Disclaimer:

"The content of this Deliverable D1.1 represents the views of the author only and is his/her sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains."

D1.1 Defined dataset for each step in the search process available on webpage

Grant Agreement number: 837354 **Project acronym:** SAVDON **Work Package number:** WP1

Organisation: World Marrow Donor Association (WMDA)

LEAR: Esther Pustjens
Project coordinator: Lydia Foeken

Tel: 0031 88 505 7900

E-mail: lydia.foeken@wmda.info

Organisation website address: www.wmda.info



Co-funded by the Health Programme of the European Union



Table of Contents

1.	. Data submission guidelines (XML)	2
2.		
۷.		
	2.1 Search and Match Registry-to-Registry Communication survey	
	2.2 Search & Match API requirements	4
	2.3 Data Dictionary	7
	2.3.1 Data Dictionary RFC Template	7
	2.3.2 Launch of new XSD version 2.2	11
3.	. Search & Match Service Data submission information	13
	XML file	14
	Minimal required data	26
	XML example files	27
	File names	29
	ZIP (compress) file	30
	Encrypt file	31
	Data Validations: Errors and warnings in processing report	35
	Cord blood banks IDs	39
	Deprecated Code User Guide	40
	WMDA forms	42



1. Data submission guidelines (XML)

In order to improve the data collection, WMDA had to implement changes to their procedures for data reception and data processing.

The file format is an XML (Extensible Markup Language) file, which is considered an industry standard that is extendable, robust and easy to use.

This Search & Match Service Data submission information (preliminary search, CT request and workup request) on WMDA Share is publicly available. It can be accessed here:

Search & Match Service Data submission information

The outline of these pages is described in **Chapter 3** - Search & Match Service Data submission information.



2. Background

2.1 Search and Match Registry-to-Registry Communication survey

In February 2019, WMDA send out a <u>survey</u> to the members of the Search, Match & Connect User Group – consisting of users worldwide who regularly make use of WMDA's Search, Match & Connect service – to investigate which data they believe is necessary for an optimal search result. The homepage of this Google survey is shown in *Figure 1*.

Figure 1. Homepage Search and Match Registry-to-Registry Communication survey

Registry to Registry Communication

Dear member of Search, Match & Connect User Group,

WMDA is starting up the third phase of development of Search, Match & Connect (BMDW).

In the first phase BMDW upgraded the matching algorithm and implemented an user friendly Search & Match Service. In the second phase BMDW extended the dataset to provide more information to search coordinators.

In the third phase WMDA aims to reduce the administrative work of a search coordinator by developing tools. For example, a search coordinator registers the HLA typing once and does not have to type the HLA typing several times. Goal is to use the EMDIS data dictionary and EMDIS semantics to achieve this.

Below you will find a questionnaire to investigate how you are currently working as search coordinator. We will explain the survey to you during the telephone conference today. The deadline of filling in the questionnaire is: Monday, March 4 by the end of the day.

Looking forward to speak with you, on behalf of

Alicia Louw, South Africa BMDR Malgorzata Urban, Anthony Nolan Kim Tohm, NMDP Mark Melchers, WMDA

*Vereist

Your email address *

Jouw antwoord

3





2.2 Search & Match API requirements

#	User stories (Pivotal chores)	MoSCoW	Description	State (Pivotal)	Pivotal ticket #	Estimated Developer hours
1	As a search coordinator I want to be able to add a new patient to the WMDA Search & Match platform using API client specification through my internal system.	Must	Given I am in my Internal Search System. Then I should be able to add a patient to the WMDA Search & Match platform (WSMS) from my internal system. This enables patients that are created in the internal system to be automatically created in WSMS.	Finished	165211608	30
2	As a search coordinator I want to be able to update existing patient details at the WMDA Search & Match platform using API client specification through my internal system.	Must	Given I am in my Internal Search System. Changes include: birth date, ABO, CMV HLA	Finished	165211646	16
3	As a Search coordinator I want to be able to update the patient status at WMDA Search & Match platform using API client specification through my internal system. This should affect the behaviour of donor/cbu searches and its refreshing in a way comparable to EMDIS.	Must	Given I am in my Internal Search System Then I should be able to change the status of the patient to the same statuses as in EMDIS. Behaviour should mirror that of EMDIS: NEW: initial state. Cannot be changed into this state from any other state. PRE: preliminary. Triggers a one-time search in Search & Match. No automatic refreshing of search results. ACT: activated. Search is performed using default settings. Automatic refreshing of results is performed nightly. SUS: suspended. Patient remains visible. Search results are NOT discarded Automatic refreshing of search results is NOT performed STP: stopped. Patient gets deactivated.	Started	165212330	20
			o SUS → STP New SUS No searches are performed even configured searches historical search results are available			





4	As a Search coordinator I want to be able to see results of my international donor search in my internal system	Must	Given I am in my Internal Search System Then I should be able to see the returned results of the international donor search that I run I should be able to see all donors identified for my patient; results should be based on the same search performed in WMDA Search & Match Service (WSMS)	Unstarted	165213707	26
5	As a Search coordinator I want to be able to see results of my international CBU search in my internal system	Must	Given I am in my Internal Search System Then I should be able to see the returned results of the international CBU search that I run I should be able to see all CBUs identified for my patient; results should be based on the same search performed in WMDA Search & Match Service (WSMS)	Unstarted	165213763	26
6	As a Search Coordinator I want to be able to see when my international search results are available.	Must	Given I am in my Internal Search System. Then I should be able to see that my international search is finished so that I can retrieve those results.	Started	165213820	28
7	As a Search Coordinator I want to be able to view the returned international donor and CBU results in detail in my internal system	Must	Given I am in my Internal Search System When the international results are received Then I should be able to view those results in detail. The detailed information ("full donor report"/ "full cord report") only needs to be with the most recent info. It does not need to be historic data, even for historic search results.	Unstarted	165213867	24
8	As a user I want to be able to test WMDA Search & Match API endpoints through a sandbox UI provided by Swagger (open standard for describing REST APIs with JSON and XML). This should give clear insight into how the API responds to parameters and options for developers and non-developers.	Should	In this case the users are actually the IT persons at the registry who are building the connection to the WMDA Search & Match API. In order to understand and test the API and its responses, there needs to be a possibility to test the API.	Finished	165211483	24
9	As a user I want to be able to authenticate with Search & Match API using Oauth2.0 server and authorization client through my internal system. For auditing purposes and practical use all activities should be marked with the userId of the user using the internal system.	Should	In this case the users are actually the IT persons at the registry who are building the connection to the WMDA Search & Match API. The majority of technical survey respondents voted for OAauth 2.0 as the authentication method. Laravel upgrade is advised. It is important that search coordinators do NOT need to use their WSMS credentials to be able to use their internal system. Authentication is done at Internal-system-to-WMDA-API level. User logs into internal system of organisation using their internal credentials. API passes email address of user during OAuth 2.0 authentication. Needs to be e-mail address of a valid WSMS user for this organisation. no Front end for client key management.	Finished	165211545	26
10	As a Search Coordinator I want to be able to view a list of all searches performed for my patient in my internal system	Should	Given I am in my Internal Search System Then I should be able to view a list of all the searches run for my patient — Both local (internal) and international donor and CBU searches	Started	165218556	18
11	As a Search coordinator I want to be able to rely on the availability of the Search & Match API for 99.99% of the time	Should	Search coordinators and organisations should be able to rely on the fact that the WMDA Search & Match Service is "always" available.	Unstarted	165218674	42
12	As a Search Coordinator I want to be notified when my international search results are available	Could	Given I am in my Internal Search System Once my international search was run I should receive a notification of received results e-mail? POST to API endpoint at organisation	Unstarted	165219456	4
13	As an admin of the internal system I should be able to delete a patient from the WMDA systems	Could	Necessary for GDPR reasons (right to be forgotten). Delete may obfuscate personal information but keep the record.	Unstarted	165219540	18
14	As a Search coordinator I want to be able to request alternative phenotype search via my internal system	Would	Given I am in my Internal Search System Then I should be able to see the results of the alternative phenotype search that I run "Non-patient search" I should be able to see all alternative phenotype donors and CBUs identified for my patient; results should be based on the same search performed in WMDA Search and March Service (WSMS)	Unstarted	165219594	26





15	stateless webservice	Could		Unstarted		
16	front end for client_id to organization id/ ION OAuth management	Could	When adding new organisations to the whitelist of allowed API clients, this needs to be managed through an .env file. This does not scale. Therefore when the pilots have been finished and more and more organisations are connecting it might be wise to have an easy to use UI to manage this whitelist	Unstarted		24
17	change detection of search results	Could		Unstarted		16
18	automated testing of the platform	Would	built-in testing would make the testing of the entire WMDA Search, Match and Connect platform very much quicker. Future development cycles would become a lot shorter. Bug and	Unstarted		
19	Audit trail	Should	As a WMDA administrator I want to be able to check the audit trail of the actions performed by users. This can be in a log file.	Unstarted		
20	Start search	Must	Given I am in my Internal Search System Then I should be able to start a new search for my patient. This search is either repeatable or not and can have various (matchengine specific) search settings	Started	#166268061	40

The results of this survey will be used to build the new API* connector. Search coordinators can obtain the same information in their local system through this API connector as they can obtain when they register a search online through the web application. This API connector will be developed as secure, robust, simple and scalable, that allows that registries to continue to use their current operational registry software application.

*API:

Application Programming Interface (see: https://youtu.be/s7wmiS2mSXY)

An application programming interface is a set of clearly defined methods of communication among various components. It facilitates the communication between computer programs. Those programs do no need to be in the same physical location. An API simplifies programming by abstracting the underlying implementation and only exposing objects or actions the developer needs. In this case the APIs on the side of WMDA enables secure and simple communication between it and the internal software of various registries.



2.3 Data Dictionary

The <u>data dictionary</u> (DD) was created to improve data exchange and communication between all organisations involved, and to ensure harmonization of data and field definitions across all domains: EMDIS, WMDA systems (Search & Match, Patient API, Connect), WMDA Form, etc.

The aims of the Data Dictionary group(s) are to ensure that:

- Data is defined by experts with relevant operational knowledge of the field to bridge the gap between operational applications and technical use
- Proposed changes of clinical or research value to patient searches and outcomes
- That all communication channels (EMDIS, Search & Match service, WMDA Forms and future systems) use a common language at user level wherever possible.

2.3.1 Data Dictionary RFC Template

The Data Dictionary Group has been established to process change requests (RFCs) that allow the community to propose changes, improvements, or additions to the data standards. The recommendations of this group are reflected in a set of domains: WMDA Data Upload (XML), WMDA Forms, EMDIS.

In March 2019, this group has created a procedure to draft an RFC (Request For Change). This Data Dictionary RFC Template (including short track) is accessible for all WMA Share members.



Quick User Guide

- Please contact any of the DD group members when you start to draft a RFC (Request For Change). DD group will assign a document owner to help you and be in charge of your request.
- The RFC template below includes full and short track templates.
- Requester only needs to fill rows that marked with "Fill by" = "REQ" and "Required" = "Yes". Welcome to fill more if you know more about your request.
- Short track template is used to request a change in validation and/or allowed new values for existing data elements/fields. The column "Shot track" identifies the rows need to be filled for a short track request.
- Full template should be used if a **NEW** data element or field is requested.
- \bullet $\$ $\$ ToDo is a column that you can identify the things still need to be discussed.

RFC#	RFC-
Document status	
Document Owner	
Request Registry	ION-
Request Date	
Target release	

	Short track	Fill by	Required	Element	Value (to be filled by requester)	Describes	Checks	ToDo
1	х	REQ	yes	RFC Title		Descriptive title of the RFC	Unique	
2	х	System	yes	RFC Number		Sequential number of the RFC within the DDG		
3				Related RFC		Reference to related RFCs (number)	should be an existing RFC	
4	x	REQ	yes	Requesting party		Group , body or individual that issued the RFC	Offical group	
5		REQ	yes	Supporting parties		Minimal 2 other parties (Organisations or Working groups)		
6	х	REQ	yes	Represented by		Person responsible for the RFC from the requesting party; main contact for questions.	Existing person	
7	x	DD	yes	Document owner (DD group)		Person responsible for the RFC from the Data Dictionary group"	Part of DD group	
8	х	DD	yes	Primary domain		In what domain the need has emerged (Forms, Emdis, WMDA, XSD, API)	Existing domain	
9	х	REQ		Data cluster		In what group of elements the request would fit?	Existing data cluster / family in the Data Dictionary	
10	х	REQ		Date of Request		YYYY-MM-DD	Valid date	
11	x (go/no go or similar)			Status of Request		Is it in preparation - consultation - elaboration - community feedback - recommendation -decision -final rejected "Further stakeholder engagement needed" is part of the consultation status.	Valid status	
12				Date of recommendation DDG		What date was the recommendation done (YYYY-MM-DD)		
13	х			Date decision		YYYY-MM-DD		



				USER STORY			
14	x (short, should already includes all parts below + reason for shorttrack + cost benefit analysis)	REQ		User story / Background / Strategic fit	Describe the user and what they are trying to achieve and why. How does this relate to the overall platform strategy? One user story per RFC.		
15				Rationale / Goals			
16				Clinical value	Describe who will benefit of having this element, add any additional documents, research or other supporting materials.		
17				Expected life span	How long will the element exist? Is it temporarily? Exit strategy?		
18				Expected fill degree	Expected Fill degree of all donors/patients.		
19				Relative cost to <i>get data</i> for delivering party.	We are talking the actual collection of the data, not the implementation (see block 4)		
20				Availability of resources to get data	Describe if there are special lab facilities or difficult available resources are needed to get the data.		
21	х			Expected outputs	Where will the value be shown, calculated, converted, aggregated, include in algorithm, transmitted, displayed or used?		,
				TECHNICAL			
22		REQ	yes	Suggested fieldname	What fieldnames have you been considering?		
23		REQ/DD	no	Proposed fieldtype	What type of field should it be (INT/BOOL/STRING/TEXT) etc. (stick to XML primitives)		
24	х	REQ/DD	no	Proposed validation	What validations should apply?		
25	x (normally short tracks do talk about changing validation and or allowed values)	REQ/DD	yes	Proposed values	What are the values that are acceptible?		
26	х	REQ/DD	yes	Required?	Is the data element required or optional, or only together with another field.		
27	х	REQ/DD	no	NULL handling	What if we have no data?		
28	x (small changes would	REQ/DD	no	Proposed (X)path	W		
29	х	DD	req	Domains affected			
30				Dictionary fit			
-						-	





31				Display methods	Is there any specific thing to be said about how you would display the value and what conversions would be needed. eg MAC/GL		
32				Similar field	Is there similar field that we can use as a template.		
33				Existing or possible duplicate, or overlapping	aboresus = abo + resus		
34				Field Naming consistency			
35	х	DD	no	Proposed versioning per domain	XSD version, EMDIS Implementation Package,		
36	x	REQ/DD	yes	Privacy considerations	Are there any privacy issues with this new information? Can it be used for de-anonymizing? Any abuse possible or foreseen? This needs to be reviewed by a qualified DPO. Please state the name of the DPO of the requesting party.		
				IMPLEMENTATION			
37	х			Will be implemented in Emdis?	Describe shortly what is the proposed implementation scope for this domain or application.	Link to corresponding EMDIS page	
38	x			Will be implemented in Optimatch?	Describe shortly what is the proposed implementation scope for this domain or application.	provide link to implementation plan for this domain.	
39	х			Will be implemented in Front end search application?	Describe shortly what is the proposed implementation scope for this domain or application.		
40	х			Will be implemented in Patient API?	Describe shortly what is the proposed implementation scope for this domain or application.		
41				Overall cost/benefit			
42	х			Questions	Open questions to be addressed		
43				References	E.g. References to scientific or other publications		



2.3.2 Launch of new XSD version 2.2

XML is the data format that is used for uploading data in the WMDA Search & Match Service (BMDW) and for inventory exchange between registries. At the moment, registries use XSD version 2.1. In October 2019, WMDA published the new XSD version 2.2. This was announced in the Connect / Up-To-Data Newsletter Issue September 2019.

From now, WMDA accepts XSD files from version 2.1 and version 2.2. This means that Search, Match & Connect user can continue using version 2.1, although WMDA strongly recommends upgrading to version 2.2.

Significant changes in version 2.2:

- DONOR/CBU: Grid/ID handling changed
- CBU: changes of bank manufacture ID and bank distribute ID
- CBU: changed and added CMV elements

Access the new XSD

XSD version 2.2 can be found HERE. The page is publicly accessible. The homepage looks as follows:

XML schema files version 2.2

Created by Riny Heijdendael, last modified by Zhihong He on Nov 12, 2019

The WMDA data-upload does not support version 2.2 yet. WMDA plans to support the new version in end November or December, 2019.
 We will send the notification in the SCM and Connect Newsletters when it is supported.

Download XSD version 2.2:

FINAL

The files below are officially approved by the Datadictionary Technical XML group.

All questions regarding this version can be asked here: https://share.wmda.info/questions/topics/341344265/xsdv22



Changes between version 2.1 and 2.2

- Download XSD version 2.2 :
 - Encoding should be included
 - Updated version number to 2.2
 - Updated GRID/ID rules
 - Changed Bank identifiers
 - Updated documentation on GRIE
 - Updated CMV elements (RFC-001S Distinguish different CMV/HIV tests (ANTI-vs. PCR)#/)





This page is currently under development. The WMDA data-upload does not support XSD version 2.2 yet. WMDA plans to support the new version at the end of November or December 2019. When this is sorted, a notification will be sent to the community in the WMDA newsletters Stem Cell Matters and Connect.





3. Search & Match Service Data submission information

The homepage looks as follows:

Pages / Search & Match Service Public Access / Data submission guidelines (XML) 🚡 🖉 ९ 🎶 Analytics

Search & Match Service Data submission information

Created by just a moment ago



1. Introduction

2. XSD schema files

- 2.1 InventoryType elements
- 2.2 ItemBaseType elements (for Donors and CBUs)
- 2.3 hlaType elements
- 2.4 kirType elements
- 2.5 idmType elements
- 2.6 donltemType elements
- 2.7 cbultemType elements
- 2.8 matType elements

The topics on this webpage include

- XML file
- Minimal required data
- XML examples
- File names
- ZIP (compress) file
- Encrypt file
- Data Validation: Errors and warnings in processing report
- Cord blood banks IDs
- Deprecated Code User Guide
- WMDA forms



XML file

1. Introduction

2. XSD schema files

- 2.1 InventoryType elements
- 2.2 ItemBaseType elements (for Donors and CBUs)
- · 2.3 hlaType elements
- 2.4 kirType elements
- 2.5 idmType elements
- · 2.6 donItemType elements
- · 2.7 cbultemType elements
- 2.8 matType elements

XML file format

1. Introduction

The overall scope of the development phase two is to receive more data from our listing organisations and to make these data available through our Search & Match Service. However, the old format (DOT20) is not an appropriate format when you have many different fields/columns. Therefore, we had to move to another file format. The new file format is an XML (Extensible Markup Language) file, which is considered an industry standard that is extendable, robust and easy to use.

Several people from the community formed a working group to create the required XML Schema Definition (XSD) files. These files define the elements that are allowed in the XML file, the order of the elements and the values that will be accepted. The names of the elements are based upon EMDIS specifications and aligns with the EMDIS Data Dictionary when appropriate. Several elements are basic elements that should be included in all files, but there are also elements that are specific for only donors or only cord blood units (CBUs).

We will now explain the composition of the XML file and how you should use the XSD reference files.

XSD schema files

We provided two XSD schema files that define the structure of your XML file: basicTypes.xsd and Inventories.xsd.

The Inventories file describes the structure of the XML file and the order of the elements. Here you can also find if a certain field is mandatory or not (minOccurs="0"-> not mandatory). This file includes many "complexTypes": an XML element that contains other elements and/or attributes. In the file you can see that the values of the elements can be defined here, like the elements GRID and ID, or that after the name of the field a "type" is defined. For example for the element with name BIRTH_DATE you see type="bareDateType". The definition of "bareDateType" is described in the basicTypes.xsd file.

We will now describe the global structure of the XML file and the elements.

Please note: For a lot of elements, we use abbreviations as allowed values. The explanation of all those abbreviations can be found in the XSD files. Most abbreviations are also the same as used for EMDIS and clarified in the EMDIS dictionary.

2.1 InventoryType elements

Field Identifier	Required	Description	Туре	Length	Comment
CREATION_TIME	Yes	Creation time stamp of the inventories (in UTC)	dateTime	minimal 20	Without fractional seconds the length is 20, for example: 2016-08-23T13:16:48Z. Additional notes: CREATION_TIME is defined as "Creation time stamp of the <inventories>" that means the time in UTC when the complete and valid file was finally created at the registry. This can be the same as SNAPSHOT_TIME.</inventories>

14



Field Identifier	Required	Description	Туре	Length	Comment
LISTING_ORGANIZATION	Yes	Organisation that lists the donor/cbu provided as ION	ionType: number between 1000 and 9999	4	Issuing Organisation Number (ION) allocated by ICBBA. This can be different from the POOL when another organisation is sending the data to WMDA.
POOL	Yes	Physical location of the donors/CBUs of the inventory provided as ION	ionType: number between 1000 and 9999	4	Physical location of the donors/CBUs of the inventory provided as ION.
CONTENT_TYPE	Yes	Type of the inventory items, i.e. donor ("D") or CBU ("C")	contentTypeType	1	The content-type is also shown in the fileName. When CONTENT_TYPE is "D", the INVENTORY must contain <donor>-blocks. When CONTENT_TYPE is "C", the INVENTORY must contain <cbu>-blocks.</cbu></donor>
UPDATE_MODE	Yes	Update mode of the inventory, i.e. FULL or DIFF	updateModeType	4	Only UPDATE_MODE "FULL" is currently supported. Always the complete inventory should be send.
SNAPSHOT_TIME	No	Timestamp of the 'data snapshot' (in UTC)	dateTime	minimal 20	Without fractional seconds the length is 20, for example: 2016-08-23T13:16:48Z Additional notes: SNAPSHOT_TIME in the element <inventory> is defined as "timestamp of the data snapshot in UTC" that means the timestamp of the creation of this part of the complete file. This can be the timestamp of the XML export and I guess that in most of the cases it will be identical to the CREATION_TIME.</inventory>
SCHEMA_VERSION	Yes	Version of the applied XML Schema Definition (XSD)	schemaVersionType		The schema version is very important as this determines the validation rules that should be applied during the processing of your file.

2.2 ItemBaseType elements (for Donors and CBUs)

Field Identifier	Required	Description	Туре	Length	Comment
ID	Yes	Unique identifier of the donor/CBU	String	17	Unique identifier of the donor/CBU: If you are an EMDIS member, you can use the same ID as you use for that system (EMDIS hub code + donor identification allocated by the associated donor registry).
					For non-EMDIS members we recommend to use two digit ISO country code of the associated donor registry + donor identification allocated by the associated donor registry. For example: AU600196166, DEGOE-35487, US087013165, SB45. However, you are also allowed to use just the donor ID allocated by your registry.
GRID	No	Global registration identifier of the donor	String	19	ONLY applicable for donors. GRID format allowed is: XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX



Field Identifier	Required	Description	Type	Length	Comment
ATTR	No	Describing attribute of the donor/CBU according to house rules of the sending organization.	String	3	
BIRTH_DATE	Yes	Date of birth of the donor/CBU	bareDateType	10	Date without timezone information, example 1988- 06-28, Date Delimiter = "-"
SEX	No	Biological gender of the donor/CBU	sexType	1	sexType: "F","M" F = Female M = Male NOTE: Mandatory for donors, optional for CBUs
ABO	No	Blood group (ABO) of the donor/CBU	aboType	2	aboType: "A","B","O","AB"
RHESUS	No	Rhesus (Rh) factor of the donor/CBU	rhesusType	1	rhesusType: "P","N" P = Positive N = Negative NOTE: "+" and "-" are not supported
ETHN	No	Ethnic group of the donor/CBU	ethnType	4	ethnType: "AFNA", "AFSS", "ASSW", "ASSO", "ASCE", "ASSE", "ASNE", "ASOC", "CAEU", "CAER", "CANA", "CAAU", "HICA", "HISA", "AF", "AS", "CA", "HI", "MX", "OT", "UK" AFNA = African: North Africa
CCR5	No	CCR5 status of the donor/CBU	ccr5Type	2	ccr5Type: "DD","WW","DW" DD = Deletion (delta 32) - homozygous DW = Deletion (delta 32) / wildtype - heterozygous WW = Wildtype - homozygous
HLA	Yes	HLA of the donor/cbu	hlaType		Explained separately at hlaType 2.3
KIR	No	KIR genotype of the donor/CBU	kirType		Explained separately at kirType 2.4
IDM	No	Infectious disease markers (IDM) and other relevant tests of the donor/CBU	idmType		Explained separately at idmType 2.5



Field Identifier	Required	Description	Туре	Length	Comment
RSV_PAT	No	Unique identifier of the patient the donor/CBU is reserved for (STATUS=RS).	String	17	The value comprises the EMDIS patient identification, where the patient search centre is an EMDIS member, otherwise the value is empty. For example: AU9654021, DE275342, US2277450. NOTE: This field is not required for status "RS" and can be transmitted as empty if privacy concerns exist.
STATUS	Yes	Status of the donor/CBU	statusType	2	statusType: "AV" ,"TU" ,"RS" ("DE" is not supported yet) AV = Available for transplantation purposes TU = Temporarily unavailable RS = Reserved DE = Deleted, permanently unavailable
STAT_END_DATE	No	Date until which the current status will be applicable	bareDateType	10	Date without timezone information, example 1968- 06-28, Date Delimiter = "-"

2.3 hlaType elements

HlaType fields can be divided in hlaSerFieldsType and hlaDnaFieldsType

hlaSerFieldsType: HLA values obtained by serological typing methods

hlaSerFieldsType = "<FIELD1>" string of max length 5 "</FIELD1>", "<FIELD2>" string of max length 5 "</FIELD2>";

Example: <SER><FIELD1>1</FIELD1><FIELD2>5</FIELD2></SER>

Serological typing results can be given for loci that are defined as hlaLocusType. These loci include HLA-A, -B, -C, -DRB1, -DQB1.

hlaDnaFieldsType: HLA values obtained by DNA based typing methods

hlaDnaFieldsType = "<FIELD1>" string of max length 20 "</FIELD1>", "<FIELD2>" string of max length 20 "</FIELD2>";

Example: <DNA><FIELD1>01:01</FIELD1><FIELD2>05:01</FIELD2></DNA>

DNA typing results can be given for loci that are defined as hlaLocusType and hlaLocusDnaOnlyType. These loci include HLA-A, -B, -C, -DRB1, -DQB1, -DRB3, -DRB4, -DRB5, -DQA1, -DPA1, -DPB1.

Finally, previously the dot20 file format allowed to submit values like 01 in DNA fields. We can no longer accept this and you have to submit the equivalent of 01, so '01:XX'.

Minimal required elements

Minimal typing values for Donor: A (either SER or DNA), B (either SER or DNA)

Minimal typing values for CBU: A (either SER or DNA), B (either SER or DNA), DRB1 (either SER or DNA)

Please note:

- · It is no longer possible to submit string HLA values; only single values are allowed.
- When a donor or CBU has homozygous alleles/values, please use the following notation:

<HLA><A><SER><FIELD1>1
/FIELD1><FIELD2 /></SER> ...
or

<DQB1><DNA><FIELD1>05:02:01G</FIELD1><FIELD2 /></DNA></DQB1>

Field Identifier	Required	Description	Туре	Length	Comment
SER	depends on content type and DNA fields provided	HLA values obtained by serological typing methods	hlaSerFieldsType	5	Each SER element contains two other elements: FIELD1 and FIELD2
DNA	depends on content type and SER fields provided	HLA values obtained by DNA based typing methods	hlaDnaFieldsType	20	Each DNA element contains two other elements: FIELD1 and FIELD2
FIELD1		HLA value of allele 1		5 or 20	Element within the element SER and DNA



Field Identifier	Required	Description	Туре	Length	Comment
FIELD2		HLA value of allele 2		5 or 20	Element within the element SER and DNA
A	Yes	HLA-A values	hlaLocusType		Both SER and DNA possible; either SER or DNA values required
В	Yes	HLA-B values	hlaLocusType		Both SER and DNA possible; either SER or DNA values required
С	No	HLA-C values	hlaLocusType		Both SER and DNA possible
DRB1	Yes (CBU) No (Donor)	HLA-DRB1 values	hlaLocusType		Both SER and DNA possible; either SER or DNA values required for CBU
DRB3	No	HLA-DRB3 values	hlaLocusDnaOnlyType		Only DNA possible
DRB4	No	HLA-DRB4 values	hlaLocusDnaOnlyType		Only DNA possible
DRB5	No	HLA-DRB5 values	hlaLocusDnaOnlyType		Only DNA possible
DQA1	No	HLA-DQA1 values	hlaLocusDnaOnlyType		Only DNA possible
DQB1	No	HLA-DQB1 values	hlaLocusType		Both SER and DNA possible
DPA1	No	HLA-DPA1 values	hlaLocusDnaOnlyType		Only DNA possible
DPB1	No	HLA-DPB1 values	hlaLocusDnaOnlyType		Only DNA possible

2.4 kirType elements

The kirType Field Definitions consists of the type: kirLocusType. This is defined as a String with 3 characters: "POS" or "NEG". "POS" means "Presence of KIR gene", "NEG" means "Absence of KIR gene".

The following elements are possible and in this specific order:

<KIR2DL1>,<KIR2DL2>,<KIR2DL3>,<KIR2DL4>,<KIR2DL5A>,<KIR2DL5B>,<KIR2DS1>,<KIR2DS2>,<KIR2DS3>,<KIR2DS4>,<KIR2DS5>,<KIR2DP1>,<KIR3DL1>,<KIR3DL2>,<KIR3DL3>,<KIR3DS1>,<KIR3DP1>.

There is another field called <KIR_GL> (URI that refers to a GL-string registered with a GL-service or direct GL-string for absence / presence) this field is not used at the moment and must be empty.

Field Identifier	Required	Description	Туре	Length	Comment
KIR gene e.g. KIR2DL1	No	KIR genotype e.g. KIR gene 2DL1	kirLocusType	3	valid values: "POS" = presence of KIR gene; "NEG" = absence of KIR gene

2.5 idmType elements

There are many infectious disease markers (IDM) possible in the element IDM. Many IDM elements can have either the values idmValueType or idmValueExtType

idmValueType includes the following values: "P","N"

idemValueExtType include the following values: "P", "G", "M", "B", "H", "O", "N"

Field Identifier	Required	Description	Туре	Length	Comment
---------------------	----------	-------------	------	--------	---------



Field Identifier	Required	Description	Туре	Length	Comment
CMV	No	CMV status	idmValueExtType	1	idmValueExtType: "P", "G", "M", "B", "H", "O", "N" P = IgG or IgM positive, test did not differentiate G = IgG positive, IgM negative M = IgG negative, IgM positive B = Both IgG and IgM positive H = IgG positive, IgM not tested O = IgG negative, IgM not tested N = Both IgG and IgM negative EMDIS data dictionary also has a 'Q' (questionable / unclear) but that will not be applicable within the data submission file.
CMV_NAT	No	CMV NAT status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
CMV_DATE	No	Date of CMV test	bareDateTyp	10	Date without timezone information, example 1968- 06-28, Date Delimiter = "-"
HBS_AG	No	Hepatitis B status (hepatitis B surface antigen)	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
ANTI_HBC	No	Hepatitis B status (antibody to hepatitis B core antigen)	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
ANTI_HBS	No	Hepatitis B status (antibody to hepatitis B surface antigen)	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
ANTI_HCV	No	Hepatitis C status (antibody to hepatitis C virus)	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
ANTI_HIV_12	No	Anti-HIV 1/2 status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
HIV_1_NAT	No	HIV-1 NAT status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
HIV_P24	No	HIV p24 status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
HCV_NAT	No	HCV NAT status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
ANTI_HTLV	No	Antibody to HTLV I/II	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
SYPHILIS	No	Syphilis status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
WNV	No	WNV status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
CHAGAS	No	Chagas status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
EBV	No	EBV status	idmValueExtType	1	idmValueExtType: "P","G","M","B","H","O","N" P = IgG or IgM positive, test did not differentiate G = IgG positive, IgM negative M = IgG negative, IgM positive B = Both IgG and IgM positive H = IgG positive, IgM not tested O = IgG negative, IgM not tested N = Both IgG and IgM negative EMDIS data dictionary also has a 'Q' (questionable / unclear) but that will not be applicable within the data submission file. Please leave blank for Q.



Field Identifier	Required	Description	Туре	Length	Comment
тохо	No	Toxoplasmosis status	idmValueExtType	1	idmValueExtType: "P","G","M","B","H","O","N" P = IgG or IgM positive, test did not differentiate G = IgG positive, IgM negative M = IgG negative, IgM positive B = Both IgG and IgM positive H = IgG positive, IgM not tested O = IgG negative, IgM not tested N = Both IgG and IgM negative EMDIS data dictionary also has a 'Q' (questionable / unclear) but that will not be applicable within the data submission file. Please leave blank for Q.
HBV_NAT	No	HBV NAT status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
PB19_NAT	No	ParvoB19 NAT status	idmValueType	1	idmValueType: "P","N" P = Positive N = Negative
ALT	No	Alanine aminotransferase status in units per litre	Short		Number, no decimals, minimal value is 1

donItemType elements

DonItemType elements contain elements that are specific for donors and not applicable for CBUs.

Field Identifier	Required	Description	Туре	Length	Comment
STAT_REASON	No	Additional information relevant to the donor status. Can only be used for "TU" status.	statReasonDonType	2	statReasonDonType: "DO", "DD", "MR", "PR", "TX", "MO", "UC", "OT", "TQ", "UK" DO = Donor is too old DD = Donor died MR = Medical reasons PR = Personal reasons TX = After transplantation MO = Donor has moved UC = Unable to contact donor OT = Other reasons TQ = Typing questionable UK = Unknown
CONTACT_DATE	No	Date of last confirmed contact - defined as the date of an active form of communication (e.g. a query about status, an address update, confirmation of their interest in donating) via any channel (e.g. email, mail, phone, website), post registration, from a donor to the registry. Any communication from the registry to the donor that does not lead to an activity of the donor suggesting his further interest in donation is explicitly excluded (e.g. annual mailing without reaction).	bareDateType	10	Date without timezone information, example 1968-06-28, Date Delimiter = "-"



Field Identifier	Required	Description	Туре	Length	Comment
CHECKUP_DATE	No	Date of the last medical checkup - defined as the date of a donor health assessment that indicates whether a donor is minimally suitable to be considered for donation, regardless if eligible for only one donation type, and includes questions about current medication and health issues (e.g. completion of a health screening questionnaire at Extended Typing or Verification Typing). The donor health assessment can be completed by any means (e.g. paper-based, online, phone). This does not require any physical examination of a donor.	bareDateType	10	Date without timezone information, example 1968-06-28, Date Delimiter = "-"
WEIGHT	No	Weight in kg Please note that current validation only allows [40-199] kg	Short		Number between 1 and 999, no decimals
HEIGHT	No	Height in cm Please note that current validation only allows [100-250] cm	Short		Number between 1 and 999, no decimals
NMBR_TRANS	No	Number of blood transfusions	Short		Number: zero or greater, no decimals
NMBR_PREG	No	Number of pregnancies	Short		Number: zero or greater, no decimals
NMBR_MARR	No	Number of marrow donations	Short		Number: zero or greater, no decimals
NMBR_PBSC	No	Number of PBSC donations	Short		Number: zero or greater, no decimals
COLL_TYPE	No	Collection type, i.e. the willingness of the donor to donate in a specific manner	String	1	collTypeType: "M", "P","B" M = Marrow P = PBSC B = Both PBSC & Marrow

2.7 cbultemType elements

CbultemType elements contain elements that are specific for CBUs and not applicable for donors.

Field Identifier	Required	Description	Туре	Length	Comment
STAT_REASON	No	Additional information relevant to the CBU status. Can only be used for "TU" status.	statReasonCbuType	2	statReasonCbuType: "QR","AD","CD","DS","XP","MR","OT","UK" Proposed reasons for Status TU: QR = Quarantined; AD = Administrative Proposed reasons for Status DE: CD = Cord Destroyed or Damaged; DS = Distributed for infusion; XP = ExpiredCD = Cord Destroyed or Damaged; MR = Medical reasons OT = Unavailable for other reasons; UK = Unknown
LOCAL_ID	No	Identification of CBU locally at the associated CBB	String	17	
BAG_ID	No	Identification as it appears on the bag. If more than one bag is available then this data attribute is not populated	String	17	



Field Identifier Required Description Type Length Comment BANK MANUF ID Unique identifier of String PLEASE NOTE: For the upload the fields BANK_MANUF_ID and BANK_DISTRIB_ID should be fulfilled with a new ID for the the CBB that manufactured the CBU. ID shown in corresponding cord blood banks (See column table in tab Cord "WO number at blood bank IDs share.wmda.info/display/WMDAREG/Database) and not with the EMDIS IDs. These IDs are important to allow WMDA to identify if the CBU is from an accredited bank which will be displayed within a search report. BANK DISTRIB ID Unique identifier of 10 PLEASE NOTE: For the upload to WMDA the String the ĊBB fields BANK_MANUF_ID and distributing the CBU. ID shown in BANK DISTRIB ID should be fulfilled with a new ID for the corresponding cord blood banks table in tab Cord (See column "WO number at blood bank IDs share.wmda.info/display/WMDAREG/Database) and not with the EMDIS IDs. These IDs are important to allow WMDA to identify if the CBÚ is from an accredited bank which will be displayed within a search report. COLL DATE No Date that the CBU bareDateType 10 Date without timezone information, example was collected 1968-06-28, Date Delimiter = "-" PROC_DATE No Date that the 10 Date without timezone information, example bareDateType 1968-06-28 Date Delimiter = "-" processing started PROC_METH No Processing procMethType 2 procMethType: "HES","DGS","CEN","FIL","FIC","PER","OTH" method used HES = Hydroxy-Ethyl-Starch DGS = Density Gradient Separation CEN = Centrifuge FIL = Filtration FIC = FICOL PER = PERCOL OTH = Other NOTE: Values "NOT" and "UNK" are not supported "NOT" can now be found in CB_PROD_MOD = "NOT", "UNK" has to be transmitted as empty (CB_PROD_MOD = "") procMethTypeType: "MAN", "SPX", "OTP", "AXP", "OTH" PROC_METH_TYPE Nο Processing procMethTypeType method type used MAN = Manual SPX = Sepax OTP = Optipress II AXP = AXP OTH = Other FREEZE DATE No Date that the CBU bareDateType 10 Date without timezone information, example 1968-06-28, Date Delimiter = "-" was frozen freezeMethType: "C", "M" FREEZE METH No Freezing method freezeMethType used C = Controlled Rate M = Manual PROD_MOD No Product prodModType prodModType: "BCE", "DNE", "PLR", "PRR", "RBR", "NOT", "OTH" modifications made BCE = Buffy Coat Enriched DNE = Density Enriched PLR = Plasma Reduced (Volume reduction only) PRR = Plasma and RBC Reduced RBR = RBC Reduced (depletion) NOT = Not reduced OTH = Other bbagTypeType BAG TYPE Νo Type of bag used 5 bagTypeType: "80/20","50/50","40/60","NS" (no (bag fractions / split) split unit)



Field Identifier	Required	Description	Туре	Length	Comment
BAGS	No	Number of bags for CBU sub units	Short		Number between 1 and 99, no decimals
BACT_CULT	No	Bacterial culture	cultValueType	1	cultValueType: "P","N","D" P = Positive N = Negative D = Not done
FUNG_CULT	No	Fungal culture	cultValueType	1	cultValueType: "P","N","D" P = Positive N = Negative D = Not done
HEMO_STATUS	No	Hemoglobinopathy screening status	hemoStatusType	2	hemoStatusType: "DN","DU","NS","CD","NC","DT","DD" DN = Screening done, normal results DU = Screening done, unusual findings NS = No screening done CD = Can be done at time of release NC = Cannot be done DT = Thalassemia DD = Drepanocytosis
VOL	No	Collected volume before processing (without additives) in ml	Short		Number between 10 and 400, no decimals
VOL_FRZN	No	Total volume frozen (post processing, prior to cryopreservation) in ml	Short		Number between 10 and 400, no decimals
TNC	No	Total number of nucleated cells (before processing)	Float		Number with decimals, minimum is 0.0E0, maximum is 999.9E7
TNC_FRZN	No	Total number of nucleated cells (post processing, prior to cryopreservation)	Float		Number with decimals, minimum is 0.0E0, maximum is 999.9E7
RED_BC_FRZN	No	Total number of nucleated red blood cells (post processing, prior to cryopreservation)	Float		Number with decimals: minimum is 0.0E0, maximum is 999.9E7
MNC_FRZN	No	Total Number of mononucleated cells (post processing, prior to cryopreservation)	Float		Number with decimals
CD34PC	No	Total number of CD34+ cells (before processing)	Float		Number with decimals
CD34PC_FRZN	No	Total number of CD34+ cells (post processing, prior to cryopreservation)	Float		Number with decimals
CFU_FRZN	No	Total count of colony forming units (post processing, prior to cryopreservation)	Float		Number with decimals, nunimum is 0.1E5, maxnum is 999.9E5
VIABILITY	No	Viability as percentage value	Short		Number between 0 and 100, no decimals



Field Identifier	Required	Description	Туре	Length	Comment
VIABILITY_DATE	No	Date that viability was tested	bareDateType	10	Date without timezone information, example 1968-06-28, Date Delimiter = "-"
VIABILITY_CELLS	No	Type of oells tested for viability	viabilityCellsType	6	viabilityCellsType: "TNC", "CD34PC", "CD45PC" NOTE: VIABILITY_CELLS = "CD34PC" corresponds to CB_VIABILITY_CELLS = "CD34" in EMDIScord. VIABILITY_CELLS = "CD45PC" corresponds to CB_VIABILITY_CELLS = "CD45" in EMDIScord.
VIABILITY_METHOD	No	Method used to calculate the viability	viabilityMethodType	2	viabilityMethodType: "7A","PI","TB","OT" 7A = 7AAD PI = Propidium Iodide TB = Trypan Blue OT = Other
ATT_SEG	No	Number of attached segments available	Short		Number between 0 and 99, no decimals
DNA_SMPL	No	DNA samples available?	Boolean		true,false
OTH_SMPL	No	Samples other than DNA available?	Boolean		true,false
CT_COMPLETE_DATE	No	Date of completion of confirmatory typing (CT)	bareDateType	10	Date without timezone information, example 1968-06-28, Date Delimiter = "-"
CT_SMPL_TYPE	No	Confirmatory typing (CT) sample type	ctSmplTypeType	2	ctSmplTypeType: "AS","WB","RC","FP","ED" AS = CBU Contiguous Attached Segment WB = Whole Blood Sample RC = Red Cell Fraction (pellet) FP = Blood Spotted Filter Paper ED = Extracted DNA
AL_RED_BC	No	Number of red cell fraction aliquots	Short		Number between 0 and 99, no decimals
AL_SER	No	Number of serum aliquots available	Short		Number between 0 and 99, no decimals
SER_QUANT	No	Total quantity of serum available in ml	Float		Number between 0.0 and 99.9, one decimal
AL_PLA	No	Number of plasma aliquots available	Short		Number between 0 and 99, no decimals
PLA_QUANT	No	Total quantity of plasma available in ml	Float		Number between 0.0 and 99.9, one decimal
MAT	No	Data of the mother of the infant associated with the CBU	matType		see further on this webpage matType

2.8 matType elements

The matType elements are a sub-element from the element CBU.

Field Identifier	Required	Description	Туре	Length	Comment
ID	No	Identification used to identify the maternal donor as assigned by the registry	String	15	





Field Identifier	Required	Description	Type	Length	Comment
ID_BANK	No	Identification used by associated CBU manufacturer to identify maternal detail	String	15	
HLA	No	HLA of the mother of the infant associated with the CBU	hlaType		see above in section 2.3 hlaType
IDM	No	Infectious disease markers (IDM) and other relevant tests of the mother of the CBU	idmType		see above in section 2.5 idmType
AL_SER	No	Number of serum aliquots available	short		Number between 0 and 99, no decimals
SER_QUANT	No	Total quantity of serum available in ml	Float		Number between 0.0 and 99.9, one decimal
AL_PLA	No	Number of plasma aliquots available	Short		Number between 0 and 99, no decimals
PLA_QUANT	No	Total quantity of plasma available in ml	Float		Number between 0.0 and 99.9, one decimal



Minimal required data

Minimal required data

Organisations providing donor or CBU data, should at least include the following elements with valid values. Without this data, the records will be rejected during the validation procedure.

A DONOR record should include:

- ID
- BIRTH_DATE
- SEX
- . HLA (including at least HLA-A (SER or DNA) and HLA-B (SER or DNA))
- STATUS

A CBU record should include:

- ID
- BIRTH_DATE
- HLA (including at least HLA-A (SER or DNA), HLA-B (SER or DNA) and HLA-DRB1 (SER or DNA))
- STATUS



XML example files

XML example files

We already provided you the XSD files, but these files do not show directly how an XML file with those definitions will look like. Therefore we created some example files: one for donors and one for CBUs.

Both files contain only 2 records, but in those two records almost all possible elements contain a value. It can help you to check the order of the elements in your own XML file. Please be aware that values like GRID are fictive and do not follow the rules for the check character. These two examples are based on the XSD files version 2.1.

Example donor file: ION-1234-D.xml

Example Donor file Expand source

Example CBU file: ION-1234-C.xml

Example CBU file Expand source

Validation of your XML files

After generating your XML files, you are advised to validate the generated XML file. You may use any XML tool that includes validation for that, for example the open source Qxmledit or XMLLINT.

You can use xmllint by invoking the program giving the XML file and the Schema file, and it will generate a report.

If the file validates, it will show:

xmllint --schema Inventories.xsd --huge --stream --sax --noout ION-1234-C.xml ION-1234-C.xml validates



If it does not validate, you will see what XSD rule is violated, together with the line where that happened eg:

```
410460: Schemas validity error : Element 'CCR5': [facet 'enumeration'] The value '' is not a
```

Read more on installing XMLLINT on Windows

Installing XMLLINT on Windows is a bit tricky, you will need the XSD files, and a set of binaries from http://xmlsoft.org/sources/win32/for 32bit system, or http://xmlsoft.org/sources/win32/64bit for 64bit system.

- Download the XSD files.
- 2. Download the required binary files.

You need the binary files from 4 .zip files, and download the latest version. They are iconv, libxml2, libxslt and zlib.

The names are as belows, some binary files may have suffix version number, that is fine, and you can keep it.

```
basicTypes.xsd
Inventories.xsd

iconv.exe
libcharset-1.dll
libiconv-2.dll
libxm12-2.dll
xm12-config
xmlcatalog.exe
xmllint.exe
zlib1.dll
```

- Create a new folder in C:\tools\libxml. (This is a suggestion from the owner). Or you can create folder "libxml" in C:\Windows\ or C:\Program Files as you prefer.
- 4. Move all binary files in the folder C:\tools\libxml
- Add "C:\tools\libxml" to the System environment variable PATH/Path.
- 6. Restart your system
- The xmllint cmds can be executed in the windows cmd or PowerShell (windows 10).

Please refer to your IT staff for installation if you can not handle this.

Viewing large files

XML files tend to be large. For quick visual inspection and search you may use Glogg (windows)

Generate XML file from excel data source

WMDA prepare a php script to generate XML file from excel/csv data source for small dataset.

If you need further help, please contact support@wmda.info.



File names

File names

Registries with data on stem cell donors and cord blood units should separate these two data sets and provide two files: one for stem cell donors, and one for cord blood units. Data of stem cell donors and cord blood units should not be combined in one file. In the filename the distinction between donor data and cord blood unit data is made clear.

The first part of the filename is "ION-" (without the quotes) followed by the ion number and either a "D-" for donors (without the quotes) or a "C-" (without the quotes) for cold blood units. This <ION> should be similar to the one provided in the field <POOL>. The extension of the file is ".XML". Using this naming convention the name of the Austrian cord blood registry is: ION-2614-C.XML and the name of the German donor registry is: ION-6939-D.XML

After encryption, the file should follow almost the same name convention as for the xml file name, but then xml is replaced to pgp. So the first part of the GPG-filename is identical to the XML-filename. The GPG software will either add a second extension ".PGP", or replace the ".XML" extension of the data file with the ".PGP" extension. As an example, the file name of an encrypted Austrian cord blood file would then be: "ION-2614-C.PGP" (without the quotes).

If you are listing organisation and are also sending data from other listing organisations (with ION), you can provide the inventory of different POOL IDs together in one file. However, you should not combine donor and CBU data together. For the file naming, please always use the ION from the organisation that is sending the data.



ZIP (compress) file

How to compress your XML file with ZIP

(i)

ZIP file is not a request anymore, even for big file. Because the encryption process will compress and encrypt the file.

When your file is larger then 200Mb, you have to compress your file by using ZIP. If your file is smaller, you are also encouraged to compress your file, because the time to upload your file will be reduced.

Please find below some methods to compress your file with ZIP when you are using any of the operating systems Windows, OS, or Linux/unix.

Please note: When you are using another method to compress your file, like tar, we cannot decompress your file during processing and we have to reject the file.

Creating a compressed zip file in Windows

- Click to highlight the file that you need to zip. Please note: WMDA can only accept your file when the zip file contains 1 file.
- Right-click the file and select Send to > Compressed (zipped) Folder.
- 3. Windows will create the zip file and will position the cursor where you can choose a unique name for the file.

It is also possible to first create your ZIP folder and then drag the file to your zip folder.

Creating a compressed zip file in OS X

- Open a new Finder window and navigate to the file.
- Click to highlight the file that you need to zip.
- Select File > Compress from the pull-down menu. Sometimes you can also click with your right mouse button on the file and use the guick menu.
- Finder will compress the selected file and will create the zip file with the same same as your file but with the extension .zip.

Creating a compressed zip file in Linux

- Open a terminal session and navigate to the location of your file
- 2. To view a listing of directory contents, enter the following command: Is
- Note the files to be zipped.
- Create the zip file by entering the command: zip {.zip-filename} {filename-to-compress} (e.g. zip ION-1234-D.zip ION-1234-D.xml)



Encrypt file

How to encrypt your XML file

WMDA will only accept pgp encrypted XML files for data upload. We will now describe how you can encrypt your xml file. If you have a very large file, you should first compress your file with ZIP before you encrypted your file. Please see the picture below for a schematic representation of the encryption en decryption process.

? Unknown Attachment

? Unknown Attachment

Encryption is performed by the organisation who is sending data to WMDA; decryption is performed by WMDA to be able to process and validate the data in your file.

For this encryption, you should use the public key.

This is the new public key: public key (download the file in this link)

You may also fetch the key from the link or the https://pool.sks-keyservers.net/pks/lookup? op=get&search=0xC44E0E7A736E374E

See more key details ..

Public key details: Key ID C44E0E7A736E374E

Fingerprint = B837 44C5 CE38 C102 65E3 6120 C44E 0E7A 736E 374

userid = "dataupload.wmda.info <dataupload@wmda.info>"

PGP Public key for XML

----BEGIN PGP PUBLIC KEY BLOCK-----

Version: GnuPG v2

mQENBFk/8/ABCACqTuKBmPUsrenEvvXkPZ69m5F7eYi8KP+Si2owf//mgHXti79X KLfNievwxmbuDRGDPYaaK3LfwX7h6vpXAWfCYsosSUUmEJYa1Km+R6/R4dFQCYCG xWLcHBVS1Zr/5ZpHWdRfYMN3V7JM4u6E2axi1rysZ0WHHBDvkNFeZk14wwkTuM01 WzxQ40I7ASo//NzZnDAV2MxqLXq/W7XMdhmOMc5DtTDSypQsj2ZH30uq+MjpfyhD 5iq7wwxtib2KLmbBKDNLTqmqjjb4MHQJ1Aw4Djp62WNjgozRY8eG6rmsu8/C5Ufa yiIN0tX/s/FjgfQZFcucFnqnOBASJ2GqWU6DABEBAAG0K2RhdGF1cGxvYWQud21k YS5pbmZvIDxkYXRhdXBsb2FkQHdtZGEuaW5mbz6JATkEEwEIACMFA1k/8/ACGwMH CwkIBwMCAQYVCAIJCgsEFgIDAQIeAQIXgAAKCRDETg56c243TmvIB/49b9e3QIZC j2gc1V/d038JkpL1NOEmf51cIc8o/oDfvunxQRqjDUOCuIsh7t2SVJ7rna5wpruH FYNEnHu1S1FQ1C0J5ChHkvMUfKTUoXDn1KM6yZJRrJxd8UQ0g1310qNZvg9Hy1Pz



Cw3GH+w1JeZJ4FbqCg2/SXLKmRTjTutW5QhmfNBSxIqUq5xdjtyNvFGgdK1BJNbA UR2FXL3BQRRo5tE+Ivsb3DFM4sW6k2PRM/8p8mXDKo736c1+M1dB/rfGkq2D0scN 2pqsFPW9YdFDsUpOxgJof1c9WFyrV18I7pNuQ4Px2+CoCeF+RwSTHOm/epUB59Wd 9bTebi0rFF/ouQENBFk/8/ABCADENObYdeHOxaZiwRhNLbnHgbgq52Qmv6pmfb9g LLWyIpaRUGY+uTP2aJmovpHGwRt3+KPheoY2OsTv5hvUPU75Wcoy57J/FzXfHnac ZuSLGfViJFa0neS8Wb5NIrADUDzcOqgkfNvP8hd5J6vYOnCEgaOuNf5/9H7at4sv xVn7RpS+gqvmCIDzsLNQj6XWAkVFeL6KVGlq6Jfxg9oTOflu1EewWadRY1KUwJVN 5+G65fMnz4hFsJ+pWRmTrkFty0ti9g1LXarTHBbJRfagLZBNHKwYGyiSvwn1YbWH zVif8JLkzfHEIgO7e5RO25GaIwg6DYdhIeuZfesrHaexLecBABEBAAGJAR8EGAEI AAkFAlk/8/ACGwwACgkOxE40enNuN057Pgf/UUBThPnf/4EoCO1nzvbIefwfBvAm N1ra5z2GLrv5cjxDsB06aFCaWM9tvEU+zcuqgDSjD+H34kG00D2vINQLNmvhfYxZ DxGFwgM1xTjPzH4/alp3tAmsDmJ1VQs0LWxeXdSby47K0g7X9DvPw+VAMJpJZotg rVyYhStEqvXF66SWtPKviq5jdguGrxD4vS/y079YAiBWfFcjghB/aycebICaA2GT APUbvh2qypyodijCmiELIto/RIznUP00P43X69ciZjb4n8+Ir9buWtGpgp8XmRa1 sgdDFVbeiSnzba9A03osA5aWsdYeEAIchw6fPhLTvmZ5gjgfrbwxBn+loQ== =zCpi ----END PGP PUBLIC KEY BLOCK-----

Please note: This key is different from the key that you used to encrypt your DOT20 file.

STEP 1: Encryption program

The DOT20 file also needed to be encrypted. The procedure is actually the same, but you now have to use the new public key.

Currently, there are several different software packages that you can use to encrypt and decrypt your files. It depends of course also of your operating system which programs you can use.

Here are some examples:

Windows: Kleopatra (download here), PGP Tool

OS (MAC): GPG Suite Linux/unix: GnuPG

GnuPG is a complete and free implementation of the OpenPGP standard as defined by RFC4880 (also known as PGP). GnuPG allows to encrypt and sign your data and communication, features a versatile key management system as well as access modules for all kinds of public key directories. GnuPG, also known as GPG, is a command line tool with features for easy integration with other applications. A wealth of frontend applications and libraries are available. GnuPG is Free Software (meaning that it respects your freedom). It can be freely used, modified and distributed under the terms of the GNU General Public License. For installation on your Linux/Unix machine, please visit the following page for HowTos: https://www.gnupg.org/documentation/howtos.html

STEP 2: Import the public key

After installing your preferred program, you have to import the public key.

- First download the public key to your computer. It is important that you save the file with extension asc (key.asc)
- Open your encryption program and look for something like import (PGP) key or import certificate. Click on this and then you have to upload the file with the WMDA public key and save the key.
- If you might get an error with importing the key, you can try to remove the text "Version: GnuPG v2" from the file and try to import the file again.

For Linux/Unix, importing of the key in your gpg keyring can be done by using the following command:



Alternatively, instead of saving this file and importing the key, you may look it up at a keyserver.

For example, if you use Kleopatra, use CTRL-SHIFT-I and search for 0xC44E0E7A736E374E

Getting the key: http://pool.sks-keyservers.net/pks/lookup?search=0xC44E0E7A736E374E

If you use gpg from the commandline, use gpg --recv-keys --keyserver pool.sks-keyservers.net --recv

STEP 3: Encrypt XML or ZIP file with the public key

Next step is to use the public key that you just imported into your encryption program to encrypt your XML or zipped XML file

- In your encryption program, go to the function called encrypt or encrypt files.
- A windows with all your files will open. Look up the file you would like to encrypt.
- Following the steps in your program and make sure you choose the public key to encrypt your file
- Some programs work together with your file exploring program like Finder for Mac or Explorer for windows. If this is
 the case, go to your file look-up program and look for your file.
- Select your file and click on the right mouse button. A quick menu will become visible. Look for something with
 encryption or GPG or with a MAC it is probably under Services. This depends on the program you installed on your
 computer. Click on that and follow the instructions on your screen.
- Make sure you choose the public key to encrypt your file.

For Linux/Unix, encryption of your file can be done by using the following command:

```
$ gpg --encrypt --recipient ID {filename to be encrypted}
```

where ID is replaced with that key's ID

The short version of the above command is:

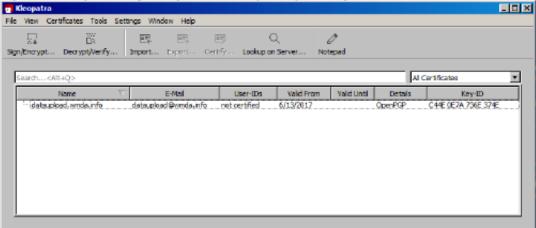
```
$ gpg -e -r ID {filename_to_be_encrypted}
```

In either case, a file is created with the same name, plus an additional .gpg file extension added to the end. Thus, if your file is ION-1234-D.xml, you will create an encrypted copy of the file named ION-1234-D.xml.gpg.

Please note: Do not sign your file. This will result in rejection of your file during processing.

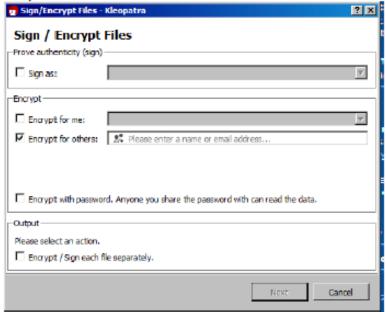
User guide for Kleopatra(Windows):

- 1. Download here. It is included in the gpg4win package.
- Import the the public key. And you will see imported public key as below:





Encrypt the file. Click "Sign/Encrypt", in the pop-up window, un-check "Sign as", "Encrypt for me" and "Encrypt with password".



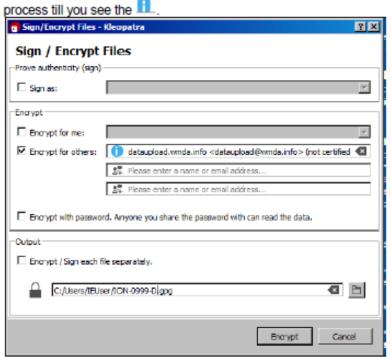
And in "Encrypt for others", type in the public key e-mails to get the public key.

Tip: if the WMDA public key is imported correctly, then after you type-in "data", the whole key name will pop-up, and can be selected.

NOTE: There will be pop-up window to notify you to certify the public key, you can ignore this. But if you do want to certify the public key, you can following the guide here.

If the public key is correctly imported, but without certify, there will be a blue "i" icon (!!) in front of the key. If you add certify, then it will be green "v" icon.

If the public key is not correctly imported, there will be a red "x" icon(). If that is the case, please repeat above



After that, the grep "next" button will turn "Encrypt", then click "Encrypt" to encrypt the file. You may get warning that you can not decrypt the file because you do not have private key, that is fine.



Data Validations: Errors and warnings in processing report

Explanation of Errors and Warnings in your processing report

(i)

INFORMATION

We are currently working on the new errors and warning you might find in your processing report. The errors and warnings shown below are mainly based on the old processing reports and are amended to our best knowledge at this moment. More information will follow soon.

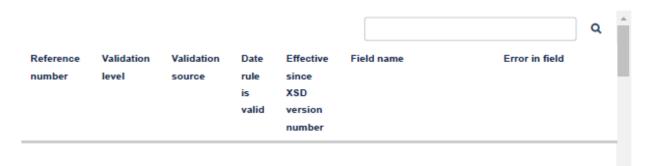
After submitting a file you will receive a processing report in your data upload service account. This document describes explanations of warnings or errors as you may find them in the processing report. In the explanations below you find references to various fields from the file format for data delivery. For details on these field names and the file format look in the tabs 'XML file' and 'File name' of this page.

The ID of the donor or CBU will be displayed as well. This helps you find the problem line in the file you have sent, and hopefully helps you correct the problem.

Business validation rules applied

WMDA has additional business validation rules in place to ensure that even though the data supplied on a field level might be correct they need to have passed the validation rules applied sometimes on multiple fields to ensure correct data is being added to the GCD database. As part of providing us the XML your organisation should also perform these checks to ensure the validity of the data you are providing.

Business validation rules in excel file



35



Reference number	Validation level	Validation source	Date rule is valid	Effective since XSD version number	Field name	Error in field
1	File	BMDW		2.0	File Name	Incorrect file naming format
2	File	BMDW		2.0	File Name	File is not encrypted
3	File	BMDW		2.0	File Name	File type is not XML
4	File	BMDW		2.0	CREATION_TIME	Date cannot be in the future
9	File	BMDW		2.0	UPDATE_MODE	Invalid update mode

Export V Records: 79 Size: 52.7 KB

File Errors

FileError: Empty file or file without data.

Explanation: When the size of the received file is zero bytes, or no data could be read from the received file, this error is returned.

Record Errors

- RecordError: No ID specified: Explanation: The ID or GRID field should be given, and should not be blank.
- RecordError: Duplicate donor/CBU ID found: Explanation: Every donor or cord blood unit should be listed only
 once. During processing of your file, a donor or cord blood ID is listed more than once, only the first occurrence has
 been accepted. The second occurrence XXXXX has been rejected.
- RecordError: Invalid GRID ID specified: Explanation: The global registration identified provided is invalid.
 PLEASE NOTE: GRID format allowed is: XXXX XXXX XXXX XXXX XXXX Also only upper case letter and numbers are allowed.
- RecordError: Duplicate GRID ID found: Explanation: Every donor or cord blood unit should be listed only once.
 During processing of your file, a GRID ID is listed more than once, only the first occurrence has been accepted.
 The second occurrence XXXXX has been rejected.
- RecordError: Invalid ION & GRID specified: Explanation: The GRID ID and Listing organisation (ION) ID are contradictory.
- RecordError: "Collected Volume before processing" (VOL) has an incorrect value: Explanation: The "Volume Collected" (VOL) should be no less than 10, or more than 400 milliliters (ml).
- RecordError: "Total Nucleated Cells" (TNC) has an incorrect value: Explanation: The "Total Nucleated Cells" should be no less than 10, or more than 999 (10^7).
- RecordError: "Collected number of CD34+ cells" (CD34PC) has an incorrect value: Explanation: The value provided for the CD34+ cell count is not a numeric value with decimal point in units of 10\(^6\).
- RecordError: "Collected number of mononuclear cells" (MNC_FRZN) has an incorrect value: Explanation: The value provided for the mononucleated cell count is rounded number of mononuclear cells in the



units of 10^7.

- RecordError: Duplicate donor/CBU ID found: Explanation: Every donor or cord blood unit should be listed only
 once. If during processing of a file, a donor or cord blood ID is listed more than once, only the first occurrence is
 accepted, the second occurrence generates this error.
- RecordError: Invalid date format for field 'field name' given: Explanation: This error may be returned for date
 fields. The format for dates field should be in the format YYYY-MM-DD.
- RecordError: Donor is either too young or too old: Explanation: The age limits of donors are set by the WMDA.
 Donor age outside range of 18-60 years are rejected.
- RecordError: Invalid Gender value found (not "M" or "F"): Explanation: Gender of donors other than "M" (for Male) or "F" (for Female) are reported as an error.
- RecordError: Invalid CMV Status value found: Explanation: The CMV status provided is not one of the possible
 values for this field.
- RecordError: BANK_MANUF_ID not provided: Explanation: The BANK_MANUF_ID allows the Search report to
 indicate that the Cord blood bank unit is accredited. If you do not provide this ID the search report will not be able
 to indicate the cord blood bank unit as being accredited or not.

HLA/DNA Related Errors

HLA/DNA-ERROR: Invalid allele value X found for DNA-"Allele".

Example: <HLA><DRB4><DNA><FIELD1>01:01/</FIELD1><FIELD2></DNA></DRB4></HLA>

Explanation: The value for DRB4 contains a slash (/) which is invalid. If such a problem is found the allele values are blanked and processing of the record continues. So, this is a warning, and not an error that requires the record to be rejected! Another examples might be an C*04:AVK (AVK is assigned bij the NMDP as 01/02/03/04/05/06) which is not valid since the C*04:02 is not assigned.

HLA/DNA-ERROR: Invalid DNA string found: "some string"

Example: For HLAA, the following value is given: 01:01/01:02/01:03/01:04.

Explanation: The DNA string given A*01:01/01:02/01:03/01:04 is not valid since the A*01:04 does not exist. Another reported problem may be that a ambiguity in the format of A*01:01/02/03 is invalid.

HLA/DNA-ERROR: Invalid HLA antigen "some value" found for field "field name".

Example: Serological HLA A30/3 is given in the file.

Explanation: The antigen or search determinant value "30/3" for HLA-A in this example is invalid.

HLA/DNA-ERROR: Incomplete typing found: HLA-A and HLA-B are required.

Explanation: At least one HLA-A and -B antigen or serological value should be available to be able to match the record. If no HLA-A or -B values (either on DNA level or on serological level) are available the record is rejected.

HLA/DNA-ERROR: "DNA allele values" does not match "serology". Equivalents for DNA alleles are ...

Example: HLA-B*15:02,15:26N does not match serological HLA-B76,62. Equivalents for the DNA-B alleles are: 75(15)

Explanation: The serology and DNA values provided, are validated separately, but also matched. If there is no match between the provided serology and the provided DNA, the record is reported.

HLA/DNA-ERROR: Number of alleles for DRB3/4/5 is more than 2; DRB3/4/5 blanked.

Example: Values are given for DRB4 for FIELD 1 and FIELD 2 and also for DRB5 for FIELD 1.

Explanation: Only two allele values are allowed for DRB3, DRB4 and DRB5 combined.



HLA/DNA-ERROR: DRB3 (or DRB4 or DRB5) does not match HLA-DR "values". DRB3 (or DRB4 or DRB5) blanked.

Example: Serology DR is 4 and 11; DRB5 is 01:XX

Explanation: DRB5 is associated with DR2, DR15(2), DR16(2) or DR1(rare). Likewise, DRB3 is associated with DR3, DR5, DR6, DR11(5), DR12(5), DR13(6), DR14(6), DR1403, DR1404, DR17(3), DR18(3); and DRB4 with DR4, DR7, DR9.



Cord blood banks IDs

Cord blood bank IDs

In the XML file format we also defined two fields that are referring to cord blood banks: BANK_MANUF_ID and BANK_DISTRIB_ID.

- The BANK_MANUF_ID is the ID corresponding to the cord blood bank that manufactured the cord blood unit (CBU)
- The BANK_DISTRIB_ID is the ID corresponding to the cord blood bank that will distribute the CBU

Please find in the table below the ID you have to use for your cord blood banks. The number always consist of 4 digits and is <u>not</u> the same as your ION number. We expect that the BANK_MANUF_ID and the BANK_DISTRIB_ID will be the same for many CBUs.

How to find your ID?

- Go to WMDA Share: the following webpage: Database
- Find the name of your cord blood bank on this page
 - If you can find the name of your cord blood bank, click on the name of your cord blood bank en you will see
 on the left side in the corner a number (for example: WO-1346); you can use the four digits of this number in
 the file you are uploading
 - If you can not find the name of your cord blood bank, contact the WMDA office at: support@wmda.info

OR

Contate the WMDA Office at: support@wmda.info

PLEASE NOTE: EMDIS is using very similar fields, but here you need to use a different ID.



Deprecated Code User Guide

Why deprecated code?

WMDA has developed HLA nomenclature guidelines, which have been published in 2007 and 2012. The HLA typing of your donors and cord blood units is validated based on these guidelines. If you list donors/cord blood units, it may happen that you receive the message that the HLA of your donor or cord blood unit is invalid.

There are two forms of 'deleted' alleles:

- Alleles that have been renamed (e.g. A*01:34N; sequence known to be expressed at low levels and renamed to A*01:01:38L)
- Alleles that have been removed without a replacement (e.g. A*2401; sequence shown to contain errors)

You can find information on: https://www.ebi.ac.uk/ipd/imgt/hla/docs/version_r3350.html

What you see in the report?

For data exchange between your organisation and WMDA, it is acceptable to use the older designation for renamed alleles for a period of one year after the renaming. During this grace period, you will receive the following message in your processing report (see below example messages):

#Warning at locus B*: renamed HLA code B*47:01:01:01 is still in grace period, New: 47:01:01:03

And if the deprecated code already passed one year grace period time, you will receive the following message in your processing report:

#Warning at locus B*: deprecated HLA code A*23:ABEBP is now invalid and passed its grace period hence blanked, New: 23:AFVGH 23:AZVXU

What to do with the deprecated code warning?

If you receive such warning messages, you have to update the HLA allele or MAC code within one year. WMDA will not mention the final date of the grace period in the processing report. You can find this information on: http://hla.alleles.org/alleles/deleted.html.

WMDA will provide you advice on how to replace the code. It may be that there are several options possible to update the allele or MAC code. WMDA recommends to perform the replacement using rules as following:

- · Always pick the same replacement code for a given deprecated code
- A possible sophisticated picking strategy is to use the code with the shortest definition string.
 For example: if you can choose between 15:01 and 15:01N, you pick the latter. The fewer items without trailing expression character in the definition string the better. Prefer generic codes over specific codes.
- Generally, definition strings containing items with trailing Q should be avoided (if possible) as alleles with expression character Q are in danger of being renamed.
 For example: in case of B*13:08Q you pick: B*13:08.



If you do not know, go for simple alphanumeric sorting

For 2018 the following alleles were deleted:

- April 2018: B*47:01:01:01 replaced by B*47:01:01:03
- July 2018: B*56:55:01:01 replaced by B*56:55:01:02
- August 2018: A*02:17:01 replaced by A*02:17:02

2019 Update

With the recent HLA Nomenclature Release 3.35.0, the allele HLA-C*03:99 has been renamed to HLA-C*01:169. While renaming of alleles usually does not affect the first two fields of an allele name, this does and as a consequence affects the HLA haplotype frequencies (HF). In fact, we use this allele for the HF sets of DE (1 haplotype) and East Asia (eas, 15 haplotypes). The affected haplotypes have a cumulated frequency of around 1 copy for each of the two HF sets.

We have one abroad patient with that HLA-C*03:99 in the Search&Match database for whom a HLA-C mismatched donor was requested in 2017.

Summary:

- The allele is generally very rare.
- Therefore the effect on patients can be assessed as negligible.
- ZKRD will replace the allele in the HF sets before it expires.

In case you need any support, we recommend to contact the WMDA at: support@wmda.info.



WMDA forms

The page that is linked to above is https://wmda.info/professionals/optimising-search-match-

WMDA forms

WMDA has developed template forms that organisations can use to cover the initial search process (preliminary search, CT request and workup request) through the delivery of stem cell at the transplant facility. Find all forms available on this page.

connect/wmda-forms/. The page includes the following forms:

Search/ HLA-typing

S10	Preliminary Search Request
S20	Request for Extended Donor HLA Typing
S30	Donor Extended Typing Test Results
S40	Blood Sample request for Verification Typing
S50	Notification of Verification Typing Sample Shipment
S60	Donor HLA Verification Typing Results
S70	Discrepant Typing Report

Formal Request for HPC, Marrow and/ or HPC, Apheresis

F10	Formal Request for Stem Cell or Lymphocyte Collection
F30	Final Compatibility Test Results

Formal Request for T-cells, Apheresis

F10	Formal Request for Stem Cell or Lymphocyte Collection
F20	Previous Transplant History

Formal Request for 2nd transplant- same recipient and same donor (HPC, Marrow and/ or HPC, Apheresis)

F10	Formal Request for Stem Cell or Lymphocyte Collection
F20	Previous Transplant History



Formal Request for 2nd transplant- same recipient and alternative donor

F10	Formal Request for Stem Cell or Lymphocyte Collection		
F20	Previous Transplant History		
F30	Final Compatibility Test		

Request for HPC, Cord

CB10	Cord Blood Unit – Information and Typing Request
CB21	HPC, Cord Blood Unit Report
CB30	Cord Blood Unit Shipment Request

Forms used by Donor Center/Collection Centre

F50	IDM Testing to be Performed During Donor Workup
F70	Verification of Cell Product
F80	Notification of Donor Clearance
C10	Collection Report
C20	Abnormal Donor Finding Letter
C30	Notification unable to clear

Forms related to Transport

T10	Courier & emergency contact information during stem cell transportation
T20	Emergency Stem Cell Storage Directions
T30	Transport of Stem Cell Product Audit
T40	Courier Letter
T50	Proforma Invoice



Forms related to Donor Follow-up

DF1 Donor Assessment Post Stem Cell Donation

DF3 Annual Donor Assessment Post Stem Cell Donation

Request for Patient's Sibling Service

F90 Request for Patient's Siblings Service

Transplant Follow Up

TF1 Stem Cell Transplantation Follow-up (month(s) after transplantation)

TF2 Stem Cell Transplantation Follow-up (year(s) after transplantation)