



COVID-19
IMMUNITY
TASK FORCE

GROUPE DE TRAVAIL
SUR L'IMMUNITÉ
FACE À LA COVID-19

CORE DATA ELEMENTS

for Reporting of Cell-Mediated Immunity
Assay (ELISpot and Flow cytometry) Results

User Guide

PURPOSE OF THIS GUIDE

This user guide was originally made by the COVID-19 Immunity Task Force (CITF), with the intent of facilitating the use of Core Data Elements for reporting cell-mediated immunity assay results and data transfer.

INTRODUCTION

The standard definition and expression of core data elements (CDE) is necessary to support the electronic recording and harmonization of data. The CITF and its Field Studies and Immune Sciences Working Parties have developed a standardized list of core data elements for evaluating COVID-19 immunity.

The accompanying Excel document ([LINK](#)) lists commonly documented fields recorded by most study cohorts to enable standardized analysis, organized into a data dictionary. It also includes dataset templates for collected immunological correlates related to serology and cell-mediated immunity (ELISPOT or Flow cytometry-based data). [LINK CDE example guide](#).

DISTRIBUTION & REVIEW

Version 1.2
March 23rd, 2022

Written by: Mariana Bego
Updated by: Jaspreet Jain, Varun Anipindi and Hanna Swail
Revised and approved by:

TABLE OF CONTENTS

	Page
1. Core Data Elements Explained	4
<u>ELISpot assays:</u>	4
1.1. Line-by-line explanations	4
1.2. Reporting schemes	6
<u>Flow cytometry-based assays:</u>	7
1.3. Line-by-line explanations	7
1.4. Reporting schemes	11
2. List of preferred terms	12
2.1. Preferred SCT codes	12
2.2. Preferred antigen nomenclature	13

Section 1- Core Data Elements Explained

Sections map:

1. Core Data Elements Explained

- **ELISpot**
- **Flow Cytometry**

2. List of preferred terms

Only one ID?

It is possible that for any given project only either the Participant ID or Specimen ID is known to the investigator.

If that is the case, please leave the field unknown as blank.

Data linkage tables:

Independently to the data entered here, it is recommended that a table should be generated by the principal investigator detailing how many collections were done for each participant and their date, linking Participant ID, Specimen ID and Specimen collection date (lines 5-7).

The type of data that can be entered is designated as:

- Free text (Unconstrained alphanumeric string),
- Categorical (a restrained selection of terms is available),
- Numerical (Unconstrained numeric string), or
- Date (YYYY[MM[DD]]).

For the categorical selections use the SCT codes provided in the Concept Unique Identifier (CUI) table I in next section.

ELISpot

1.1. Line-by-line explanations

Each line is assigned a Number and a Label. They describe a field or variable in the dataset to be shared with the CITF. See explanation for each one below.

Line 1: Project ID

Unique ID assigned to each project and connected to funding documents. See complete list of CITF Projects IDs in Appendix 1.

Line 2: Laboratory/Facility name

Name of Laboratory or facility that performed the assay for e.g., Decaluwe lab/UdeM

Line 3: Laboratory/Facility zip code

Zip code of laboratory or facility that performed the assay e.g., H3T 1C5

Line 4: Specimen source

Select from one of the following SCT code:

- *Whole Blood (258580003)*
- *PBMCs (404798000)*
 - *Fresh*
 - *Frozen*

Line 5: Participant ID

Each participant ID must be a unique value, also used to identify the participant in the other Core data elements tables. It is recommended, however not required, to use HL7 naming rules. You can read more about Naming rules for H7L here: <https://www.hl7.org/implement/standards/>

Sections map:

1. Core Data Elements Explained

- ELISpot
- Flow Cytometry

2. List of preferred terms

More details on samples that do not pass Quality control can be added to the line 9, Comment.

Sections blocks:

Lines 10 to 18 can be repeated multiple times to accommodate reporting from multiple sequential assays, and multiplex conformations.

See next page for details.

Line 6: Specimen ID

It is possible that a single participant has several sample collections; for example, in longitudinal studies, where samples are collected at different time point. A single specimen ID should be generated for each collection, with the option of using HL7 naming rules. Often the specimen ID is reflected in the tube/sample barcode stated as Numerical (Unconstrained numeric string)

Line 7: Specimen collection date

Date of sample collection, stated as YYYY[MM[DD]].

Line 8: Overall description

Collective descriptive assessment of the individual readings of the assay. Select from one of the following SCT code

- *Positive (10828004)*
- *Negative (260385009)*
- *Inconclusive (419984006)*

Line 9: Comments

Space for additional Error codes or notes not covered in the other lines (e.g., batch ID). For error code reporting, please refer to the Unique concept Identifiers list provided in next section.

Line 10: CITF Assay ID

A unique number assigned to the different facilities for each assay can be requested to the CITF. A protocol assigned to the assay can be linked to this number. The list of the CITF assay ID can be accessed here. If your assay is not found in the list, please contact the CITF secretariate (project liaison) with assays details.

Line 11: Date assay run

Date when the assay was run, stated as YYYY[MM[DD]].

Line 12: Number of cells

The number of cells used for the assay stated in millions for e.g., 1E+6 cells

Line 13: Cytokine tested

When reporting for several cytokines, in order to harmonize result reporting and data analysis we recommend to use the following order #1: IFN- γ , #2: IL-2, #3: TNF ... (if more cytokines are further tested keep filling them out following the data schema)

Line 14: Peptide used

Specify manufacturer (or write 'home-made') and use preferred antigen nomenclature from Table II (Pg. 12).

Sections map:

1. Core Data Elements Explained

- ELISpot
- Flow Cytometry

2. List of preferred terms

Line 15: Result

Numerical value obtained for this sample in this assay. Mention results are Spot forming units (SFU)/1M cells/PBMCs (depending on the specimen type used for the assay)

Line 16: Cut-off

Value above which the samples is considered positive.

Line 17: Maximum

Highest measurable value for semiquantitative data or higher value of standard curve for quantitative data.

Line 18: Results description

Categorical description based on the numerical value obtained for a particular cytokine/peptide in the assay. Select from one of the following SCT codes:

- Positive (10828004)
- Negative (260385009)
- Inconclusive (419984006)
- Weakly positive (60408008)
- Borderline high (442777001)
- Borderline low (442779003)

Line 19: Pass Qc?

Results from internal quality control should be indicated here. Select from one of the following SCT code:

- Yes (373066001)
- No (373067005)

1.2. Reporting schemes or dataset structure

When reporting a simple assay (i.e., single peptide, ELISpot TNF), you need to use only the fields on lines 1-18 and ignore the remainder.

When reporting more complex assays (i.e., multiple peptides for one cytokine, or multiple cytokines for one peptide), use as many assay blocks (lines 10-19) as needed. The dataset can be in wide or long format. Note that if long format is used, CITF assay ID, Date run, and QC lines should be repeated.

Sections map:

1. Core Data Elements Explained

- ELISpot
- Flow Cytometry

2. List of preferred terms

Only one ID?

It is possible that for some projects only either the Participant ID or Specimen ID is known to the investigator.

In that case, please leave the field unknown as blank.

Data linkage tables:

Independently to the data entered here, it is recommended that a table should be generated by the principal investigator detailing how many collections were done for each participant and their date, linking Participant ID, Specimen ID and Specimen collection date (lines 5-7).

Flow cytometry-based assays

1.3. Line-by-line explanations

Each line is assigned a Number and a Label. See explanation for each one below.

Line 1: Project ID

Unique ID assigned to each project and connected to funding documents. See complete list of CITF Projects IDs in Appendix 1.

Line 2: Laboratory/Facility name

Name of Laboratory or facility that performed the assay e.g., Decaluwe lab/UdeM

Line 3: Laboratory/Facility zip code

Zip code of laboratory or facility that performed the assay e.g., H3T 1C5

Line 4: Specimen source

Select from one of the following SCT code:

- Whole Blood (258580003)
- PBMCs (404798000)
 - Fresh
 - Frozen

Line 5: Participant ID

Each participant ID must be a unique value, also used to identify the participant in the other Core data elements tables. It is recommended, however not required, to follow HL7 naming rules. You can read more about Naming rules for H7L here: <https://www.hl7.org/implement/standards/>

Line 6: Specimen ID

It is possible that a single participant has several sample collections; for example, in longitudinal studies, where samples are collected at different time point. A single specimen ID should be generated for each collection, with the option of using HL7 naming rules. Often the specimen ID is reflected in the tube/sample barcode.

Line 7: Specimen collection date

Date of sample collection, stated as YYYY[MM[DD]].

Line 8: CITF Assay ID

A unique number assigned to the different facilities for each assay can be requested to the CITF. A protocol assigned to the assay can be linked to this number. The list of the CITF assay ID can be accessed here. If your assay is not found in the list, please contact the CITF secretariate (project liaison) with assays details.

Line 9: Date assay run

Date when the assay was run, stated as YYYY[MM[DD]].

Sections map:

1. Core Data Elements Explained

- ELISpot
- Flow Cytometry

2. List of preferred terms

More details on samples that do not pass Quality control can be added to the line 13, Comment.

Sections blocks:

Lines 19-21 can be repeated multiple times to accommodate reporting several intracellular cytokines (max 3, recommended order #1: IFN- γ , #2: IL-2, #3: TNF.

Line 10: Stimuli (Peptide used or unstimulated)

Specify manufacturer (or write 'home-made') and use preferred antigen nomenclature from Table II (Page 12).

Line 11: Overall description

Select from one of the following SCT code:

- Positive (10828004)
- Negative (260385009)
- Inconclusive (419984006)

Line 12: Pass Qc?

Results from internal quality control should be indicated here.

Select from one of the following SCT code:

- > Yes (373066001)
- No (373067005)

Line 13: Comments

Space for additional Error codes or notes not covered in the other lines (e.g., batch ID). For error code reporting, please refer to the Unique concept Identifiers list provided in next section.

Core Results (lines 14-17)

Line 14: Total number of cells

Add total number of cells, or gated events in the live gate (in millions).

Line 15: % CD3+ cells

Enter percentage of cells from live gate that are CD3+.

Line 16: % CD3+/CD4+ cells

Enter percentage of cells from the CD3+ gate that are CD4+.

Line 17: % CD3+/CD8+ cells

Enter percentage of cells from the CD3+ gate that are CD8+.

Intracellular staining (ICS) assay blocks (lines 18-27)

Line 18: ICS assay?

Select from one of the following SCT code:

- No (373067005)
- Yes (373066001)- If so, complete lines 19-21 as many times as needed

Line 19: Cytokine tested

When reporting for several cytokines, in order to harmonize result reporting and data analysis we recommend to use the following order #1: IFN- γ , #2: IL-2, #3: TNF, #4+: Specify the names in data table.

Line 20: % CD3+/CD4+ cells

Enter percentage of cells from the CD3+/CD4+ gates that are positive for that cytokine.

Sections map:

1. Core Data Elements Explained

- ELISpot
- Flow Cytometry

2. List of preferred terms

Because of high result variability across studies and consequentially low harmonization potential, AIM assay data will not be harmonized across studies. However, it is still recommended to follow the format described.

If this CDE form is used to report

AIMs without Memory phenotypes:

mark Yes in line 28, No in line 29, leave blank line 30 (memory population), and complete only lines 31 to 34.

Sections blocks:

Lines 30-34 can be repeated 4 times to accommodate the 4 different memory populations

Line 21: % CD3+/CD8+ cells

Enter percentage of cells from the CD3+/CD8+ gates that are positive for that cytokine.

Lines 22-27:

Repeat 19-21 as needed for multiple cytokines

Activation induced marker (AIM) assay and memory phenotype blocks (lines 28 and up)

Line 28: AIM assay?

Select from one of the following SCT code:

- No (373067005)
- Yes (373066001)- If so, complete lines 30 and up as needed

Line 29: Memory phenotype?

Select from one of the following SCT code:

- No (373067005)
- Yes (373066001)- If so, complete lines 30 and up as needed

Two option for results are suggested: style 1 (activation of the different memory subsets), or style 2 (memory phenotype of the activated cells). For both types of results cells are first gated in CD3+, then on CD4+ or CD8+.

- For style 1, cells are then gated using memory markers (CCR7 and CD45RA), outlining the memory population mentioned below in line 30, and finally in each population the percentage of activated cells are extracted (using activation markers for CD4+ or CD8+ populations, like CD134+/CD137+ or CD69+/CD137+, respectively.

Results Style 1- Activation of different memory subsets (lines 30-49). Repeat blocks as needed (max 4 blocks).

Line 30: Memory population

Enter markers or population name for each of the 4 memory populations: **Naïve** (Tn) (CCR7+/CD45RA+), **Effector Memory** (Tem) (CCR7-/CD45RA-), **CD45RA+ Effector Memory** (Temra) (CCR7-/CD45RA+), or **Central Memory** (Tcm) (CCR7+/CD45RA-).

Line 31: Activation markers for CD3+/CD4+

Enter activation markers used for CD3+ and CD4+ cells, for example: CD134+/CD137+.

Line 32: % cells

Enter percentage of activated cells for each memory population (cells from the CD3+/CD4+ gates that had that particular memory phenotype from line 30 and the activation marker from line 31).

Sections map:

1. Core Data Elements Explained

- ELISpot
- Flow Cytometry

2. List of preferred terms

Sections blocks:

Lines 50-55 can be repeated 4 times to accommodate the 4 different memory populations

- **Line 33: Activation markers for CD3+/CD8+**
- Enter activation markers used for CD3+ and CD8+ cells, for example: CD69+/CD137+.
- **Line 34: % cells**
- Enter percentage of activated cells for each memory population (cells from the CD3+/CD8+ gates that had that particular memory phenotype from line 30 and the activation marker from line 33).
- *For style 2, cells are then gated using activation markers (for example CD134+/CD137+ for CD4+ cells and CD69+/CD137+ for CD8+ cells), and finally the percentage of the different memory subtypes of the activated cells is recorded using memory markers (CCR7 and CD45RA), outlining the memory population mentioned in line 51.*

Results Style 2- Memory phenotype of activated cells (lines 50-73). Repeat blocks as needed (max 4 blocks).

Line 50: Activation markers for CD3+/CD4+

Enter markers used for CD3+ and CD4+ cells, for example: CD134+/CD137+.

Line 51: Memory population

Enter markers or population name for each of the 4 memory populations: **Naïve** (Tn) (CCR7+/CD45RA+), **Effector Memory** (Tem) (CCR7-/CD45RA-), **CD45RA+ Effector Memory** (Temra) (CCR7-/CD45RA+), or **Central Memory** (Tcm) (CCR7+/CD45RA-).

Line 52: % cells

Enter percentage of activated cells for each memory population (cells from the activated CD3+/CD4+ gates that had that particular memory phenotype, sum of all four phenotypes or blocks should be 100%).

Line 53: Activation markers for CD3+/CD8+

Enter markers used for CD3+ and CD8+ cells, for example: CD69+/CD137+

Line 54: Memory population

Enter markers or population name for each of the 4 memory populations as on line 51.

Line 55: % cells

Enter percentage of activated cells for each memory population (cells from the activated CD3+/CD8+ gates that had that particular memory phenotype, sum of all four phenotypes or blocks should be 100%).

Sections map:

1. Core Data Elements Explained

- ELISpot
- Flow Cytometry

2. List of preferred terms

Sections blocks:

Lines 50-55 can be repeated 4 times to accommodate the 4 different memory populations

1.4. Reporting schemes

Investigators could choose between different reporting schemes, from very simple (only % of CD4+ and CD8+ cells, filling up to line 17), adding intracellular staining (lines 18-27), and/or adding AIMs without Memory phenotypes (lines 28-34).

If they choose to add memory phenotypes, depending on the gating and reporting styles proposed, they can add fields (wide format) or block of rows (long format) corresponding to lines 30-49 or 50-73.

Sections map:

1. Core Data Elements Explained

- ELISpot
- Flow Cytometry

2. List of preferred terms

What does SCT stand for?

The **S** stands for SNOMED or Systematized Nomenclature of Medicine,

and **CT** stands for Clinical Terms

Section 2- List of preferred terms

For the categorical selections use the SCT codes provided in the Concept Unique Identifier (CUI) Table I in the next section. CUI are unique numerical codes assigned for each concept, based on SCT codes. For the description of the antigen used in your assay, if possible, use the terms defined in table II.

2.1. Preferred SCT codes

If you are unable to use the SCT code provided, please replace them with the content text from the table below. In the case that this is not possible, please submit with your data a single flat file that maps the terminology used to the coding below.

Table I- Concept Unique Identifier for categorical selections.

Line	Label	Content	SCT Code
4	Specimen source	Whole blood	258580003
		PBMCs (fresh/frozen)	404798000
8	Overall description	Positive	10828004
14	Results description	Negative	260385009
		Inconclusive	419984006
15	Pass QC?	Yes	373066001
		No	373067005
<i>* If No, it is possible to add more information on line 9-Comments, see below.</i>			
9	Comments	Weakly positive	60408008
		Borderline high	442777001
		Borderline low	442779003
		Details for QC fail:	
		Specimen unsatisfactory for evaluation	125154007
		Insufficient sample	281268007
		Sample contaminated	123840003
		Sample decomposed	733480003
		Sample unsuitable for this assay	281273001
		Improper storage of sample	712743008

Sections map:

1. Core Data Elements Explained

- ELISpot
- Flow Cytometry

2. List of preferred terms

2.2. Preferred antigen nomenclature

In order to harmonize result reporting and perform data analysis, we suggest using the following nomenclature for the antigens used in your assay. This name is also recorded in the table “CITF Assay List.” When a new assay for a new antigen is added to the CITF Assay list (see Appendix that points to Assay list), the new antigen will be added to the list below.

If this is not possible, please submit with your data a single flat file that maps the terminology used to the coding below.

Table II- Preferred antigen nomenclature. Also included as part of the CITF assay ID

Nomenclature	Manufacturer	Virus	Antigen
JPT-SARS-CoV-2-Full	JPT	SARS-CoV-2	full virus
MLTBIO-SARS-CoV-2-S	JPT	SARS-CoV-2	Spike
JPT-SARS-CoV-2-S1	JPT	SARS-CoV-2	S1
JPT-SARS-CoV-2-S1_NTD	JPT	SARS-CoV-2	S1 NTD
JPT-SARS-CoV-2-S1_RBD	JPT	SARS-CoV-2	S1 RBD
JPT-SARS-CoV-2-S1/S2	JPT	SARS-CoV-2	S1/S2
JPT-SARS-CoV-2-S2	JPT	SARS-CoV-2	S2
JPT-SARS-CoV-2-N	JPT	SARS-CoV-2	Nucleocapsid
MLTBIO-SARS-CoV-2-Full	Miltenyi Biotec	SARS-CoV-2	full virus
MLTBIO-SARS-CoV-2-S	Miltenyi Biotec	SARS-CoV-2	Spike
MLTBIO-SARS-CoV-2-S1	Miltenyi Biotec	SARS-CoV-2	S1
MLTBIO-SARS-CoV-2-S1_NTD	Miltenyi Biotec	SARS-CoV-2	S1 NTD
MLTBIO-SARS-CoV-2-S1_RBD	Miltenyi Biotec	SARS-CoV-2	S1 RBD
MLTBIO-SARS-CoV-2-S1/S2	Miltenyi Biotec	SARS-CoV-2	S1/S2
MLTBIO--SARS-CoV-2-S2	Miltenyi Biotec	SARS-CoV-2	S2
MLTBIO-SARS-CoV-2-N	Miltenyi Biotec	SARS-CoV-2	Nucleocapsid

Nomenclature for variants for which assays are available - either JPT or Miltenyi.

Sections map:

- 1. Core Data Elements Explained**
- 2. List of preferred terms**
- 3. CITF Assay List**

Appendix 1

Section 3- CITF Assay List

The CITF Assay List tab in the Excel document contains all the information related to the different assays. In it, both commercial as in-house developed assays can be found.

The CITF Assay catalogue provides coding for many data elements: Antibody and antigen tested (including lots when available), assay units, device or equipment Identifier, LOINC test codes, SCT test description, SCT specimen source. When available, it also provides links to Assay Protocols.

Each assay is given a unique name and number. Assays with the same name but having changes in the number can reflect, for example, changes in equipment, protocols, or reagents lots.