# CDC Blood Donation Serosurvey Data Guide - WARPSURVEY

Your participation in the CDC National Blood Donor Serosurvey Study, part of Operation Warp Speed, is greatly appreciated. This study represents a huge collaborative effort to capture data from a large number of sources for rapid analysis and reporting to the CDC. This document serves as a guide to Blood Collection Organizations (BCOs) for submitting data and should be read in conjunction with the codebooks and example data files. In order to integrate and analyze these data on very tight timelines, it is critical that you strictly follow the agreed process and comply with the codebooks. Please do not hesitate to contact the study team with any problems or questions.

You will submit these primary data reports:

1. **Tested Sample Report:** The line listing report for samples tested. Every row in the line listing report corresponds to one (1) SAMPLEID. Testing results fields to include in the tested sample reports, and when the submissions are due, differ between data submission groups and are detailed in section II.
2. **Requisition Forms for Confirmatory Testing:** After samples are screened on Vitros, requisition forms will be provided by VRI for samples requiring further confirmatory testing. Complete the form and submit to VRI prior to shipping samples to VRI for further confirmatory testing.
3. **Donation Summary Report:** A frequency summary report of all donations at the collection site for the month. This report is due within the first 10 business days of every month.

*Please refer to the corresponding sheets in the codebook for details on the fields in each of these reports.* [1]

This guide consists of six sections:

- **I.** Conditions for data reporting
- **II.** Data submission groups
- **III.** Requisition of data and samples for confirmatory testing
- **IV.** Details on fields
- **V.** File naming convention for data files
- **VI.** Points of contact for questions

## I. Conditions for Submission

Data submissions must adhere to the codebook and follow the conditions listed below or will be rejected with a request for correction and resubmission.

1. Remove all pre-approved Convalescent COVID-19 Plasma (CCP) donations, of any procedure type, for the **Tested Sample Report** and **Donation Summary Report**
2. There should be no missing or null values for any variable. As specified in the codebook, the code for "blank/unavailable" should be used when applicable (e.g., use "9" for blank/unavailable sex of donor).
3. Data files must be named as per naming conventions in Section V (five).
4. Submissions must contain all fields listed in the codebook, in adherence to the data submission groups' specifications, with no other variables.

[1] Codebook's filename: "*CDC Serosurvey Codebook 20201120 v2.0*"

5. Submissions must have variables with the correct data type (e.g. “numeric”, “character”, “date”, etc.) as designated in the codebook.
6. Values for every variable must be within the ranges (numeric or date) or contain only the codes (categorical) specified in the codebook.
7. Submission files must be in comma separated values (.csv) format.
8. Data Summary Reports are submitted to Westat. Tested Sample Reports are submitted via VRI’s SFTP. Login credentials for VRI’s SFTP and instructions are provided separately.
9. Monthly Summary data files are submitted to Westat through a WinSCP tool to the secure file transfer protocol site (SFTP). Only files named according to the file naming convention will be validated. The system will perform a data validation check, such as formats, missing values, consistency in headers, etc. An automatically generated email will be sent to the data submitter letting them know whether the files have been accepted. If the files fail one or more validation checks, the data submitter will be directed to an error log detailing the errors. It is critical that data submitters address errors and re-upload files within 1 business day in order to facilitate rapid turnaround of analysis and datasets for CDC.

## II. Data Submission Groups

Depending on a BCO’s universal testing policy and assay used, BCOs are binned into the data submission groups below to ensure proper validation is done on the submitted data files.

| Data Group | CENTERIDs | Initial Submission | | | | | Subsequent Submission | | | | Sample Requisition |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Metadata | Vitros Results (Both Fields) | Roche Results | | Submission Frequency | Metadata | Roche Results | | Submission Frequency | |
| | | | | S/CO | Interpretation | | | S/CO | Interpretation | | |
| 1 | OB, CBC, ARC | ✓ | ✓ | ✓ | ✓ | Monthly | | | | N/A | |
| 2 | BSSM, BBOH, VTL | ✓ | ✓ | | | Monthly | ✓ | ✓ | ✓ | Monthly | |
| 3 | TBC, GCBC | ✓ | | ✓ | ✓ | Per Shipment | | | | N/A | |
| | VST | ✓ | | | ✓ | Per Shipment | | | | N/A | |
| 4 | BCKC, BBOA, BWNW, LSV | ✓ | | | | Per Shipment | | | | N/A | |
| 5 | NYBC, RIBC, BBDA, LSO | ✓ | ✓ | | | Monthly | | | | N/A | ✓ |

**Table *1.*** Data submission groups and the BCOs included in each of the group. Fields to include and when the reports are due for the initial and subsequent tested sample report submissions are indicated. BCOs in data submission group **5** will receive a sample requisition form with every submission and are expected to return a filled form before samples are sent to VRI San Francisco

**Submission Group 1: OneBlood, Carter Blood Center, and the American Red Cross**

BCOs in submission group 1 will only submit one tested sample report, due within the first 10 business days of the month. The tested sample report should include metadata, Ortho-Vitros Total Ig S1 results, and Roche Elecsys Total Ig NC results. ARC and OB are expected to submit Roche results for all Vitros reactives, using the explicit null values for Roche testing to fill in blanks in the report, while CBC is expected to submit Roche results for all samples included in the report. Use of explicit nulls for Roche testing in CBC’s submissions will result in rejection.

**Submission Group 2: Banco de Sangre Servicios Mutuos, Blood Bank of Hawaii, and Vitalant**
BCOs in submission group 2 will submit two tested sample report for each sampling month. The initial tested sample report submission (screening report) will be due within the first 10 business days of the month and includes all of the metadata fields and the two Vitros testing results fields, while the subsequent submission should include all of the metadata fields and the two Roche testing results fields. The subsequent submission is due when Roche results become available.

**Submission Group 3: The Blood Center, Gulf Coast Blood Center, Versiti**
BCOs in submission group 3 will only submit one tested sample report, due on every sample shipment to VRI. The tested sample report submission should include metadata as well as both of the Roche testing result fields. The use of explicit nulls for Roche testing results is permitted.

**Submission Group 4: Blood Center of Kansas City, Blood Bank of Alaska, BloodWorks Northwest, LifeServe**
BCOs in submission group 4 will only submit one tested sample report, due on every sample shipment to VRI. The tested sample report submission should include metadata for every sample shipped to VRI San Francisco.

**Submission Group 5: New York Blood Center Enterprise and LifeSouth**
BCOs in submission group 5 will only submit one tested sample report, due monthly within the first 10 business days of the month. When a submission passes a validation, a sample requisition form will be sent out to the appropriate persons at the BCO. Details on sample requisition can be found in section III.

## III. Requisition of Samples for Confirmatory Testing

BCOs in submission group 5 are asked to complete the requisition forms provided by indicating which samples are included in the shipment, its aliquot type, box position, and return the completed form to a directory on the SFTP ([username]/warpsurvey/upload/requisition) prior to shipping samples to VRI for further confirmatory testing. Details on the requisition form can be found in the codebook and examples of an empty and filled out requisition form is included in the data package[3]

For questions and/or concerns on the requisition process, contact Lois Fisher or Hasan Sulaeman at VRI. (Contact info provided in Section VI.)

[3] EXAMPLE_BCO_202010_20201119_requisition.csv, EXAMPLE_BCO_202010_20201119_requisition_form.csv

## IV. Details on Fields

**SAMPLEID**
If your BCO received sample labels from VRI, please use the provided SAMPLEID for samples. Otherwise, you may assign a unique SAMPLEID in the format specified in the codebook

**SEX**
Sex of donor. BCOs with sex data that is neither M nor F are asked to report any other entries as "9"

**CENTERID**
Study assigned abbreviation for participating BCOs

**FIN**
The FIN (ISBT facility identification number) field is for the first 4-digits of the FIN, not including the first letter. (e.g. W0410 is listed as 0410)

**DONDATE**
Donation date is written in YYYYMMDD format. (e.g. 23 Mar 2020 is listed as 20200323)

**AGE**
Age of donor at the time of donation in the proper range (16-105)

**RACETH**
Self-reported race and ethnicity of the donor. As different BCOs document race and ethnicity differently, participating BCOs should fit entries into the given RACETH (Race and Ethnicity) codes

**ZIP5**
5-digit zip code of donor residence. 99999 denotes blank/unavailable. ZIP5 should have leading zeros, where applicable

**SCRNHIST**
Donor screening history – whether the donor is a first-time donor or a repeat donor. Note that donors with no prior screening results are considered first-time donors.

**SITETYPE**
Donation site type – whether the donation is from a fixed location or from a mobile drive

**DONTYPE**
Donation type. Most collections are of voluntary allogeneic/homologous donations, but it is important to capture when the donation is autologous (for the donor's own use), directed (designated for use by a particular recipient), research collections, etc.

**DONPROC**
Donation procedure. Whole blood donation or one of various kinds of apheresis procedure.

**ABO_RH**
Blood type of donor denoted by the ABO blood group letter and (+/-) signs. (e.g., A positive is listed as A+)

**VACC**
Donor vaccination status. The VACC field specifically denotes whether the donor have been vaccinated for COVID-19 and, if applicable, the timing of the vaccination.

For BCOs collecting vaccination data from more than 8 weeks ago, the possible entries are Not Available (**NAV**), Vaccinated Recently (**VRE**), Vaccinated (ever; **VEV**), Never Vaccinated (**NVE**).

For BCOs collecting vaccination data only from 8 weeks prior, the possible entries are Not Available (**NAV**), Vaccinated Recently (**VRE**), and Not Vaccinated Recently (**NVR**).

For BCOs not colleting vaccination data or for BCOs collecting vaccination data but the data is not extractable, the field should still be included but with Not Available (**NAV**) as an entry on every sample.

## Note on Test Result Fields

The following fields are only applicable to the tested sample report. Refer to Section II (two) for more info.

**VITROS_COV2T_SCO**
S/CO values for samples tested using the Ortho Vitros CoV2T assay

**VITROS_COV2T_INTERPRETATION**
interpretation of the Vitros CoV2T assay. Omit field if no data is available

**ELECSYS_COV2T_SCO**
S/CO values for samples tested using the Roche CoV2T assay

**ELECSYS_COV2T_INTERPRETATION**
Interpretation for the Roche CoV2T assay.

## Note on explicit nulls

Some fields in the codebook have the number “9” as a possible entry. The number “9” is an example of an entry to be used as an explicit null. The use of explicit nulls ensures data points left blank during reporting are due to data being unavailable and not due to errors in reporting or data entry. Other examples of explicit nulls can be found in the codebook provided.

## V. File Naming Convention

Filenames use the variables below. File naming convention stays the same for resubmissions.
Center Abbreviation: CENTERID blood collection organization abbreviation
Year Month: Year and month of the data report. Formatted as YYYYMM (e.g. 202007 for July 2020)
Date of Creation: Year, month and day of file creation. Formatted as YYYYMMDD

**Tested Sample Reports:**
[Center Abbreviation]_[Year Month]_[Date of Creation]_linelist
**Examples:**
VTL_202003_20200730_linelist.csv

**Requisition of Vitros Low Reactive samples:**
[Center Abbreviation]_[Year Month]_[Date of Creation]_requisition_form (blank form)
[Center Abbreviation]_[Year Month]_[Date of Creation]_requisition (completed form)
**Examples:**
ARC_202011_20201119_requisition_form.csv
ARC_202011_20201119_requisition.csv

**Donation Summary Report:**
It is the same file name format as the Tested Sample Report, without "linelist"
**Example:** ARC_202003_20200730.csv

## VI. Contact List

**Hasan Sulaeman, MSc**
Data Manager | Vitalant Research Institute
hsulaeman@vitalant.org

**Eduard Grebe, PhD**
Investigator | Vitalant Research Institute
egrebe@vitalant.org

**Howard King**
Data Manager | Westat
howardking@westat.com

**Matt Levy, PhD**
Project Manager | Westat
matthewlevy@westat.com

**Mars Stone, PhD**
Co-Principal Investigator | Vitalant Research Institute
mstone@vitalant.org