8	100	5,128	545	595	87	877	
Z9	n.d	–	1,316	1,274	n.d	1,786	
Z10	94	–	2,597	1,360	86	1,439	
Z11	99	4,675	1,667	877	58	588	
Z12	35	–	526	645	26	625	
Detected by	10/10	5/12	8/12	12/12	10/10	11/12	(III)
Detected %	100	42	67	100	100	92	

(I) Minimum, mean and maximum frequencies of antigen-specific T-cells

(II) Results obtained from the individual centers Z1–Z12. All frequencies are indicated as 1 per *x* counted CD8$^+$ T-cells

(III) Number and percentage of centers which detected a given reactivity in donors D1, D2, D3 and D5

centers Z1–Z12, and (III) the number and percentage of centers that detected each reactivity. As described in the "Materials and methods", results of spot-forming cells per seeded PBMC were accepted as a positive reaction only when passing statistical testing and when the number of antigen-specific spots exceeded the number of spots in the background wells by at least a factor of two. IFNγ-producing cells reactive against CMV were detected by 10 of the 11 laboratories in donor D1 (mean reactivity was 1 per 1,855 PBMC $\pm$ 825) but only by 8 of 11 in donor D5 (mean reactivity was 1 per 4,405 PBMC $\pm$ 3,762). The influenza-specific T-cells present in subject D3 were detected by six laboratories, while the responses in the healthy subjects with markedly lower numbers of peripheral specific

Table 2 Overview of the IFNγ ELISPOT results from phase I/2005 of the CIMT monitoring panel

	D1 CMV	D1 Flu	D2 Flu	D3 Flu	D5 CMV	D5 Flu	
Min	3,061	62,500	55,555	33,333	11,428	50,000	(I)
Mean	1,855	38,141	43,589	17,547	4,405	30,811	
Max	888	10,256	30,769	8,571	1,039	14,705	
Z1	1,006	–	–	8,571	–	–	(II)
Z2	2,439	–	–	–	2,816	–	
Z3	1,295	62,500	–	22,727	1,412	27,727	
Z4	1,312	–	–	10,909	1,980	–	
Z5	–	–	–	–	–	–	
Z6	233*	–	–	1,960*	253*	–	
Z7	1,895	10,256	44,444	33,333	11,428	50,000	
Z8	3,061	–	–	–	6,896	–	
Z9	888	41,666	55,555	12,195	1,039	14,705	
Z10	ND	ND	ND	ND	ND	ND	
Z11	1,769	–	30,769	–	–	–	
Z12	3,030	–	–	–	5,263	–	
Detected by	10/11	3/11	3/11	6/11	8/11	3/11	(III)
Detected %	91	27	27	55	73	27	

(I) Minimum, mean and maximum frequencies of antigen-specific cells

(II) Results obtained from the individual centers Z1–Z12. All frequencies are indicated as 1 per x seeded PBMC except for Z6* where it is indicated as 1 per x seeded CD8+ T-cells

Results from Z6 were not included for calculation of the mean frequency of antigen-specific T-cells in D1, D3 and D5

(III) Number and percentage of centers that detected a given response in donors D1, D2, D3 and D5

ND not determined

T-cells (D1, D2 and D5) were detected by three laboratories only.

Subgroup analysis reveals that the number of CD8+ T-lymphocytes analyzed affects the sensitivity of the tetramer staining

Although the tetramer stainings were performed with centrally prepared reagents following set guidelines, centers were left free to select several parameters according to their own protocols, and this could have influenced the test results (see "Materials and methods"). Most of the participants used monoclonal antibodies specific for CD3 and CD8 to co-stain the cells. There were no obvious differences in the performance of the centers depending on which antibody clones, antibody combinations or cytometer were used (data not shown).

There was a high degree of variability in the number of CD8+ cells which were analyzed per staining, ranging from only 0.5×10^4 to about 19×10^4 (inter-center variation). In addition, a non-negligible intra-center variation was observed for the number of counted CD8+. We therefore analyzed each individual staining independently of the center that performed it and focused on the number of CD8+ T-cells that had been counted. For the

six different antigen-specific populations detectable, a total of 68 tests was performed by the group (see Table 1). Overall, antigen-specific T-cell reactivities were reported in 82% of the tests (56/68, mean of duplicate stainings). When less than 30,000 CD8+ T-cells were counted, only 70% of all responses were found. In contrast, 89% of all responses were manifest when more than 30,000 CD8+ T-cells were counted (Fig. 2a). When antigen-specific T-cells were present at high frequency, the number of cells counted did not influence the result, because CMV-specific T-cells from donors D1 and D5 were detected irrespective of the number of CD8+ T-cells in the test. However, for the influenza-specific cells, positivity was registered in only 75% of all tests performed (36 of 48 tests). Strikingly, we observed a marked difference for the results derived from those tests involving less than 30,000 CD8+ T-cells (56% success in detection) as compared to tests performed with more than 30,000 CD8+ T-cells (84%).

In conclusion, the ability to detect antigen-specific T-cell reactivities by tetramer staining was mainly affected by the number of CD8+ T-cells stained and analyzed, especially when the antigen-specific T-cells were present at low or moderate frequencies. We therefore modified our guidelines for the tetramer assay and recommended staining at
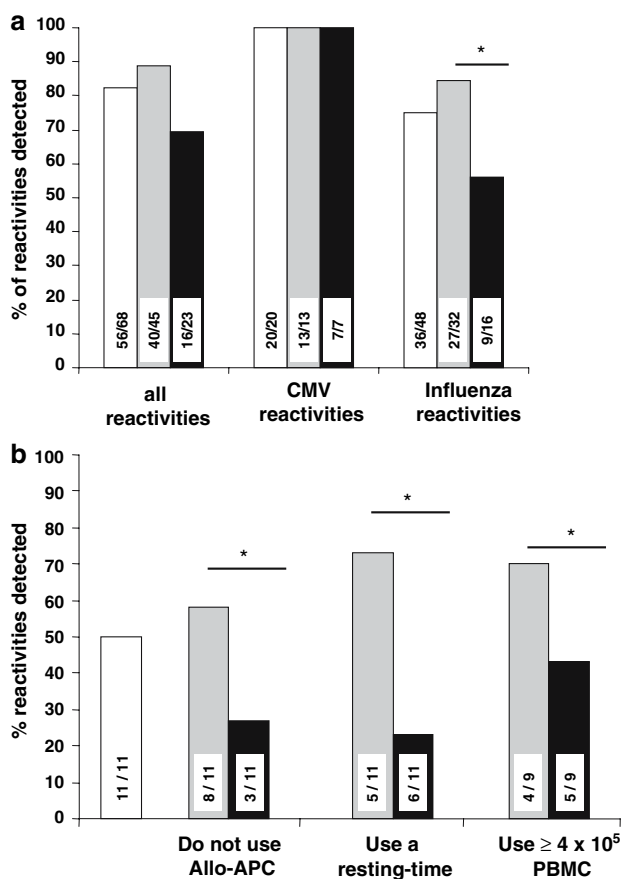
Fig. 2 **a** Subgroup analysis of tetramer results from phase I/2005. *Bars* indicate the percentage of positives that could be detected by tetramer staining. The first group of *bars* shows the results for all of the six detectable positives, the second group shows results from stainings with the CMV-tetramer and the third group of columns shows results from stainings with the influenza-tetramer. The *open bars* in each group represent all tests performed, *grey bars* represent results obtained in tests that were performed on more than 3×10^4 CD8$^+$ T-cells and *black bars* represent results obtained in tests that were performed on less than 3×10^4 CD8$^+$ T-cells. The *boxes* within each *bar* indicate the fraction of tests with a positive result. The *asterisk* indicates a *P*-value < 0.05 by Chi-square analysis. **b** Subgroup analysis of ELISPOT results from phase I/2005. The *bars* indicate the percentage of positive reactivities detected by IFNγ ELISPOT assays. The *open bar* shows the percentage of all reactivities detected by all 11 centers that performed the ELISPOT assay as the functional test. Criteria for division of centers into two subgroups were based on the following requirements: do not use allo-APC (first subgroup analysis), use a resting time (second subgroup analysis) or use equal or more than 400,000 PBMC per well (third subgroup analysis). *Grey bars* always represent centers that are in conformity with the indicated minimum requirement, *black bars* show results from centers that did not fulfil that requirement. The *boxes* within each column indicate the fraction of centers in each category. The *asterisks* indicate a *P*-value < 0.05 in Chi-square analysis

least 1×10^6 PBMC and analyzing all cells in the tube. In addition, we provided an example of how optimal cell gates and dot-plot quadrants could be selected.

ELISPOT assays are heterogeneous and require standardization

The ELISPOT analyses were performed according to 11 more or less different protocols. The most discernible differences that were observed in these protocols concerned (1) the different types of multi-screen plates, (2) the serum origin, (3) the use of duplicates, triplicates or quadruplicates, (4) the use of allogeneic APC, (5) the inclusion of a resting phase after thawing the PBMC, (6) the number of PBMC per well, (7) the type of antibodies used, (8) the type of spot-reader, and the (9) enzyme and substrate for staining of the spots. Each center also used a different plate protocol (distribution of the wells, number of replicates, control tests).

The influence of each of these parameters on the number of positive responses was studied by further analysis in which the laboratories were divided into two subgroups. As a result, several criteria were identified which could help to improve the sensitivity and comparability of detection.

All data sets (duplicates, triplicates or quadruplicates) were first analyzed by Student *t* test for unpaired samples ("Materials and methods"). In our panel, one center used quadruplicates, nine centers used triplicates and one center performed the ELISPOT analysis in duplicates. Due to the variety in the replicates, responses measured by duplicate wells failed to pass the Student *t* test more often as compared to triplicates.

Overall, the 11 centers were able to detect 50% of all possible reactivities in this panel phase (Table 2; Fig. 2b). In a subgroup of three laboratories (Z5, Z6 and Z8), an allogeneic APC population (T2 or K562-A*0201 cells) was added for binding and presentation of the synthetic peptides. The three centers that used allo-APC detected only 28% of all responses, while the other centers detected 58% of all responses.

In five laboratories (Z3, Z4, Z7, Z8, and Z9) PBMC were thawed, and then incubated in culture medium at 37°C. After this resting phase of 2–20 h, living cells were washed, counted and seeded into ELISPOT plates. Laboratories using a resting phase detected 73% of the positive reactivities (22 out of 30 potentially positive tests). No significant difference in the ability to detect antigen-specific T cells was found using shorter or longer resting-times. In contrast, the laboratories that did not use a resting procedure detected only 30% of all positives (Fig. 2b).

Finally, the number of cells seeded per well differed considerably between all participants and ranged from 1 to 6×10^5 PBMC. We divided the laboratories arbitrarily into two groups, those using either more than 4×10^5 PBMC (Z4, Z7, Z8, and Z9) or less than 4×10^5 PBMC (Z1, Z2, Z3, Z11 and Z12). The first group detected 71% of all positive samples, whereas the second group was able to detect

only 43% of all positives (Fig. 2b). Centers Z5 and Z6 used a defined number of separated CD8$^+$ T-cells in the ELISPOT and were therefore not included in this subgroup analysis.

None other of the nine depicted protocol variables had any obvious impact on the detection of specific T-cells. As a conclusion from these results, four minimum requirements were formulated for the ELISPOT protocol: (1) perform triplicates for each test antigen (2) do not use allo-APC (3) add a resting time to increase the proportion of living cells seeded and (4) use a minimum number of 4×10^5 PBMC per well.

Phase II/2006 of the interlaboratory testing project—general aspects

To formally prove that the requirements formulated for tetramer staining and ELISPOT analysis increase the ability of the participants to detect antigen-specific CD8$^+$ T-cells and reduce the inter-center variability, we decided to repeat the analysis in a second phase of the panel, with the same participants (phase II/2006). In this round, all groups were asked to follow our modified guidelines for the tetramer- and the ELISPOT-assays.

Again, all PBMC samples were prepared and pre-tested in one central lab and peptide antigens and PE-conjugated tetramers were also provided from one source. As one investigator had meanwhile moved to another lab, we added a 13th center to the group. PBMC from seven selected healthy HLA-A*0201-positive donors and 1 HLA-A*0201-negative donor (D3) were required to be analyzed

for the presence of HLA-A*0201-restricted CMV-specific T cells and for influenza-specific T cells. The mean number of recovered cells after thawing was sufficient to perform the tests. When all the data were combined, it became clear that subjects D2, D5 and D8 possessed CMV-specific CD8$^+$ T-cell subsets, and D1, D2, D4, D6 and D7 possessed influenza-specific CD8$^+$ T-cells. Therefore, each laboratory could theoretically have measured eight positives ($3\times$ CMV and $5\times$ Influenza) in this second phase.

Analysis of CD8$^+$ T-cell tetramer binding using the new guidelines

In the second phase, a total of 104 tests were performed to detect the eight possible tetramer reactivities. Following the modified guidelines for tetramer staining, the mean number of CD8$^+$ T-cells that were counted in each separate test increased markedly (+36%): a mean of about 49,000 CD8$^+$ cells were analyzed in the phase I ($n = 68$ tests) and a mean of 67,000 CD8$^+$ T-cells in phase II ($n = 104$ tests). The number of cells per test ranged from 12,000 to 467,000 CD8$^+$. In 81% (84 of 104) of the tests >30,000 CD8$^+$ were counted (compared to 66% of all relevant tests in the first phase). Table 3 shows (I) the minimum, mean and maximum frequencies of antigen-specific T-cells, (II) the results obtained from the individual centers Z1–Z13, and (III) the number and percentage of centers that detected each T-cell specificity. Donors D2, D5 and D8 showed very strong reactivities with the CMV-tetramer, with mean frequencies of 1/45 CD8$^+$ T-cells, 1/37 CD8$^+$ T-cells, and 1/19 CD8$^+$ T-cells, respectively. All 13 laboratories were able to detect

Table 3 Overview of tetramer results from phase II/2006 of the CIMT monitoring panel

	D1 Flu	D2 CMV	D2 Flu	D4 Flu	D5 CMV	D6 Flu	D7 Flu	D8 CMV	
Min	7,143	77	10,000	10,000	60	3,333	869	42	(I)
Medium	3,739	45	3,573	5,278	37	1,116	347	19	
Max	1,538	28	1,250	2,500	30	571	202	8	
Z1	3,333	45	3,333	5,000	39	769	294	10	(II)
Z2	4,000	30	3,333	5,000	30	1,100	250	21	
Z3	4,000	47	–	5,000	45	588	270	20	
Z4	1,538	71	2,857	6,666	33	769	263	42	
Z5	6,666	54	2,500	–	43	1,428	377	24	
Z6	5,000	77	–	10,000	60	1,666	869	27	
Z7	2,857	47	3,333	6,666	38	833	290	22	
Z8	2,000	35	3,333	3,333	31	666	244	8	
Z9	3,333	31	10,000	3,333	32	3,333	625	20	
Z10	3,333	28	2,222	–	30	952	202	13	
Z11	1,666	36	1,250	2,500	32	571	215	15	
Z12	7,413	37	–	–	35	714	260	8	
Z13	–	53	–	–	34	–	–	20	
Detected by	12/13	13/13	9/13	9/13	13/13	12/13	12/13	13/13	(III)
Detected %	92	100	69	69	100	92	92	100	

(I) Minimum, mean and maximum frequencies of antigen-specific T-cells

(II) Results obtained from the individual centers Z1–Z13. All frequencies are indicated as 1 per x counted CD8$^+$ T-cells

(III) Number and percentage of centers that detected each of the eight possible responses

these populations (Table 3). All but one center detected the influenza-specific cells present at high frequencies in donors D6 (1/1116 CD8[+] T-cells) and D7 (1/347 CD8[+] T-cells). Donors D1, D2 and D4 possessed fewer specific cells (1/3,739, 1/3,573 and 1/5,278 CD8[+] T-cells) which were found by 12, 9 and 9 centers, respectively. Three laboratories also reported influenza tetramer-binding CD8[+] cells in D5 or D8. According to the results of the other centers as well as from the ELISPOT (see below), these stainings were considered as false positive (not shown). One center (Z13) was not able to detect any of the influenza-specific CD8[+] T-cell reactivities. Finally, no tetramer[+] cells were described in the HLA-A*0201-negative donor (D3).

Analysis of CD8[+] T-cell responses by ELISPOT following the introduction of a set of four rules

In this second phase, all laboratories performed ELISPOT analysis following local protocols, all of which conformed to the newly introduced minimum requirements. Table 4 shows (I) the minimum, mean and maximum frequencies of antigen-specific cells, (II) the results obtained from the individual centers Z1–Z13, and (III) the number and percentage of centers that detected the response. High frequency T-cell responses against CMV could readily be detected by all 13 centers in donors D5 and D8 and by 12 of 13 in donor D2. Failure of center Z4 to detect the CMV reactivity in donor D2 was due to a very high background of the medium control. The number of spots representing IFNγ-producing cells after influenza-peptide stimulation was generally lower, and consequently, the influenza-

specific T-cell responses in subjects D1, D2, D4 and D6 were detected by fewer laboratories (four centers for D1, three centers for D2, two centers for D4 and ten centers for D6). The high numbers of influenza-specific T-cells present in D7 were detected by all 13 laboratories (Table 4).

Comparison of the results obtained in both phases

When the mean frequencies of all T-cell responses in both testing rounds were compared, it became clear that there was a difference in the distribution of reactivities (Fig. 3). In the tetramer assay, the mean T-cell frequency of the six possible positives in the first phase was 1 per 2,083 CD8[+] T-cells. This value was 1 per 1,769 CD8[+] T-cells for the eight possible positives in the second phase. Similarly, the mean T-cell frequency of the responses detected in IFNγ ELISPOT was 1 per 22,369 PBMC for Phase I/2005 but 1 per 14,653 PBMC for Phase II/2006. To allow a comparison of the overall performance in both phases of the panel, we therefore decided to define theoretical thresholds for high, moderate and low T-cell responses and then to compare data of the participating laboratories within these groups.

In order to define such thresholds for low, medium and high T-cell responses, we first displayed the probability of detecting each of the 14 different reactivities as a value in a coordinate system and inserted a trendline. For both the tetramer assay and the ELISPOT assay, we observed a clear correlation between the frequencies of antigen-specific T-cells and the number of participating centers that were able to detect these populations. We then calculated the

Table 4 Overview of IFNγ ELISPOT results from phase II/2006 of the CIMT monitoring panel

	D1 Flu	D2 CMV	D2 Flu	D4 Flu	D5 CMV	D6 Flu	D7 Flu	D8 CMV	
Min	44,118	1,791	33,803	58,824	1,745	48,387	14,720	1,698	(I)
Medium	28,823	1,088	16,395	49,960	999	14,265	4,669	1,023	
Max	10,345	396	1,231	41,096	391	4,458	1,706	269	
Z1	44,118	596	–	–	596	48,387	14,720	318	(II)
Z2	–	1,732	–	–	1,745	4,739	2,222	1,698	
Z3	–	1,333	–	–	1,117	–	2,982	1,292	
Z4	–	–	–	–	774	–	2,273	447	
Z5	–	997	–	–	1,157	17,647	3,827	1,209	
Z6	10,345	1,031	14,151	–	1,006	–	1,706	1,005	
Z7	–	1,317	–	–	1,061	10,000	6,000	1,661	
Z8	32,258	396	–	58,824	391	4,458	3,623	269	
Z9	28,571	966	–	41,096	912	14,423	8,955	1,081	
Z10	–	847	1,231	–	806	7,426	3,052	696	
Z11	–	1,044	33,803	–	1,087	20,339	3,670	1,273	
Z12	–	1,791	–	–	1,387	10,619	3,057	1,583	
Z13	–	1,008	–	–	949	4,615	4,615	770	
Detected by	4/13	12/13	3/13	2/13	13/13	10/13	13/13	13/13	(III)
Detected %	31	92	23	15	100	77	100	100	

(I) Minimum, mean and maximum frequencies of antigen-specific cells

(II) Results obtained from the individual centers Z1–Z13. All frequencies are indicated as 1 per *x* seeded PBMC

(III) Number and percentage of centers that detected each of the eight possible responses

Fig. 3 Distribution of antigen-specific T-cell frequencies in the two testing phases as obtained by tetramer staining (**a**) and IFNγ ELISPOT assays (**b**). The figure shows the six reactivities (*filled circle*) and the calculated mean of all reactivities from phase I/2005 (*filled line*) as well as the eight reactivities (*open circle*) and calculated mean of all reactivities from phase II/2006 (*open line*). The frequency of antigen-specific T-cells is indicated on the *y*-axis as 1 per *x* counted CD8⁺ T-cells for the tetramer test and as 1 per *x* seeded PBMC for the ELISPOT assay
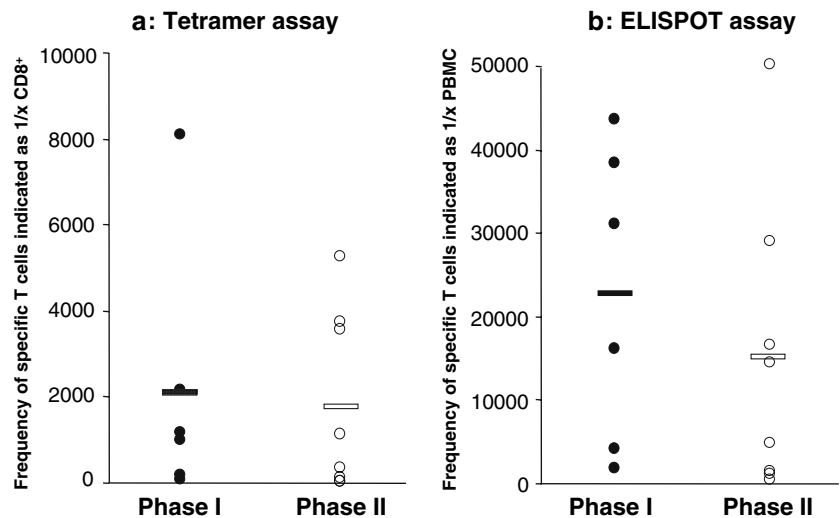
theoretical frequencies at which 90% ($y = 90$) and 50% ($y = 50$) of all participants could detect a given response (Fig. 4a, b) and used these two thresholds to divide all reactivities into three distinct classes of T-cell responses ("high", "moderate" and "low").

For the tetramer assay, T-cell frequencies exceeding 1 per 1,200 CD8⁺ T-cells were therefore classified as "high", whereas frequencies of less than 1 per 7,650 CD8⁺ were classified as "low" (Fig. 4a). Following the same rules for the ELISPOT assay, T-cell responses of at least one IFNγ spot per 2,850 PBMC can be considered as "high" and T-cell responses of less than one spot per 19,000 PBMC as "low" (Fig. 4b).

With these calculated assay-specific thresholds for high, moderate and low T-cell responses, we compared the results obtained in the two phases. For the tetramer assay, the ability to detect high frequency T-cells (>1 per 1,200 CD8⁺) did not differ in the two phases, and was not influenced by the number of CD8⁺ analyzed, as previously seen for each of the two phases separately (Fig. 5a). However, for moderate and low T-cell frequencies, we found that they could be successfully detected in only 54% of cases in the first phase but this improved to 77% in the second phase. Moreover, here, the number of cells counted did have an impact on the ability to detect low frequency T-cells. In the first phase, only 14% were detected when less than 30,000 CD8⁺ were counted, as compared to 71% when more than 30,000 CD8⁺ T-cells were counted. The same trend was observed in phase II/2006, but in this case 40% of assays with less than 30,000 CD8⁺ successfully detected the moderate to low T-cell frequencies compared to 83% counting more than 30,000 CD8⁺ (Fig. 5a).

We then analyzed the capacity of the laboratories to measure either high T-cell responses (>1 per 2,850 PBMC) or low to moderate T-cell responses (<1 per 2,850 PBMC) in the ELISPOT assay. This analysis was performed for
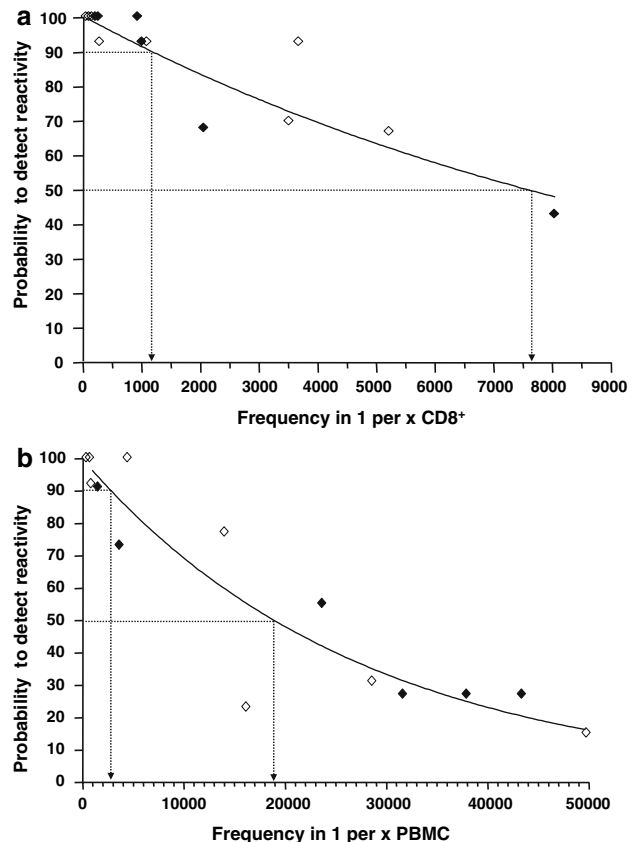

Fig. 4 Probability of detecting a reactivity by **a** tetramer staining, or **b** IFNγ ELISPOT assay. A trendline was inserted on the basis of results from all 14 reactivities from both phases of the panel. The figure shows the six reactivities from phase I/2005 (*filled squares*) and the eight reactivities from phase II/2006 (*open squares*). The frequency of antigen-specific T-cells is shown on the *x*-axis in 1 per *x* counted CD8⁺ T-cells for the tetramer assay (**a**) or 1 per *x* seeded PBMC for the ELISPOT assay (**b**). *X*-values for $y = 90\%$ and $y = 50\%$ are indicated by the *broken lines*
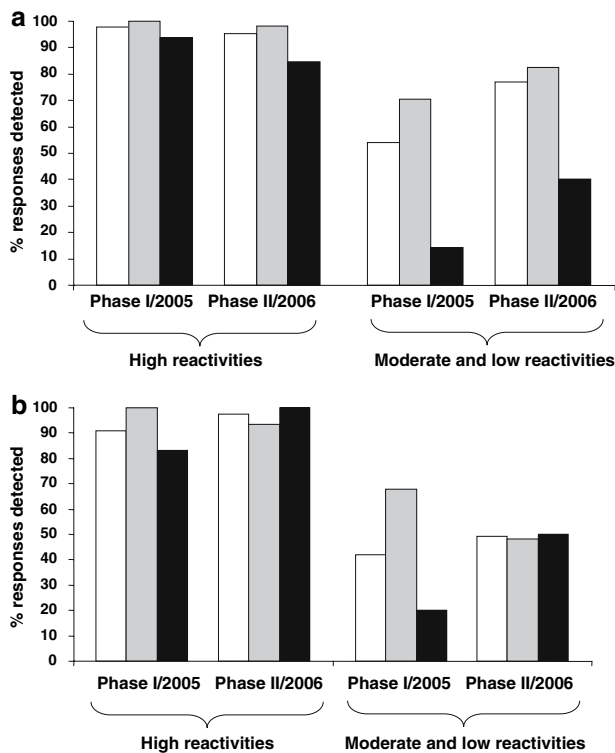
Fig. 5 **a** Percentage of reactivities actually detected by tetramer staining. The first two groups of *bars* show the detection rate for the nine high reactivities (>1 per 1,200 CD8$^+$ T-cells) in phase I/2005 and phase II/2006. The next two groups of *bars* show the detection rate for five moderate to low reactivities (<1 per 1,200 CD8$^+$) in phase I/2005 (third group) or phase II/2006 (fourth group). The *open bars* represent all tests performed, *grey bars* represent results obtained in tests that were performed on more than 3 × 10^4 CD8$^+$ T-cells and *filled bars* represent results obtained in tests that were performed on less than 3 × 10^4 CD8$^+$ T-cells. **b** Percentage of reactivities detected in IFNγ ELISPOT assays. The first two groups of *bars* show the rate of detection of the four high reactivities (>1 per 2,850 PBMC in phase I/2005 and phase II/2006. The next two groups of *columns* show the rate of detection for the ten moderate to low reactivities (<1 per 2,850 PBMC) in phase I/2005 and phase II/2006. The *open bars* represent the performances of all centers in the respective panel phase, *grey bars* represent results obtained from the five centers that already fulfilled at least three of the four minimum criteria in phase I/2005 and *filled bars* represent results obtained from centers that fulfilled less than three of the four minimum criteria in phase I/2005

two defined subgroups of participants. The first subgroup included those five centers (Z3, Z4, Z7, Z8 and Z9) that already fulfilled three or four of the requirements in the first phase of the panel. These five centers did not have to introduce any change or at least no major changes to their protocol for the repetition of the experiments in phase II. The second subgroup included the new center Z13 (led by a colleague that had been in a laboratory that only fulfilled one of four requirements in phase I) and all others that had fulfilled only one or two of the four requirements in the first phase. All laboratories in this second group had to introduce marked changes to their locally established protocols.

Similar to the tetramer analysis, the new requirements were not necessary to detect antigen-specific responses among the category of high T-cell frequencies in either the first or second phases (Fig. 5b). However, applying the set of rules defined in phase I markedly improved the capacity of centers to detect the low to moderate T-cell responses. The first subgroup detected a total of 68% of the low to moderate reactivities in phase I, whereas the second subgroup detected only 20% (Fig. 5b). After harmonization of the protocols, both subgroups performed equally well. In addition, the inter-group variability in detecting positive responses was reduced in phase II (percentage of detected responses ranged from 38 to 88% with a mean of $67 \pm 16\%$) as compared to phase I (percentage of detected responses ranged from 0 to 100% with a mean of $55 \pm 33\%$).

Experience does not equal performance

Among the 13 centers that had participated in phase II, tetramer stainings had been performed for 1–8 years. Similarly, the experience in the ELISPOT technology varied between 1 and 10 years. For both techniques, we could not find any correlation between the years of experience and the ability to detect T-cell responses, not even among the subgroups of moderate or low T-cell responses (not shown).

Discussion

Whenever new techniques are introduced to the scientific community, they are first only available to a small group of expert laboratories. If these assays are robust and applicable for specific research or routine applications, they spread to the international community. In general, the "original" protocol then undergoes several adaptations in order to meet specific needs. On the one hand, changes can be beneficial and result in the improvement of protocols. On the other, this evolutionary process leads to employment of many different protocol variants, limiting comparison of the study results obtained by different laboratories. Thus, standardization approaches should be omitted during the initial development but are absolutely required when assays have become firmly established. In recent years, several activities aiming at the harmonization of techniques used to monitor the presence of antigen-specific T-cells have been initiated for ELISPOT [28–31], tetramer staining [32] and ICS [33–36]. While these studies showed the feasibility, general applicability and the diversity of performance among participants, they were not designed to either systematically investigate the influence of distinct protocol variables nor to test whether changes to these parameters can lead to a global improvement of the group. The CIMT

monitoring panel is the first initiative that has now introduced the two-step approach proposing a strategy where technical variables that influence the performance of a defined assay are first systematically identified ("first step") followed by a new testing phase where resultant protocol changes are validated under controlled conditions within the same group of investigators ("second step").

As soon as a number of protocol variables that might have influenced the sensitivity and the quality of the tests were identified in the first phase of our study, it was decided to validate this finding in a second phase. Because this two-step approach was not initially foreseen, the second phase was performed with PBMC samples obtained from different donors than those used in the first round. The distribution and the frequencies of detectable T-cell responses directed against the chosen model antigens were different in the first and second group of donors (Fig. 3) precluding a direct comparison of the results obtained in both phases of the panel. To circumvent this problem, two assay-specific frequency thresholds were introduced that allowed us to distinguish classes of T-cell responses (low, moderate and high) (Fig. 4a, b). Clearly, high-frequency T-cell responses were detected irrespective of the protocol used and as such did not allow the identification of factors that exert a strong influence on the sensitivity and variability of the protocols used. Relevant parameters could only be detected when the comparison was focused on the detection of T-cells that are present at low to moderate frequencies in PBMC. This finding should be taken into account when selecting model antigens for use in monitoring panels [37], in particular by laboratories that are interested in the detection of peripheral tumor-specific T-cells, which are often present at low frequencies, even after vaccination.

Although our experiments do not specifically address the question of detection limits for the ELISPOT and tetramer assays, we could detect a high variability in the sensitivity of protocols used by the different participants. The majority of labs ($y = 90\%$) is able to detect responses with a frequency above 1 per 1,200 CD8$^+$ T cells in the tetramer assay or responses with a frequency above 1 per 2,859 PBMC in the ELISPOT. Note that some of the centers could reliably detect a response with a frequency of about 1 per 8,000 CD8$^+$ T cells in the tetramer assay and about 1 per 40,000 PBMC in the ELISPOT assay. These low frequencies are in the range of that is commonly reported as the detection limit for internally validated protocols for both technologies [39, 40, own unpublished observations]. Another important task of standardization efforts should be to decrease the variation of results obtained in a group of several laboratories down to the stable and low values (15–30%) that can be reproducibly found within single labs. In order to quantify the variation of results among laboratories we calculated the coefficient of variation for all

14 reactivities of the two panel phases. The CVs were determined on the base of centers that were able to detect the respective T cell response and the results are shown in supplementary Tables S1a, b. As expected, the CVs we found in our inter-laboratory testing project were higher than those reported from intra-center analysis [39, 40].

In the ELISPOT assay, the background spot numbers obtained by the different participants varied greatly, but we were unable to correlate this finding to a distinct variable. Since the spontaneous cytokine secretion impacts significantly on the sensitivity of this assay, factors that especially influence the non-specific spot production, possibly the medium type or serum source, will need to be systematically analyzed in a separate study.

The main conclusions from our study have been drawn on the basis of subgroup analyses. Although the CIMT panel in general (13 centers in this initial action), and consequently the subgroups formed during the analysis were rather small, we could already identify statistically significant differences in the ability to detect positive responses. We concluded that the number of counted CD8$^+$ T-cells is the most influential crucial factor for the tetramer assay and that the combination of a resting-time and a high number of PBMC leads to increased sensitivity in the ELISPOT assay. This suggests that the impact of the identified technical variables on the quality of the assays is high. In order to identify those protocol variables that lead to more subtle differences, a larger group of participants would be needed.

In addition to the systematic identification of variables that correlate with sensitivity/insensitivity of various assays, inter-laboratory testing projects also allow the rapid evaluation of individual performance among a group. Interestingly, the finding that experienced laboratories did not perform better than laboratories which recently applied these techniques strongly suggests that non-optimal protocols, once established in a lab, can commonly be maintained for several years. Periodic comparison of local protocols with those of other centers is recommended. Even if a new staff member uses an established protocol, it is recommended to have them participate in inter-laboratory testing/teaching exercise. Regular participation in multi-center comparisons could thereby help to optimize and validate participants' performance over time and to maintain sensitive protocols or minimal standards. This is of great importance when material from expensive clinical trials has to be analyzed.

All data from the CMV-serology, from the pre-testing experiments and from the results generated by the participating laboratories in ELISPOT and tetramer staining were taken together for each donor in order to qualitatively validate the presence of CMV- and influenza-specific T-cells. To estimate the quantity, i.e. the frequency of specific T-cells in each donor, we calculated the average of all qualitatively positive results, as well as the standard deviations.

This procedure constitutes only an approximation of the real number of antigen-specific cells present in a given sample, and cannot be taken as a method for determining absolute T-cell frequencies. Cell samples that contain pre-defined numbers of antigen-specific T-cells (e.g. spiked T-cell clones), especially tumor-reactive T-cells, are not easily available for use in multi-center comparisons, although such standard samples are urgently needed. We see this as one major bottle-neck for the optimization and standardization of immunomonitoring techniques. Methods to generate such standard samples for broader use will therefore be elucidated with high priority in the near future for the next phases of this international collaboration. Another big challenge will be to define accepted rules for the settings of the equipment used in these analyses (flow cytometer or ELISPOT reader) in order to uniformly process and analyze the raw data. Ten from eleven laboratories that performed the ELISPOT assay in the first phase used an ELISPOT reader for spot counting. It is known that spot counts between centers can differ significantly and this may be explained by the use of different reading machines, different settings for the same type of machine or by the experience of the operator. Within this group, four different commercially available reading systems were used (supplementary Table S2). We were not able to identify differences between the types of ELISPOT readers. A new ELISPOT panel phase is currently in preparation, that will specifically focus on the performances of different ELISPOT readers and try to introduce tools to control inter center variation. In addition, none of the participant reported on the use of live/dead cell discrimination on thawed PBMC samples for the FACS-based experiments. Whether the combination of staining with Ab/HLA-tetramers and vital dyes or with a resting phase is beneficial for increasing the sensitivity of the tetramer staining assay could be addressed in future testing actions.

Results from a proficiency panel of 36 laboratories from nine different countries in which the ELISPOT assay was validated are now also being reported [41]. This initiative, conducted under the aegis of the Cancer Vaccine Consortium (CVC), was mainly designed to offer an external validation to the participating laboratories but the in depth analysis of the obtained data sets lead to similar findings and recommendations as the CIMT monitoring panel. It confirmed that a resting phase of cells prior to addition to the ELISPOT plates is advantageous and should therefore be generally recommended. Furthermore, a long year experience in a technology did not guarantee for a sensitive test and failure to detect specific T cell responses concentrated on the weak responses. The fact that two independent initiatives come to similar findings is surely notable and shows the necessity to carry on running proficiency panels.

Last but not least, we would like to stress that even the best guidelines and protocols alone cannot guarantee good performance. Monitoring of antigen-specific T-cell responses requires skills as well as experience. Participation in immunomonitoring panels cannot compensate for the need to constantly educate and train staff and to develop specific expertise for covering individual needs. Nevertheless, we strongly believe that by organizing further two-step interlaboratory testing projects, the CIMT monitoring panel will be able to improve the sensitivity of the assays used for immunomonitoring as well as to actively participate in the harmonization of these assays, which is required to enable the comparison of immunotherapeutic trials performed in different centers.

References

1. Mocellin S, Lise M, Nitti D (2005) Targeted therapy for colorectal cancer: mapping the way. Trends Mol Med 11:327
2. Mocellin S, Mandruzzato S, Bronte V, Marincola FM (2004) Cancer vaccines: pessimism in check. Nat Med 10:1278
3. Restifo NP, Rosenberg SA (2005) Use of standard criteria for assessment of cancer vaccines. Lancet Oncol 6:3
4. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, Tosolini M, Camus M, Berger A, Wind P, Zinzindohoue F, Bruneval P, Cugnenc PH, Trajanoski Z, Fridman WH, Pages F (2006) Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science 313:1960
5. Romero P, Cerottini JC, Speiser DE (2004) Monitoring tumor antigen specific T-cell responses in cancer patients and phase I clinical trials of peptide-based vaccination. Cancer Immunol Immunother 53:249
6. Dunn GP, Old LJ, Schreiber RD (2004) The immunobiology of cancer immunosurveillance and immunoediting. Immunity 21:137
7. Boon T, Coulie PG, Van den Eynde BJ, van der Bruggen P (2006) Human T-cell responses against melanoma. Annu Rev Immunol 24:175
8. de Jong A, van Poelgeest MI, van der Hulst JM, Drijfhout JW, Fleuren GJ, Melief CJ, Kenter G, Offringa R, van der Burg SH (2004) Human papillomavirus type 16-positive cervical cancer is associated with impaired CD4+ T-cell immunity against early antigens E2 and E6. Cancer Res 64:5449
9. van Poelgeest MI, van Manon S, van Marc B, Kwappenberg KM, Heijmans-Antonissen C, Drijfhout JW, Melief CJ, Kenter GG, Helmerhorst TJ, Offringa R, van der Burg SH (2005) Detection of human papillomavirus (HPV) 16-specific CD4+ T-cell immunity in patients with persistent HPV16-induced vulvar intraepithelial neoplasia in relation to clinical impact of imiquimod treatment. Clin Cancer Res 11:5273
10. Czerkinsky C, Andersson G, Ekre HP, Nilsson LA, Klareskog L, Ouchterlony O (1988) Reverse ELISPOT assay for clonal analysis of cytokine production. I. Enumeration of gamma-interferon-secreting cells. J Immunol Methods 110:29
11. Herr W, Schneider J, Lohse AW, Meyer zum Buschenfelde KH, Wölfel T (1996) Detection and quantification of blood-derived

CD8+ T lymphocytes secreting tumor necrosis factor alpha in response to HLA-A2.1-binding melanoma and viral peptide antigens. J Immunol Methods 191:131

12. Altman JD, Moss PA, Goulder PJ, Barouch DH, Heyzer-Williams MG, Bell JI, McMichael AJ, Davis MM (1996) Phenotypic analysis of antigen-specific T lymphocytes. Science 274:94

13. Jung T, Schauer U, Heusser C, Neumann C, Rieger C (1993) Detection of intracellular cytokines by flow cytometry. J Immunol Methods 159:197

14. Assenmacher M, Schmitz J, Radbruch A (1994) Flow cytometric determination of cytokines in activated murine T helper lymphocytes: expression of interleukin-10 in interferon-gamma and in interleukin-4-expressing cells. Eur J Immunol 24:1097

15. Stone JD, Demkowicz WE Jr, Stern LJ (2005) HLA-restricted epitope identification and detection of functional T-cell responses by using MHC-peptide and costimulatory microarrays. Proc Natl Acad Sci USA 102:3744

16. Betts MR, Brenchley JM, Price DA, De Rosa SC, Douek DC, Roederer M, Koup RA (2003) Sensitive and viable identification of antigen-specific CD8+ T-cells by a flow cytometric assay for degranulation. J Immunol Methods 281:65

17. Chattopadhyay PK, Yu J, Roederer M (2005) A live-cell assay to detect antigen-specific CD4+ T-cells with diverse cytokine profiles. Nat Med 11:1113

18. Frentsch M, Arbach O, Kirchhoff D, Moewes B, Worm M, Rothe M, Scheffold A, Thiel A (2005) Direct access to CD4+ T-cells specific for defined antigens according to CD154 expression. Nat Med 11:1118

19. Scriba TJ, Purbhoo M, Day CL, Robinson N, Fidler S, Fox J, Weber JN, Klenerman P, Sewell AK, Phillips RE (2005) Ultrasensitive detection and phenotyping of CD4+ T-cells with optimized HLA class II tetramer staining. J Immunol 175:6334

20. Chattopadhyay PK, Price DA, Harper TF, Betts MR, Yu J, Gostick E, Perfetto SP, Goepfert P, Koup RA, De Rosa SC, Bruchez MP, Roederer M (2006) Quantum dot semiconductor nanocrystals for immunophenotyping by polychromatic flow cytometry. Nat Med 12:972

21. Speiser DE, Baumgaertner P, Barbey C, Rubio-Godoy V, Moulin A, Corthesy P, Devevre E, Dietrich PY, Rimoldi D, Lienard D, Cerottini JC, Romero P, Rufer N (2006) A novel approach to characterize clonality and differentiation of human melanoma-specific T-cell responses: spontaneous priming and efficient boosting by vaccination. J Immunol 177:1338

22. Devevre E, Romero P, Mahnke YD (2006) Livecount assay: concomitant measurement of cytolytic activity and phenotypic characterisation of CD8(+) T-cells by flow cytometry. J Immunol Methods 311:31

23. Batard P, Peterson DA, Devevre E, Guillaume P, Cerottini JC, Rimoldi D, Speiser DE, Winther L, Romero P (2006) Dextramers: new generation of fluorescent MHC class I/peptide multimers for visualization of antigen-specific CD8+ T-cells. J Immunol Methods 310:136

24. Offringa R, van der Burg SH, Ossendorp F, Toes RE, Melief CJ (2000) Design and evaluation of antigen-specific vaccination strategies against cancer. Curr Opin Immunol 12:576

25. Coulie PG, Karanikas V, Colau D, Lurquin C, Landry C, Marchand M, Dorval T, Brichard V, Boon T (2001) A monoclonal cytolytic T-lymphocyte response observed in a melanoma patient vaccinated with a tumor-specific antigenic peptide encoded by gene MAGE-3. Proc Natl Acad Sci USA 98:10290

26. Keilholz U, Weber J, Finke JH, Gabrilovich DI, Kast WM, Disis ML, Kirkwood JM, Scheibenbogen C, Schlom J, Maino VC, Lyerly HK, Lee PP, Storkus W, Marincola F, Worobec A, Atkins MB (2002) Immunologic monitoring of cancer vaccine therapy: results of a workshop sponsored by the Society for Biological Therapy. J Immunother 25:97

27. Keilholz U, Martus P, Scheibenbogen C (2006) Immune monitoring of T-cell responses in cancer vaccine development. Clin Cancer Res 12:2346S

28. Scheibenbogen C, Romero P, Rivoltini L, Herr W, Schmittel A, Cerottini JC, Wölfel T, Eggermont AM, Keilholz U (2000) Quantitation of antigen-reactive T-cells in peripheral blood by IFNgamma-ELISPOT assay and chromium-release assay: a four-centre comparative trial. J Immunol Methods 244:81

29. Cox JH, Ferrari G, Janetzki S (2006) Measurement of cytokine release at the single cell level using the ELISPOT assay. Methods 38:274

30. Janetzki S, Cox JH, Oden N, Ferrari G (2005) Standardization and validation issues of the ELISPOT assay. Methods Mol Biol 302:51

31. Janetzki S, Schaed S, Blachere NE, Ben-Porat L, Houghton AN, Panageas KS (2004) Evaluation of Elispot assays: influence of method and operator on variability of results. J Immunol Methods 291:175

32. Heijnen IA, Barnett D, Arroz MJ, Barry SM, Bonneville M, Brando B, D'hautcourt JL, Kern F, Totterman TH, Marijt EW, Bossy D, Preijers FW, Rothe G, Gratama JW (2004) Enumeration of antigen-specific CD8+ T lymphocytes by single-platform, HLA tetramer-based flow cytometry: a European multicenter evaluation. Cytometry B Clin Cytom 62:1

33. Maecker HT, Rinfret A, D'Souza P, Darden J, Roig E, Landry C, Hayes P, Birungi J, Anzala O, Garcia M, Harari A, Frank I, Baydo R, Baker M, Holbrook J, Ottinger J, Lamoreaux L, Epling CL, Sinclair E, Suni MA, Punt K, Calarota S, El-Bahi S, Alter G, Maila H, Kuta E, Cox J, Gray C, Altfeld M, Nougarede N, Boyer J, Tussey L, Tobery T, Bredt B, Roederer M, Koup R, Maino VC, Weinhold K, Pantaleo G, Gilmour J, Horton H, Sekaly RP (2005) Standardization of cytokine flow cytometry assays. BMC Immunol 6:13

34. Maecker HT, Moon J, Bhatia S, Ghanekar SA, Maino VC, Payne JK, Kuus-Reichel K, Chang JC, Summers A, Clay TM, Morse MA, Lyerly HK, DeLaRosa C, Ankerst DP, Disis ML (2005) Impact of cryopreservation on tetramer, cytokine flow cytometry, and ELISPOT. BMC Immunol 6:17

35. Maecker HT, Trotter J (2006) Flow cytometry controls, instrument setup, and the determination of positivity. Cytometry A 69:1037

36. Hobeika AC, Morse MA, Osada T, Ghanayem M, Niedzwiecki D, Barrier R, Lyerly HK, Clay TM (2005) Enumerating antigen-specific T-cell responses in peripheral blood: a comparison of peptide MHC tetramer, ELISpot, and intracellular cytokine analysis. J Immunother 28:63

37. Currier JR, Kuta EG, Turk E, Earhart LB, Loomis-Price L, Janetzki S, Ferrari G, Birx DL, Cox JH (2002) A panel of MHC class I restricted viral peptides for use as a quality control for vaccine trial ELISPOT assays. J Immunol Methods 260:157

38. Britten CM, Meyer RG, Kreer T, Drexler I, Wölfel T, Herr W (2002) The use of HLA-A*0201-transfected K562 as standard antigen-presenting cells for CD8(+) T lymphocytes in IFN-gamma ELISPOT assays. J Immunol Methods 259:95

39. Speiser DE, Pittet MJ, Guillaume P, Lubenow N, Hoffman E, Cerottini J-C, Romero P (2004) Ex vivo analysis of human antigen-specific CD8+ T-cell responses: Quality assessment of fluorescent HLA-A2 multimer and interferon-γ ELISPOT assays for patient immune monitoring. J Immunother 27:298

40. Comin-Anduix B, Gualberto A, Glaspy JA, Seja E, Ontiveros M, Reardon DL, Renteria R, Englahner B, Economou JS, Gomez-Navarro J, Ribas A (2006) Definition of an immunologic response using major histocompatibility complex tetramer and enzyme-linked immunospot assays. Clin Cancer Res 12:107

41. Janetzki S, Panageas KS, Ben-Porat L, Boyer J, Britten CM, Clay TM, Kalos M, Maecker HT, Romero P, Yuan J, Kast WA, Hoos A (2007) Results and harmonization guidelines from two large-scale international elispot proficiency panels conducted by the Cancer Vaccine Consortium (CVC/SVI). Cancer Immunol Immunother. (doi:10.1007/s00262-007-0380-6)