Harmonization and qualification of an IFN-γ Enzyme-Linked ImmunoSpot assay (ELISPOT) to measure influenza-specific cell-mediated immunity within the FLUCOP consortium

Gwenn Waerlop1*, Geert Leroux-Roels1, Teresa Lambe2, Duncan Bellamy2, Donata Medaglini3, Elena Pettini3, Rebecca Jane Cox4, Mai-Chi Trieu4, Richard Davies4, Geir Bredholt4, Emanuele Montomoli5,6, Elena Gianchecchi6 and Frédéric Clement1

1Center for Vaccinology (CEVAC), University Hospital, Ghent University, Ghent, Belgium, 2Nuffield Department of Medicine, The Jenner Institute, University of Oxford, Oxford, United Kingdom, 3Laboratory of Molecular Microbiology and Biotechnology, Department of Medical Biotechnologies, University of Siena, Siena, Italy, 4Influenza Centre, Department of Clinical Science, University of Bergen, Bergen, Norway, 5Department of Molecular and Developmental Medicine, University of Siena, Siena, Italy, 6VisMederi srl, Siena, Italy

Influenza continues to be the most important cause of viral respiratory disease, despite the availability of vaccines. Today’s evaluation of influenza vaccines mainly focuses on the quantitative and functional analyses of antibodies to the surface proteins haemagglutinin (HA) and neuraminidase (NA). However, there is an increasing interest in measuring cellular immune responses targeting not only mutation-prone surface HA and NA but also conserved internal proteins as these are less explored yet potential correlates of protection. To date, laboratories that monitor cellular immune responses use a variety of in-house procedures. This generates diverging results, complicates interlaboratory comparisons, and hampers influenza vaccine evaluation. The European FLUCOP project aims to develop and standardize assays for the assessment of influenza vaccine correlates of protection. This report describes the harmonization and qualification of the influenza-specific interferon-gamma (IFN-γ) Enzyme-Linked ImmunoSpot (ELISpot) assay. Initially, two pilot studies were conducted to identify sources of variability during sample analysis and spot enumeration in order to develop a harmonized Standard Operating Procedure (SOP). Subsequently, an assay qualification study was performed to investigate the linearity, intermediate precision (reproducibility), repeatability, specificity, Lower and Upper Limits of Quantification (LLOQ-ULOQ), Limit of Detection (LOD) and the stability of signal over time. We were able to demonstrate that the FLUCOP harmonized IFN-γ ELISpot assay procedure can accurately enumerate IFN-γ secreting cells in the analytical
within this consortium, efforts have been made to develop assessment of influenza vaccination and SARs. To date, laboratories that perform cellular immune responses need to be harmonized in order to avoid diverging results, complicating interlaboratory comparisons, and generate reliable data. This harmonization step is crucial for vaccine development and public health decisions. In addition, a reliable assay for the detection of influenza-specific cellular immune responses is needed to investigate influenza-specific cellular immune responses induced by natural infection or vaccination and can be an additional instrument in the search for novel correlates of protection.

**Introduction**

Influenza continues to be the most important cause of viral respiratory disease associated with millions of hospitalizations and hundreds of thousands of deaths, despite the availability of vaccines (1, 2). The current seasonal human influenza vaccines are moderately effective in certain populations but require annual updating and administration. Additionally, the vaccine effectiveness varies depending on the match between the vaccine and the circulating strains. To overcome these shortcomings, there is an urgent need for new or improved influenza vaccines and efforts are already being made to design long-lasting universal influenza vaccines effective against different variants (3). Today’s evaluation of influenza vaccines mainly focuses on the quantitative and functional analyses of antibodies to haemagglutinin (HA) and neuraminidase (NA), the major surface glycoproteins of the virus. Cellular immune responses, primarily mediated by T cells, not only target the mutation-prone surface proteins but also internal proteins that are generally more conserved and shared by heterologous viral strains across influenza A subtypes. Consequently, vaccines inducing cellular immune responses are more likely to elicit broad protection against heterologous viral strains.

Reliable detection and quantification of these cellular responses are of key interest. Therefore, cellular immune assays need to be qualified and, if possible, validated to demonstrate assay precision, robustness and specificity before being applied in clinical trials. To date, laboratories that monitor cellular immune responses use a variety of in-house procedures. This generates diverging results, complicates interlaboratory comparisons, and hampers influenza vaccine evaluation (4–6). The European FLUCOP project, supported by the Innovative Medicines Initiative Joint Undertaking (IMI-JU, under Grant Agreement 115672), aims to develop and standardize assays for the assessment of influenza vaccine correlates of protection (7). Within this consortium, efforts have been made to develop harmonized Standard Operating Procedures (SOPs) for influenza-specific interferon-gamma (IFN-γ) Enzyme-Linked Spot (ELISpot) and Intracellular Cytokine Staining (ICS) assays, followed by assay qualification. These two cell-mediated immunity (CMI) assays allow for the detection and quantification of antigen-specific cytokine responses to vaccination and infection. The ELISpot assay specifically aims at quantifying IFN-γ producing cells, such as CD4+ Th1 and CD8+ T cells which are the prime subsets of interest when examining influenza-specific responses induced by vaccination and/or infection. Note that other cell populations such as NK- and NK T cells can secrete IFN-γ and contribute to the spot formation in the plate (7–9). Depending on the research questions asked, other cytokines (e.g., IL-2, IL-4, IL-5) (10) can be investigated by ELISpot. The harmonization and qualification of the ICS assay are described in a separate report in this special topic issue (Begue et al., 2022. Harmonization and Qualification of Intracellular Cytokine Staining to Measure Influenza-Specific CD4+ T Cell Immunity Within the FLUCOP Consortium (submitted)).

This report describes the harmonization and qualification of the IFN-γ ELISpot assay. First, two pilot studies were conducted to identify sources of variability during sample analysis and spot enumeration with the aim of developing a harmonized SOP. Finally, an assay qualification study was performed to investigate the linearity, intermediate precision (reproducibility), repeatability, specificity, Lower and Upper Limits of Quantification (LLOQ-ULOQ), Limit of Detection (LOD) and the stability of signal over time.

**Materials and methods**

**Samples**

Peripheral blood mononuclear cells (PBMC) were isolated from 12 buffy coats obtained from healthy blood donors (Red Cross Flanders). PBMC were also isolated from blood sampled...
from 27 healthy volunteers that participated in a clinical vaccine trial that was carried out specifically for these studies. For this, venous blood was collected in heparin-coated blood collection tubes prior to and 7 days after the administration of a seasonal influenza vaccine (alfa-RIX-Tetra® (season 2015/2016 or 2016/2017)). Ethical approvals for this study and the use of blood collected from Red Cross donors were given by the Ethical Committee of the Ghent University Hospital.

PBMC were isolated according to the standardized procedure FLUCOP SOP for PBMC isolation and cryopreservation, available as Supplementary Material (Appendix 1). In brief, venous blood samples were diluted 1:2 in Hanks buffered salt solution (HBSS), anduffy coats were brought to a total volume of 300 mL in HBSS. PBMC were isolated by isopycnic centrifugation using Lymphoprep™. Subsequently, cells were washed twice in HBSS, suspended in freezing medium [10% dimethyl sulfoxide (DMSO)/90% fetal bovine serum (FBS)], frozen at a concentration of ≥5 to ≤20 million cells/mL within 24 h (buffy coats) or 6 h (whole venous blood samples) after blood collection and finally stored in liquid nitrogen until use.

All cryovials were identified with unique codes without any reference to their source. All samples later distributed to other laboratories were selected from this PBMC biobank based on their pre-examined CMI immune responses against influenza antigens.

### Pilot studies

**Pilot study 1**

In pilot study 1, different methods applied to analyse samples were carefully examined. A panel of 24 coded PBMC samples was distributed to 5 consortium partners. Each vial had a unique code allowing only the organizing center to link it with the original specimen identifier. Hereafter, sample identifiers 1 to 24 will be used. Participating laboratories were also provided with antigens for *in vitro* stimulations of the PBMC. These were recombinant Hemagglutinin (HA) H1 A/California/07/2009 (catalog number 3006) and recombinant Hemagglutinin (HA) B/Phuket/3073/2013 (catalog number 3006), both from Protein Sciences Corporation (Swiftwater, PA) and kindly provided by Sanofi. The recombinant proteins were produced in insect cells using the baculovirus expression vector system and purified to at least 90%. Each laboratory was asked to analyze the samples using their in-house procedure for IFN-γ ELISPOT and preferred reagents. An online worksheet was filled out to collect information regarding the number of vials thawed, thawing medium, thawing medium temperature, thawing process, FBS validation status, cell counting technique, the technique applied for determination of cell viability, cell resting time and medium as well as the concentration of cells during the resting period (where applicable), culture medium, the use of self- or precoated plates, coating antibody, conjugated detection antibody, cell concentration in the well, stimulation/incubation time, substrate for staining, the process of washing the wells, stop reaction, ELISPOT reader and related software, Quality Control (QC) process, validation criteria on background conditions, and any comments/deviations that may have occurred. Minimal instructions on data reporting were provided and linked to the lab ID, stimulation condition, mortality percentage, sample ID, investigated marker and counted spots.

**Pilot study 2**

In pilot study 2, the variation in spot interpretation and data reporting was investigated. An ELISPOT plate prepared by one partner (University of Oxford) was distributed to 6 other partners for read-out within 30 days. Storage and transport of this plate were performed at room temperature. Each lab was asked to read the plate shortly after reception according to their in-house procedure, report the data according to minimal instructions provided by the organizing center, and send the plate to the next lab according to the predefined distribution schedule.

Results from both pilot studies and information provided via the online worksheets were collected and centrally analyzed by the organizing center (Center for Vaccinology, Ghent University, Belgium). The processed data is reported in a blinded manner in this study.

**IFN-γ ELISPOT assay procedure and antigen titration**

For the pilot studies, the consortium partners were asked to use their in-house procedures and preferred reagents. During assay qualification experiments performed by one FLUCOP partner, the SOP for IFN-γ ELISPOT developed by FLUCOP was applied. This procedure is available as Supplementary Material (Appendix 2). Briefly, the FLUCOP IFN-γ ELISPOT assay was performed using the human IFN-gamma ELISPOT PRO kit (3420-2APW-2, Mabtech). On the first day of the 2-day procedure, the pre-coated plates were re-hydrated by washing 4 times with PBS 1X and were blocked by adding 200 µL culture medium (RPMI supplemented with penicillin/streptomycin, L-glutamine, sodium pyruvate, MEM, β-mercaptoethanol, FBS). Plates were stored for 2 to 4 hours at 37°C. Meanwhile, the samples were thawed in a water bath at 37°C and washed in culture medium. The cell suspension was centrifuged and cells were washed again in culture medium supplemented with benzamide. After centrifugation, cells were resuspended in culture medium. Cell concentration was measured using a Sysmex hematology analyzer (XN-L-350) and cell viability was determined with propidium iodide (PI; staining of dead cells) using a flow cytometer. After centrifugation, cells were resuspended in culture medium to obtain a concentration of 4 x 10⁶ PBMC/mL. The blocking solution was removed from the assay plates and 100 µL stimulation reagent was added to each well. PBMC were plated by adding 200 000 PBMC per well (in 50 µL). The plates were
stored overnight (18-24 hours) in an incubator at 37°C and 5% CO2. The next day, cell suspensions were removed and plates were washed 5 times with 200 µL lab-grade water per well. Next, 100 µL of detection antibody was added and the plates were incubated for another 2 hours at room temperature. Plates were washed 3 times with 200 µL/well lab-grade water, followed by 3 more washes with 200 µL/well PBS. Substrate (50 µL) was added to each well to visualize the spots. Plates were stored at room temperature and were air-dried in the dark. Spots were enumerated within 7 days after the start of the assay using an automated spot counter (ImmunoSpot® S5, Cellular Technology Limited), followed by manual verification of the identified spots (Quality Control).

The influenza-specific stimulating agent selected for the assay qualification experiments was split A/California (H1N1) virus (batch FA593899), kindly provided by Sanoﬁ. Antigen was titrated to determine the optimal stimulation concentration (Supplementary Figure 1) and investigate any potential cellular toxicity. A final concentration of 1.25 µg/mL was found to be the most appropriate stimulating condition. At this concentration, no cellular cytotoxicity was observed (data not shown) and therefore, this concentration was selected for further PBMC stimulations in ELISpot assays. In some qualification experiments, tetanus toxin (T3194-25UG, Sigma-Aldrich) was used as a control antigen to stimulate the cells in vitro at a final concentration of 2 µg/mL. The mock or background condition consisted of PBMC cultured in medium only. Each condition was tested in triplicate and mean responses were reported.

Data analysis

Results from the IFN-γ ELISpot assay were expressed as Spot Forming Units (SFU) per million PBMC and reported as the mean of triplicates, duplicates or as singletons, depending on the lab and/or test condition. For the evaluation of antigen-specific responses, background-subtracted results were used. Negative values were corrected to 1 SFU/10^6 cells. Results from the pilot studies and the specificity experiment were log-transformed prior to further calculations. Individual lab results were compared to the geomean, descriptive statistics were performed and a Z-score was attributed.

The Z-score expresses the relationship of a reported value to the mean of a group of values in terms of standard deviations and shows where that result is positioned in the distribution of all reported values. The Z-score is therefore a good tool to assess the performance of an individual lab within a group. Still, it does not give information on the total imprecision of the group data. A Z-score of 0 indicates that the reported result is identical to the mean result. A Z-score of 1 indicates that the reported result differs one standard deviation from the mean. Z-scores can be positive or negative, with a positive value indicating the score is above the mean and a negative score indicating it is below the mean.

\[ z = \frac{x - \mu}{\sigma} \]

\( \mu = \text{mean}; \sigma = \text{standard deviation} \). As an arbitrary rule of thumb, an absolute Z-score ≤ 1 can be considered as excellent, a Z-score between 1 and 2 as acceptable and a Z-score >2 should be considered as poor performance requiring further investigation.

Results

Pilot study 1

In pilot study 1, a large variation was already observed in the unstimulated conditions (Figure 1). Two out of 5 labs reported background responses exceeding 50 SFU/million PBMC in more than half of the samples (Figure 1A). This 50 SFU/million PBMC threshold is a commonly applied acceptance or quality criterion (11–13). This large variation is reflected in the Z-scores (Figure 1B) where lab A generally reported higher responses, lab D lower responses, lab C showed a dispersed profile and labs B and E reported values close to the group mean (Z-scores
between -1 and 1). The overall mean coefficient of variation (CV) of the background responses was 64% (Figure 1C).

Significant variation was also observed in the antigen-stimulated cultures. The Z-scores demonstrated that all labs reported values relatively close to the group mean, in other words, no lab generated data deviating with more than 2 standard deviations from the group mean. However, the imprecision, expressed as the mean CV, of cultures stimulated with A/California and B/Phuket demonstrated high levels of variation, with 63% and 34% CV, respectively (Figure 2). The lower mean CV upon stimulation with B/Phuket can be explained by the higher range of responses observed, 48.9-2069 SFU/million PBMC, whereas A/California elicited responses in the range of 8.3-319.9 SFU/million PBMC.

Pilot study 2

In pilot study 2, the variation that can be induced upon the enumeration of IFN-γ secreting cells by interpreting developed spots was examined. Figure 3 shows that the Z-scores of the reported counts varied between -1 and 1 (panel A) and that the highest CV values were observed in the lower part of the analytical range (< 5 SFU/well or 25 SFU/million PBMC; panel B). Although limited variation was observed in the enumeration of the spots (mean CV of 25.8%), high diversity in reporting conditions with spots too dense to count was noticed. Some labs reported these particular conditions with a symbol (“*” or “°”), one lab with a code (“-2”), one lab with an abbreviation (TNTC, too numerous to count) and two labs with zero (“0”). This observation indicates that careful thought should be given to how such a condition should be flagged and reported. Especially the zero (“0”) or “-2” codes can lead to misinterpretation when evaluating (background-subtracted) results as this will still generate a (negative) number. It is important to clearly distinguish non-numerical values (e.g. in case of too dense spots preventing accurate counting) from numerical results (i.e. actual data). A harmonized reporting code needs to be defined.

Establishment of the FLUCOP SOP

Based on information collected via the online worksheet (pilot study 1) and a thorough comparison of all applied procedures, a set of potential critical parameters was identified. These were: the cell count technique, application and duration of a cell resting phase, the use of self- versus pre-coated plates, the concentration of cells in the wells, the time of incubation, the medium used during stimulation, the wash step, the ELISpot reader and the FBS validation status (Table 1).
Within the consortium, a consensus was reached to define specific parameters that had to be strictly applied in the standardized operating procedure for IFN-γ ELISpot testing, whereas for other parameters only a recommendation was proposed (Appendix 2).

**IFN-γ ELISpot assay qualification**

**Linearity**

Linearity of an assay is its ability to provide test results that are directly proportional to the concentration of the measurand (quantity to be measured) in a test sample. A standard approach to assess the linearity of a laboratory method consists of diluting a test sample in a negative sample matrix to demonstrate the assay’s ability to reproduce the initial result after recalculation of the result obtained after dilution. In complex cell-based assays, where no reference materials are available, this approach requires the availability of non-responding cells to dilute the cell(s) of interest. In the IFN-γ ELISpot assay, where *in vitro* stimulation occurs immediately prior to the read-out and in the same final reaction vessel, this would require that the diluting cells do not induce any allogeneic or bystander effect on the "tested cells". As this cannot be achieved, an adapted method was designed.

The linearity of the IFN-γ ELISpot was assessed using 4 PBMC samples that were independently fractionally diluted in culture medium without (background condition) or with split A/California virus as stimulating antigen. This was done in 3 replicates and repeated 3 (sample 4) or 4 (samples 1-3) times.

**TABLE 1** Overview of the parameters that were identified as potentially critical for the outcome of the IFN-γ ELISpot assay.

| Lab ID | Cell count technique | Cell resting | Duration of cell resting (hours) | Plates | Concentration of cells in wells | Incubation time (hours) | Culture medium used in cell resting and during stimulation | Wash step | ELISpot reader | FBS validation |
|--------|----------------------|--------------|----------------------------------|--------|---------------------------------|------------------------|------------------------------------------------------------|----------|---------------|---------------|
| Lab 01 | Automated            | Yes          | 4                                | Self-coated | 200.000                      | 18                    | cRPMI                                                      | Automated | AID-ELISpotSpectrum | Validated     |
| Lab 02 | Automated            | No           | No                               | Self-coated | 500.000                      | 22                    | cRPMI                                                      | Manually  | ImmunoSpot S6, CTL | Validated     |
| Lab 03 | Automated            | Yes          | 19                               | Pre-coated | 300.000                      | 24                    | cRPMI                                                      | Manually  | ImmunoScan, CTL   | No info       |
| Lab 04 | Manual               | Yes          | 3                                | Self-coated | 250.000                      | 19                    | cRPMI                                                      | Manually  | ELISpot 7.0 - iSpot | Validated     |
| Lab 05 | Automated            | Yes          | 1.5                              | Pre-coated | 200.000                      | 19.5                  | cRPMI                                                      | Manually  | ImmunoSpot S6, CTL | No           |
| Consensus protocol | Validated method | Yes          | 3                                | Free of choice | 200.000                      | Overnight (18-24 h) | Complete RPMI (cRPMI) | Your validated method | Your validated equipment | Validated FBS |
| Fixed or as recommended | Recommended | Fixed          | Fixed                            | /       | Fixed                          | Fixed                   | Recommended                                                 | Recommended | Recommended | Fixed |

**FIGURE 3**

Pilot study 2. An IFN-γ ELISpot plate with 84 wells to be scored was distributed to 6 laboratories within one month after preparation. Z-scores per well per lab (A) and the CV per well (B) are shown. The dotted line in the right panel indicates the CV of 40%, commonly applied as a threshold of acceptable variation in CMI assays.
Fractional dilution of the samples was done by plating 40 000 to 400 000 cells per well. The microculture with 200 000 cells per well was selected as the reference condition for calculating the recovered response.

At this condition of 200 000 cells per well, 127.8 – 15.3 – 15.5 and 13.6 SFU/well were counted in samples 1, 2, 3 and 4, respectively (Table 2). Expressed as SFU per million PBMC, this translated into 639.2 – 76.7 – 77.5 and 67.9 SFU for samples 1, 2, 3 and 4, respectively.

As a rule of thumb, a recovery value between 50 and 150% of the observed reference value, or in other words the ability to detect a decrease or increase of the signal with 50%, is regarded as "linear". Table 2 clearly shows that the responses were linear within the assay conditions ranging between 120 000 cells and 360 000 cells per well. Within this technical assay range, the lowest background-subtracted response observed was 3.5 SFU/80 000 cells. The proportionality of the method was further assessed by calculating the ratios of the number of plated cells versus the observed number of spots per well. Two variables are considered proportional if their corresponding elements have a constant ratio, which is called the coefficient of proportionality. The correlation curves demonstrated high correlations in all 4 measured samples with R² values ranging from 0.9734 to 0.9916 (Figure 4).

Within the linear range of 120 000 to 360 000 cells per well and taking into consideration the acceptable recovery range of 50–150%, the lowest and highest acceptable number of spots per well are 3.5 and 331.3 SFU. If this is extrapolated to the standard condition of 200 000 cells per well, this can be further extrapolated to the standard condition of 200 000 cells per well. The microculture with 200 000 cells per well was selected as the reference condition for calculating the recovered response.

## Intermediate precision and repeatability

The IFN-γ spot forming responses elicited by A/California and TT antigens were measured in 10 PBMC samples. Each sample was tested in duplicate by 2 operators performing each 2 runs on different days, resulting in 8 measurements per sample. The mean responses ranged from 34.4 to 327.5 SFU/10⁶ cells and from 2.5 to 201.5 SFU/10⁶ cells after stimulation with A/California and TT, respectively (Figure 5A). The variability was assessed by calculating the CV.

Inter-assay or intermediate precision was assessed as the average CV% of all individual measurements per sample after stimulation with TT and A/California and ranged from 14 to 107%, with a mean of 41%. The intra-assay precision or repeatability was assessed and determined across all mean responses obtained per run and ranged from 0 to 36%, with a mean of 21%. The inter-operator precision was also assessed and determined across all mean responses collected per operator for a particular test condition and ranged from 0 to 60%, with a mean of 12%. The level of variation increased substantially in the lower part of the analytical range of this assay, as shown in Figure 5C.

The Lower Limit of Intermediate Precision (LLOIP) was determined by plotting the mean background-subtracted responses to TT and A/California against the mean inter-assay CV% per sample (Figure 5B). The LLOIP was defined as the lowest value with a CV of 40%. A log-log line was considered as the best-fitted curve with an R² = 0.8729 and y=10^((-0.3820*log (x))+2.189), where y is the mean CV% and x the mean SFU/10⁶ cells per sample. The lowest background-subtracted response with a CV of 40% was determined at 34.4 SFU/10⁶ cells.

The Lower Limit of Quantification (LLOQ) was defined as the highest value of the LLOL and LLOP and therefore set at 34.4 SFU per million cells. In the white paper published by Corsaro

### Table 2 Observed SFU per well at each dilution and recovery.

| Number of cells plated | Sample 1 | Sample 2 | Sample 3 | Sample 4 |
|------------------------|----------|----------|----------|----------|
|            | Observed mean SFU/well | Observed mean SFU/well | Observed mean SFU/well | Observed mean SFU/well | Observed mean SFU/well | Observed mean SFU/well | Observed mean SFU/well | Observed mean SFU/well | Observed mean SFU/well |
| 40 000      | 7.6      | 1.0      | 0.4      | 1.5      |
| 80 000      | 26.3     | 3.7      | 3.2      | 3.5      |
| 120 000     | 58.3     | 5.0      | 5.1      | 6.9      |
| 160 000     | 89.2     | 10.0     | 9.0      | 8.4      |
| 200 000     | 127.8    | 15.3     | 15.5     | 13.6     |
| 240 000     | 155.9    | 17.3     | 18.3     | 17.0     |
| 280 000     | 191.4    | 21.7     | 22.7     | 22.4     |
| 320 000     | 221.8    | 27.7     | 26.7     | 25.0     |
| 360 000     | 270.7    | 30.7     | 36.2     | 31.3     |
| 400 000     | 331.3    | 38.3     | 41.6     | 34.9     |

Recovery is calculated as the percentage of mean SFU/10⁶ cells observed in the reference condition of 200 000 cells plated per well expressed as SFU/10⁶ cells (indicated in bold).
et al. (14), the acceptance criterion for the intermediate precision is defined as ≤40% for ≥80% of the samples having mean spot forming units/million cells exceeding the LLOQ. In this study, 96 out of 160 obtained values exceeded the LLOQ determined at 34.4 SFU/million cells, resulting in an intermediate precision of 20% (ranging from 4 to 39%).

Specificity

Assay specificity was determined by stimulating PBMC from 25 paired samples collected before and 7 days after the administration of a seasonal influenza vaccine. PBMC were stimulated with split A/California virus and tetanus toxin, a control antigen unrelated to the vaccine. Specificity was demonstrated if a significant increase in signal was observed between the pre- and post-vaccination samples upon stimulation with the selected influenza-specific antigen (p < 0.05) and if no significant increase was observed after stimulation with the control antigen (p ≥ 0.05). Figure 6 shows a significant increase of the influenza-specific response (p < 0.001; Wilcoxon signed-ranked test) and no difference in the TT-specific response (p = 0.2872, Wilcoxon signed-rank test).

ULOQ

PBMC from two samples were plated each on one plate. The cells were added to the plates in a serial dilution ranging from 100 000 to 48 cells, with 8 repeats per dilution. All wells were stimulated with the superantigen SEB (staphylococcal enterotoxin B) at a final concentration of 0.25 µg/mL. The Upper Limit of Quantification (ULOQ) is the highest number of counts per well that can be detected and reliably quantified, i.e., a CV% ≤ 40. For both samples, reliable counts were reported up to and including the test condition where 50 000 PBMC were stimulated with SEB with CVs of 4 and 6% Figure 7 (sample 1) and Supplementary Figure 2 (sample 2). The wells seeded with 100 000 PBMC generated too many spots preventing an accurate spot count. The determination of the ULOQ is primarily defined by the instrument and the operator(s) and depends on the availability of strong positive samples. The highest responses that could be enumerated accurately in this analyzing lab were 452.25 and 458.75 SFU/50 000 PBMC or 9045 and 9175 SFU/10^6 PBMC. Based on this information, it can be proposed that samples or conditions resulting in too many spots are reported as ‘> 10 000 SFU/10^6 PBMC’.

Limit of detection (LOD)

The Limit of Detection (LOD) is the lowest detectable analyte concentration that can be reliably distinguished from analytical noise (15). The technical LOD is calculated as the 95th percentile of the non-specific responses after mock stimulation (medium only). Based on 402 values from the linearity, intermediate precision, repeatability and specificity experiments, the LOD was determined at 12.70 SFU/well or
FIGURE 5
Determination of Lower Limit of Intermediate Precision (LLOIP). (A) Distribution of background-subtracted responses after in vitro stimulation of cells with A/California and tetanus toxin (TT). Each symbol represents a measurement of a certain sample, indicated on the x-axis. A set of 10 samples was tested each in duplicate by 2 operators performing each 2 runs on different days, resulting in 8 measurements per sample. Measurements from operator 1 are indicated with circles and from operator 2 with triangles. Responses obtained after stimulation with A/California are indicated in blue, with TT in green. (B) LLOIP was determined by plotting the mean background-subtracted responses to A/California and TT against the mean inter-assay CV% per sample. The LLOIP was defined as the lowest response with a CV of 40% and was determined at 34.4 SFU/10⁶ cells. (C) The inter-assay, intra-assay and inter-operator variability were assessed by the mean CV% observed for the mean responses of each sample. Each horizontal dotted line represents the cut-off of 40% CV and each vertical dotted line represents the LLOIP determined at 34.4 SFU/10⁶ cells.

FIGURE 6
Determination of assay specificity. PBMC from 25 paired samples collected before and 7 days after the administration of a seasonal influenza vaccine were stimulated with influenza (split A/California) and control (tetanus toxin; TT) antigens. Background-subtracted values are shown as box plots. Differences in IFN-γ spot forming responses between the pre- and post-vaccination samples following in vitro stimulation with either influenza or TT antigen were examined. Wilcoxon matched-pairs signed-rank test was applied and p < 0.05 was considered statistically significant.
63.50 SFU/million PBMC. This cut-off could be applied as a criterion of acceptance for background responses instead of the commonly used arbitrary albeit more stringent cut-off of 50 SFU/million PBMC.

Validity of signal stability over time

A set of 9 developed plates was re-read every other week for 24 weeks. The stability of the signal was demonstrated as the counts did not change significantly during this period (Figure 8). The total counts of 96 wells per plate varied with CVs between 2 to 5%, with a decrease of the total sum of counts in 8 out of 9 plates compared to the first read (-1 to -4%) and one plate with an increase of 14%. The absence of striking differences in the reported values for these 3 different operators appears to be an indication of a low inter-operator variability. It is assumed that all involved operators were equally qualified in spot enumeration and because of the limited number of reads per operator this was not further examined. The effect of time was statistically analyzed by performing one-way ANOVA tests. The tests demonstrated a statistically significant effect of time as of week 7 (p < 0.0001), meaning that the plates remain stable and can be stored up to and including 6 weeks after development.
A summary of all qualification results of the IFN-γ ELISpot assay is shown in Table 3.

**Discussion and conclusion**

The IFN-γ ELISpot and intracellular cytokine staining (ICS) assays are frequently used to examine cellular immune responses elicited by influenza infection or vaccination (16–21). A better insight into the magnitude, quality and durability of cell-mediated immune responses can improve our understanding of the immunological mechanisms underlying viral clearance or vaccine effectiveness and may contribute to the identification of new correlates of protection. The first prerequisite to generate reliable data is the quality of the samples to be studied. To produce high-quality PBMC, it is desirable to apply a standardized procedure for the isolation and cryopreservation of PBMC that has been proven to be compatible with the envisaged downstream analyses (22–25). Secondly, laboratories that monitor cellular immune responses with the IFN-γ ELISpot assay apply a variety of in-house procedures. This generates diverging results, complicates interlaboratory comparisons, and hampers a reliable evaluation of the immunogenicity of influenza vaccines (4–6). For these reasons, there is still a great need to harmonize the procedures and provide guidance on how to report the assay results in a standardized manner. Within the European FLUCOP consortium, we first developed SOPs for influenza-specific IFN-γ ELISpot (described in this paper) and ICS assays (published in Begue et al. 2022. *Harmonization and Qualification of Intracellular Cytokine Staining to Measure Influenza-Specific CD4+ T Cell Immunity Within the FLUCOP consortium (submitted)*) and subsequently performed qualifications of the assays.

Two pilot studies performed in 5 to 6 labs allowed us to identify critical and less critical parameters that can induce assay variation. When every lab applied its in-house procedure, the observed variation, expressed as the coefficient of variation, was 64% in the unstimulated conditions and 63 and 34% in the cell cultures stimulated with split A/California and B/Phuket virus, respectively. Lower variation (mean CV of 25.8%) was observed when only the spot enumeration was assessed. However, high diversity was noticed in conditions where the spots were too dense to count. Based on the reported data and observed ULOQ, \( > 10,000 \text{ SFU/10}^6 \text{ PBMC} \) is proposed as a harmonized reporting code. The overall heterogeneity observed in both pilot studies was considered modest. This may be due to the use of commercial IFN-γ ELISpot kits by several laboratories, a very advantageous possibility not available for all immunoassays. Several assay variables were identified as potential critical parameters: the cell counting technique, the use and duration of a cell resting phase, the use of self- versus pre-coated plates, the cell

| Assay parameter         | Acceptance criteria                                                                 | Qualification outcome                  |
|-------------------------|--------------------------------------------------------------------------------------|----------------------------------------|
| Intermediate precision  | The CV for reproducibility should be ≤ 40% (*The CV for reproducibility should be ≤ 40% for ≥80% of the samples having SFU/10^6 PBMC > LLOQ.*) The Lower Limit of Intermediate Precision is defined by the lowest measurement with a CV of 40%. | Stimulation with A/California: Inter-assay CV is 27% (*20%) Intra-assay CV is 10% (*8%) Inter-operator CV is 8% (*8%) Stimulation with TT: Inter-assay CV is 55% (*19%) Intra-assay CV is 31% (*14%) Inter-operator CV is 17% (*6%) Pooled data: Inter-assay CV is 41% (*20%) Intra-assay CV is 21% (*9%) Inter-operator CV is 12% (*7%) LLOIP = 34.4 SFU/10^6 PBMC |
| Linearity               | The range of cell densities (plated cells) through which the recovery was between 50 and 150%. | Linear range: from 120 000 to 360 000 cells/well |
| LLOQ                    | The highest value observed of the Lower Limit of Precision and Lower Limit of Linearity. | LLOQ = 34.4 SFU/10^6 PBMC |
| Specificity             | Comparison of pre- and post-vaccination samples: Significant increase after influenza-specific stimulation of the cells (p < 0.05) AND Non-significant increase after stimulating the cells with control (non-vaccine) antigen (p ≥ 0.05) | Stimulation with A/California: p < 0.0001 Stimulation with TT: p = 0.7260 |
| ULOQ                    | The highest number of counts per well that can be detected and reliably quantified, i.e. with a CV ≤ 40. | ULOQ = 458.75 SFU/50 000 PBMC |
| LOD                     | The 95th percentile of the responses after mock stimulation. | LOD = 12.70 SFU/well or 63.50 SFU/million PBMC |
| Signal stability over time | No significant waning of counts over time. | No significant effect of time up to and including 6 weeks after development (p ≥ 0.0001). |
In conclusion, the FLUCOP harmonized IFN-γ ELISpot assay procedure can accurately enumerate IFN-γ secreting cells in the analytical range of 34.4 SFU/million cells up to the technical limit of the used reader and in the linear range from 120 000 to 360 000 cells per well, in plates stored up to 24 weeks after development. We hope this harmonized IFN-γ ELISpot procedure will become a useful and reliable tool to investigate influenza-specific cellular immune responses induced by natural infection or vaccination and will be an aid in the search for novel correlates of protection. We estimate that this harmonized assay may also be applied to cellular immune responses against other (respiratory and non-respiratory) infectious pathogens.

**Data availability statement**

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.
Ethics statement

The studies involving human participants were reviewed and approved by the Ethical Committee of the Ghent University Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

FC, GL-R, and GW contributed to the conception and design of the study, data analysis and interpretation, and wrote the first draft of the manuscript. TL, DB, FC, GL-R, and GW contributed to the data acquisition. RC contributed to the critical revision of the manuscript. All authors have read and approved the final version for submission.

Funding

The FLUCOP project is supported by the Innovative Medicines Initiative-Joint Undertaking under grant agreement 115672, resources of which are composed of financial contribution from the European Union’s Seventh Framework Programme (FP/2007-2013) and EFPIA companies’ kind contribution.

Acknowledgments

The authors wish to thank Annelies Goussaert, Peter Vander Linden, Sanne Foubert and Sibyl Couvent for the data acquisition and Sara Tete for the critical review of the manuscript. The authors would like to acknowledge their FLUCOP consortium collaborators for their assistance: Catherine Caillot, Barbara Camilloni, Maria Rita Castrucci, Marco Cavaleri, Annalisa Ciabattini, Simon De Lusignan, Oliver Dibben, Othmar Engelhardt, Susanna Maria Roberta Esposito, Marzia Facchini, Felipa Ferreira, Sophie Germain, Sarah Gilbert, Sammy Ho, Katja Hoschler, Sarah L Jalloh, Stefan Jungbluth, Marion Koopmans, Manuela Mura, Nedzad Music, Martina Ochs, Thierry Ollinger, Albert Osterhaus, Anke Pagnon, Giuseppe Palladino, Elena Pettini, Ed Remarque, Leslie Reperant, Hanna Sediri Schön, Sarah Tete, Alexandre Templier, Claudia Trombeta, Serge van de Witte, Ralf Wagner, Joanna Waldock, Brenda Westerhuis, Fan Zhou.

Conflict of interest

Authors EM and EG are employed by VisMederi srl. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2022.984642/full#supplementary-material

References

1. de Courville C, Cadarette SM, Wissinger E, Alvarez FP. The economic burden of influenza among adults aged 18 to 64: A systematic literature review. Influenza Other Respir Viruses (2022) 16:376–85. doi: 10.101111/trv.12963
2. Iuliano AD, Roguski KM, Chang HH, Muc sabotello DJ, Palekar R, Tempia S, et al. Estimates of global seasonal influenza-associated respiratory mortality: A modelling study. Lancet (2018) 391(10127):1285–300. doi: 10.1016/S0140-6736 (17)33293-2
3. Krammer F, Smith GD, Foucher RAM, Peiris M, Kedziorska K, Doherty PC, et al. Influenza. Nat Rev Dis Primers (2018) 4(1):3. doi: 10.1038/s41572-018-0002-z
4. Cox NH, Ferrari G, Kalams SA, Lopaczynski W, Oden N, D’Souza MP. Results of an ELISPOT proficiency panel conducted in 11 laboratories participating in international human immunodeficiency virus type 1 vaccine trials. AIDS Res Hum Retroviruses (2005) 21(1):68–81. doi: 10.1089/aid.2005.21.68
5. Moodie Z, Price L, Janetzki S, Britten CM. Response determination criteria for ELISPOT: Toward a standard that can be applied across laboratories. Methods Mol Biol (2012) 792:185–96. doi: 10.1007/978-1-61779-325-7_15
6. Britten CM, Janetzki S, van der Burg SH, Gouttefangeas C, Hoos A. Toward the harmonization of immune monitoring in clinical trials: Quo vadis? Cancer Immunol Immunother (2008) 57:285–8. doi: 10.1007/s00262-007-0379-z
7. FLUCOP.
8. Desombere I, Clement F, Rigole H, Leroux-Roels G. The duration of in vitro stimulation with recall antigens determines the subset distribution of interferon-γ-producing lymphoid cells: A kinetic analysis using the interferon-γ secretion assay. J Immunol Methods (2005) 301(1–2):124–39. doi: 10.1016/j.jim.2005.04.008
9. Desombere I, Meuleman P, Rigole H, Willems A, Irsch J, Leroux-Roels G. The interferon gamma secretion assay: A reliable tool to study interferon gamma production at the single cell level. J Immunol Methods (2004) 286(1–2):167–85. doi: 10.1016/j.jim.2004.01.001
10. Frank K, Paust S. Dynamic natural killer cell and T cell responses to influenza infection. Front Cell Infect Microbiol (2020) 10. doi: 10.3389/fcimb.2020.00425
Cell mediated immune responses following revaccination with an influenza A/HSN1 vaccine. Vaccine (2016) 34(4):547–54. doi: 10.1016/ j.vaccine.2015.11.055

22. Higdon LE, Lee K, Tang Q, Maltzman JS. Virtual global transplant laboratory standard operating procedures for blood collection, PBMC isolation, and storage. Transplant Direct (2016) 2(9):e101. doi: 10.1097/ TXD.0000000000000613

23. Weinberg A, Song LY, Wilkening C, Sevin A, Blais B, Louzao R, et al. Optimization and limitations of use of cryopreserved peripheral blood mononuclear cells for functional and phenotypic T-cell characterization. Clin Vaccine Immunol (2009) 16(8):1176–86. doi: 10.1128/ CJIM.00342-08

24. Kierstead LS, Dubey S, Meyer B, Tobery TW, Moggi R, Fernandez VR, et al. Enhanced rates and magnitude of immune responses detected against an HIV vaccine: Effect of using an optimized process for isolating PBMC. AIDS Res Hum Retroviruses (2007) 23(1):86–92. doi: 10.1089/aid.2006.0129

25. Bull M, Lee D, Stucky J, Chiu YL, Rubin A, Horton H, et al. Defining blood processing parameters for optimal detection of cryopreserved antigen-specific responses for HIV vaccine trials. J Immunol Methods (2007) 322(1–2):57–69. doi: 10.1016/j.jim.2007.02.003

26. Britten CM, Gouttefangeas C, Welters MJ, Pawelec G, Koch S, Ottenheimer C, et al. The CIMT-monitoring panel: A two-step approach to harmonize the enumeration of antigen-specific CD8+ T lymphocytes by structural and functional assays. Cancer Immunol Immunother (2008) 57(5):289–302. doi: 10.1007/s00262-007-0378-0

27. Janetzki S, Rabin R. Enzyme-linked Immunospot (ELISPot) for single-cell analysis. In: AK Singh and A Chandrasekaran, editors. Single cell protein analysis: Methods and protocols. New York, NY: Springer New York (2015). p. 27–46. doi: 10.1007/978-1-4939-2987-0_3

28. Barabas S, Spindler T, Kiener R, Tonar C, Lugner T, Batzilia J, et al. An optimized IFN-γ ELISPot assay for the sensitive and standardized monitoring of CMV protein-reactive effector cells of cell-mediated immunity. BMC Immunol (2017) 18:14. doi: 10.1186/s12865-017-0195-y

29. Yang F, Patton K, Kasprzyk T, Long R, Gupta S, Zoog SJ, et al. Validation of an IFN-γ ELISPot assay to measure cellular immune responses against viral antigens in non-human primates. Gene Ther (2022) 29(1–2):41–54. doi: 10.1038/s41434-020-00214-w

30. Janetzki S, Panageras KS, Ben-Portz L, Boyer J, Britten CM, Clay TM, et al. Results and harmonization guidelines from two large-scale international elispot proficiency panels conducted by the cancer vaccine consortium (CVC/SVI). Cancer Immunol Immunother (2008) 57(3):303–15. doi: 10.1007/s00262-007-0380-6

31. Islam R, Vance J, Poirier M, Zimmer J, Khadang A, Williams D, et al. Recommendations on ELISPot assay validation by the GCC. Bioanalysis (2022) 14(4):187–93. doi: 10.4155/bio-2022-0010

32. Maerck HT, Hasler J, Payne JK, Summers A, Comatas K, Ghanayem M, et al. Precision and linearity targets for validation of an IFNγ ELISPOT, cytokine flow cytometry, and tetramer assay using CMV peptides. BMC Immunol (2008) 9. doi: 10.1186/1471-2172-9-9