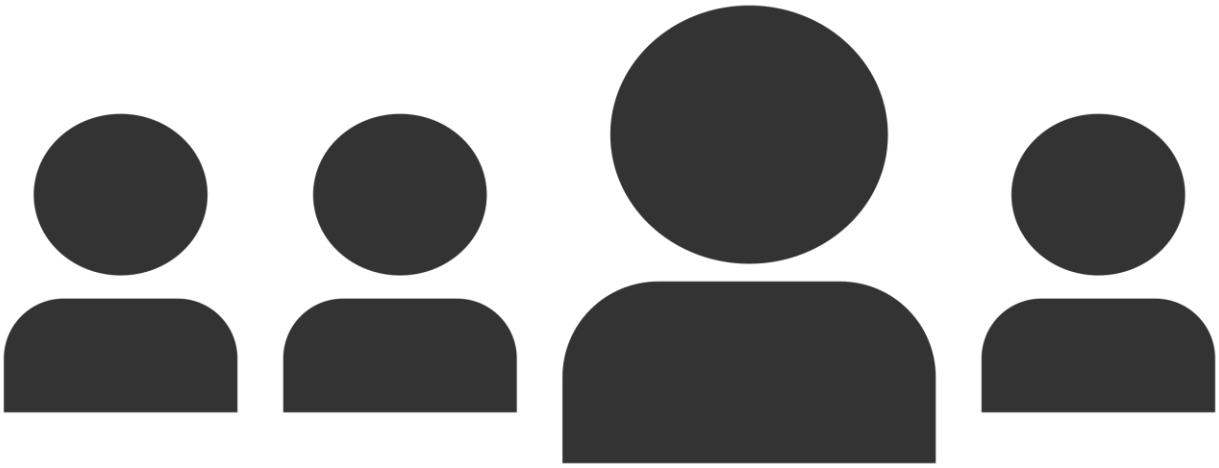
Disclaimer:
"The content of this Deliverable D2.2 represents the views of the author only and is his/her sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains."

D2.2 Educational resources



Grant Agreement number: 811126
Project acronym: SAVDON
Work Package number: WP2

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

"This Deliverable D2.2 of an activity received funding under an operating grant from the European Union's Health Programme (2014-2020)."

Product:

https://share.wmda.info/x/NYIRD

Introduction:

Search & Match Service provides a fast-preliminary search facility to find the best matched haematopoietic stem cell donor or cord blood unit in the world for a patient in need of a haematopoietic stem cell transplant. In January 2017, 'Search & Match Service' became a product of WMDA.

Benefits of Search & Match Service:

- 'Search & Match Service' is the primary source of all haematopoietic stem cell donors and cord blood units globally
- 'Search & Match Service' holds over 33 million donors and over 700,000 cord blood units from 98 registries and cord blood banks from 54 countries
- 'Search & Match Service' helps the community find a donor as quickly as possible by providing:
    - User friendly interface and functionalities
    - Probability matching for each donor/cord powered by OptiMatch
    - Additional information on each donor and cord blood unit (if provided by organisation)

The 'Search & Match Service' public pages contain the user guide for the Search & Match Service, Frequently Asked Questions (FAQ), a demo video showing how to use the system, archive of our newsletter BiteSize Learning, and other useful links that can assist you during your search activities.

Additionally, we provide information about the file format used for data submission, XML.

Future directions of Search & Match Service

WMDA has organised several workshops to work on the future of the Search & Match Service. During the workshops the members of the WMDA worked on:

- Flowcharts and process flows of the end-to-end search, match and connect process
- Explanation of the improved matching algorithm
- Description on how searches are organised within EU Member States. Including a map of the most efficient and effective search, match and connect journey and identification of opportunities and priorities for automation and streamlining (i.e. removal of redundant / non-value adding steps).

This report indicates three outcomes of the workshops:

1. Registry communication workshop outputs, in collaboration with Red Badger
2. Improved Search Algorithm
3. Search & Match Service Data submission information, on WMDA Share

1. Registry communication workshop outputs, in collaboration with Red Badger

RED BADGER

WMDA

matching donors • serving patients

Registry communication workshop outputs

22nd - 23rd October 2018

Registry Communications Workshop

1. Introduction to Red Badger
2. Summary of the workshop
3. Workshop approach
4. Understanding the user; jobs, pains, gains
5. Mapping the user journey and creating the service blueprint
6. Defining the problem statement
7. Project vision/goal
8. Project metrics; KPIs
9. Sketches and prioritisation
10. To-be service map
11. Recommendations for engaging the community

Who Red Badger are

We **bring together the best** in strategic services, user experience and technical delivery **using Lean and Agile processes**.

We're dedicated to **doing the right thing**, which is why we bring our award-winning services to **non-profit organisations** too.

Summary of the workshop

Time to (donor) delivery is key to saving a life. However registry communication can be slow, complex and inaccessible for many.

Eight representatives from both small and large registries worldwide spent two days creating the problem statement, vision and a proof of concept for an enhanced registry communication solution.

Key outcomes:

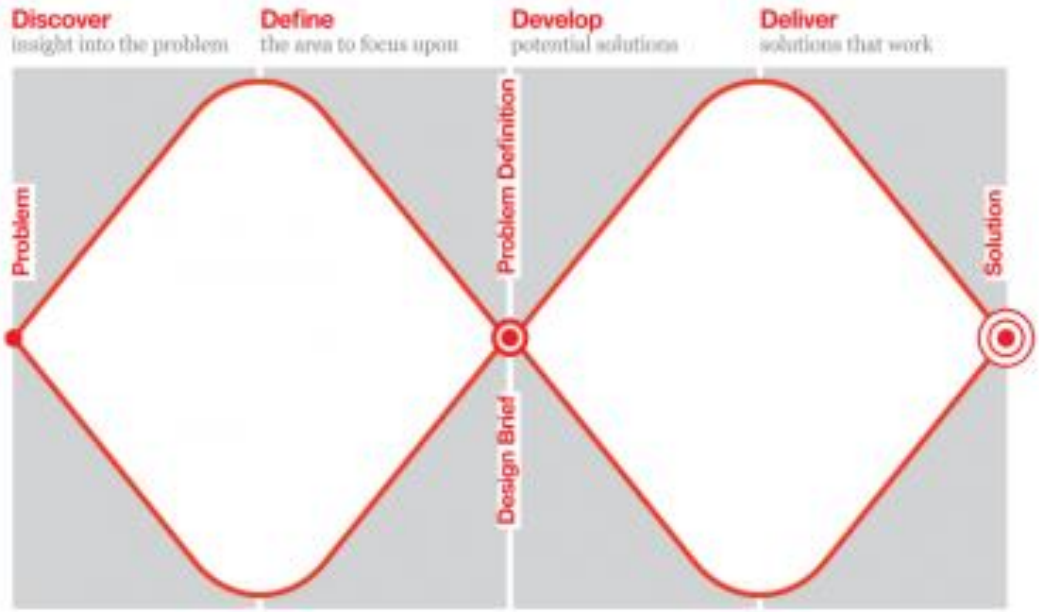
- **Alignment** with selection of stakeholders, and agreement of the **problem definition**
- Agree **the vision**, define what good looks like and the key aspects required to enable it
- **Collaborative ideation** around what a solution might look like
- Feasibility checks as part of a high-level **proof of concept**

Workshop approach: The Double Diamond

The Double Diamond is a visual model of the design process. It's split into two phases, problem definition & solution design.

Each of these phases has two stages; Divergent thinking, where we ask **'yes and'**, building on ideas, and convergent thinking, where we are more critical, to refine ideas by asking **'yes but'.**

1. Diverge to initially discover more about the problem
2. Converge to define the problem
3. Diverge to develop potential ideas
4. Converge to evaluate and get to a solution



Registry Communications Workshop

Understanding the user

1. Understand the environment
   - How do registries currently communicate with one another?
   - What are the key constraints?

2. User deep dive
   - Key jobs, pains and gains

3. Mapping the existing user journey
   - A service blueprint mapping existing journey, touchpoints, highlights and lowlights in the experience

Jobs

Things the user is trying to get done:

- Check donor availability
- Send and receive donor information
- Find a donor match
- Patient follow up
- Invoicing and payments
- Blood sample updates

Pains

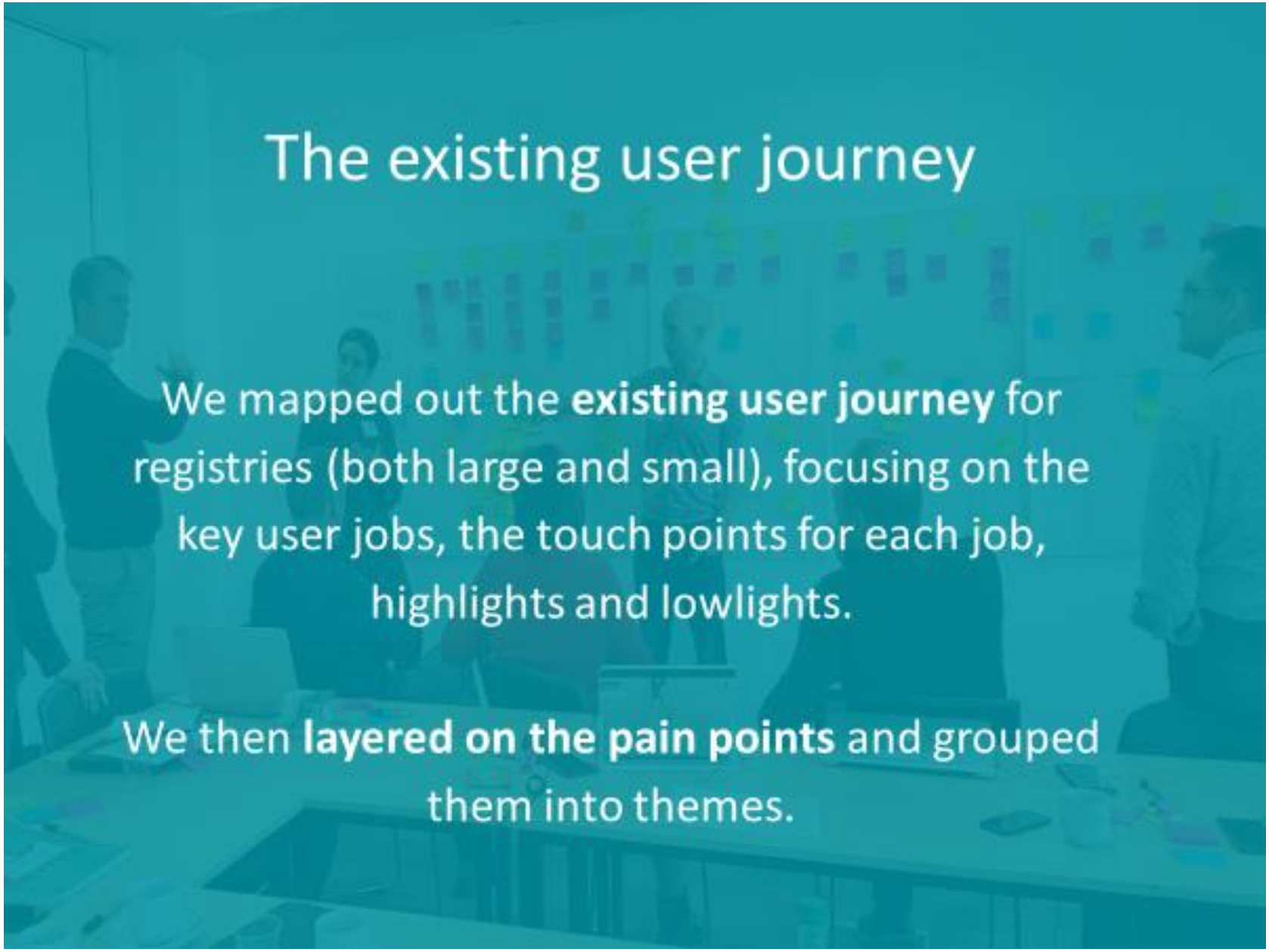
Things that may prevent, delay or block the user from getting their job done.

- No single source of truth
- Response time slow
- Data quality (not up to date)
- Too many communication paths
- Complex system
- Data Protection challenges
- Invoicing process

Gains

Potential gains we can provide the customer which will enable them to get their job done:

- Easy to use
- Faster, flexible and efficient
- Reliable data
- Automated comms
- Ready and easy to access data
- Customised donor search


The existing user journey
We mapped out the existing user journey for registries (both large and small), focusing on the key user jobs, the touch points for each job, highlights and lowlights.
We then layered on the pain points and grouped them into themes.

The existing user journey

Doing

Touchpoints

Feelings

Themes

Pain points

Themed pain points

Lack of data quality, automation, visibility, standardisation & agility key

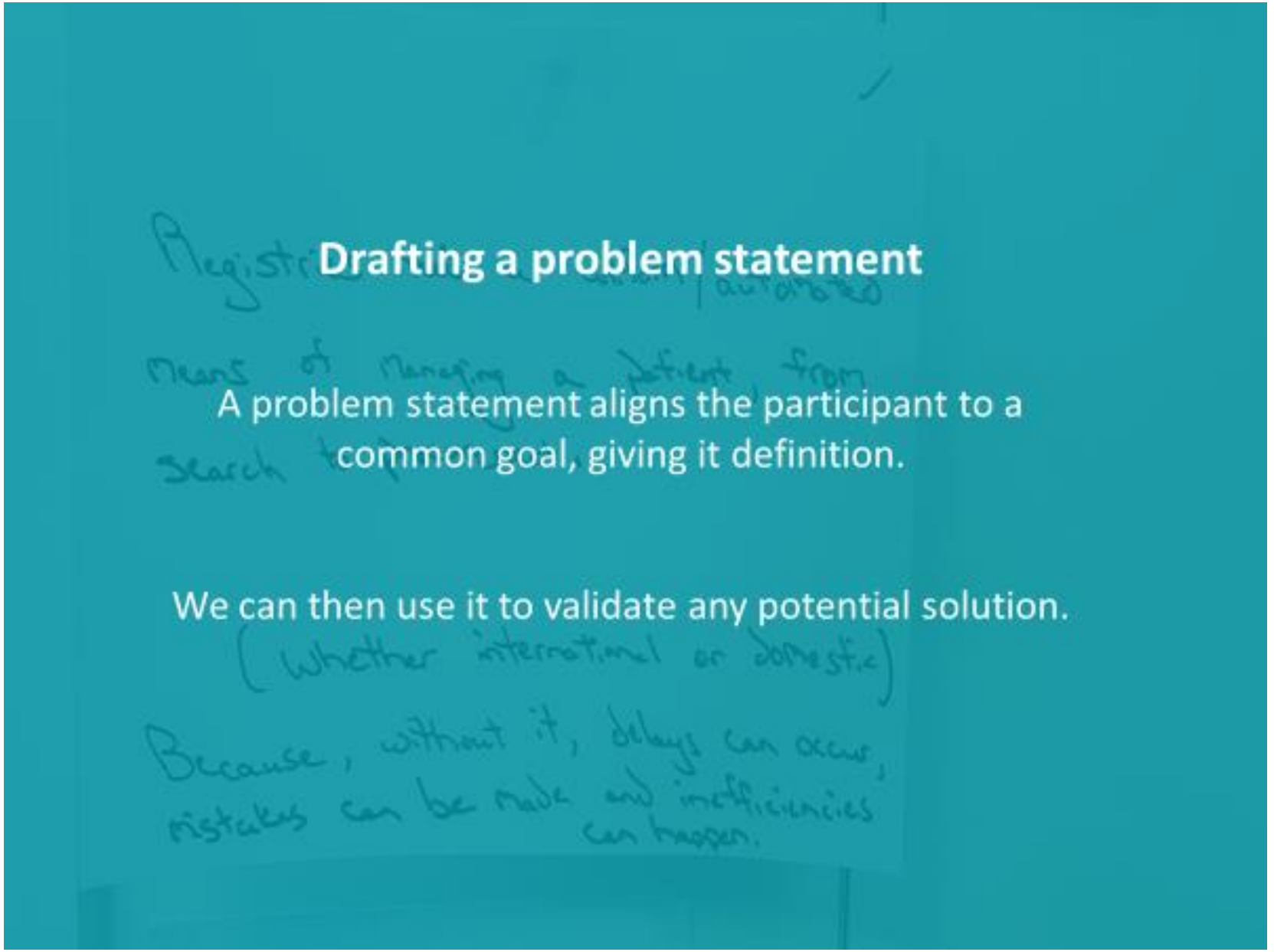
Registry Communications Workshop

Defining

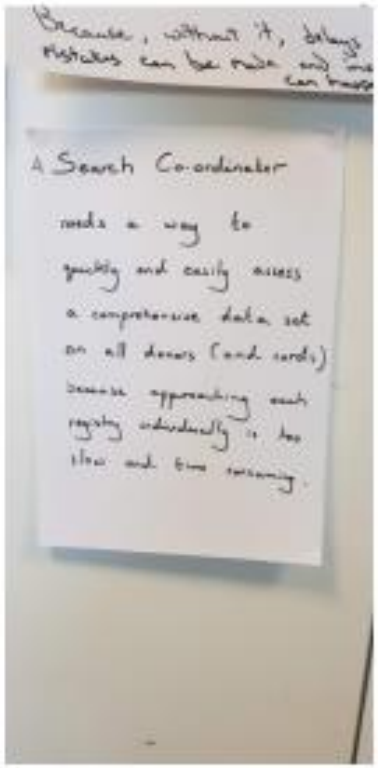
1. Define a problem statement
- Based on prioritised, themed pain points

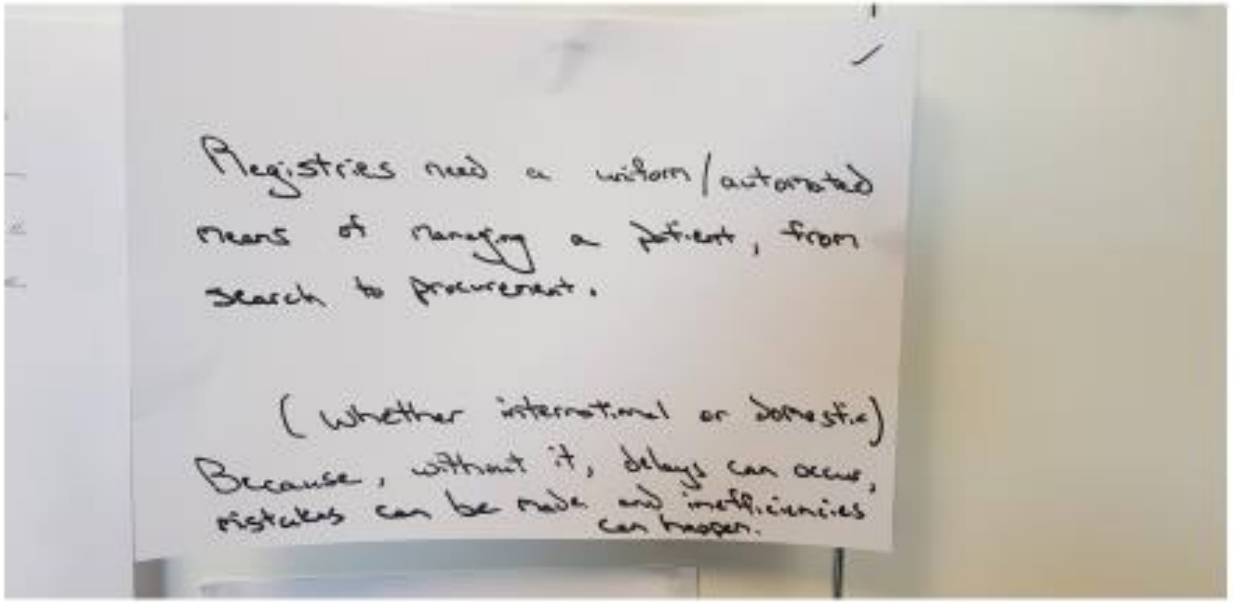
2. Agree the goal
- How Might We.. activity to help frame the problem

3. Define how we will measure success (Key Performance Indicator)
- How will we know that we have achieved what we set out to?

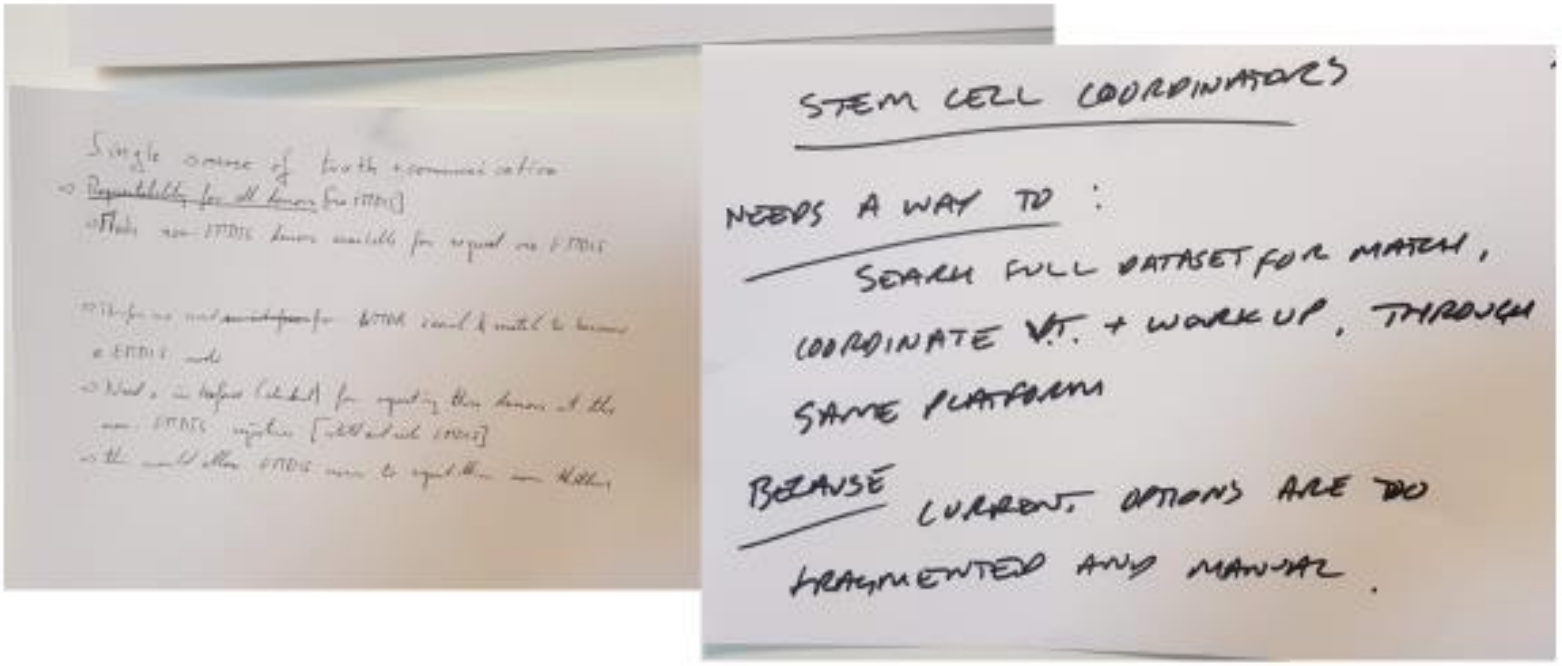

Drafting a problem statement
A problem statement aligns the participant to a common goal, giving it definition.
We can then use it to validate any potential solution.

Problem statement (activity)

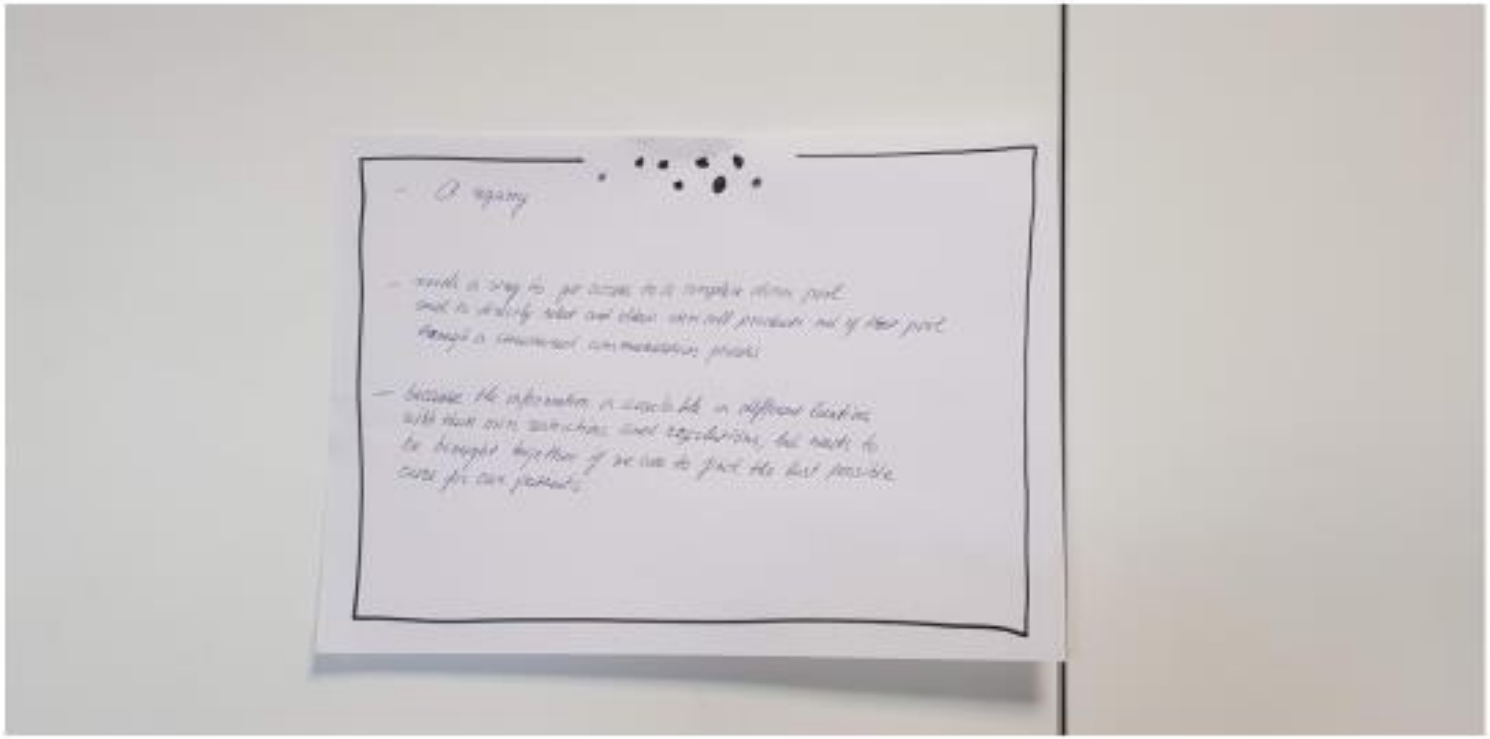




Problem statement (activity)



Problem statement (top voted)



Problem statement

Following the workshop, Red Badger iterated on the problem statement and created a second, simplified version:

A registry

User's name / description

Needs a way to...

Get access to a complete donor pool and to directly select and obtain stem cell products out of that pool through a structure communication process

Because... *User's need*

The information is available in different locations with their own restrictions and regulations, but needs to be brought together if we are to find the best possible cure for our patients.

Insight

A registry

User's name / description

Needs a way to...

Access a complete donor pool and to select and obtain stem cell products out of that pool

But *User's need*

The information is available in different locations, each having its own constraints.

Insight


'How Might We' activity output
Agreed statement:
"How might we create an integrated system that allows search and requests in small & large registries?"

Key performance indicators

There are three key areas to focus on when measuring effectiveness of registry to registry communication tools and services; adoption, efficiency and usability.

In order to create a target, a baseline measurement would need to be taken. Here are some sample KPIs which could be used to measure success.

Adoption:

- 90% of registries participating and responsive within 14 days

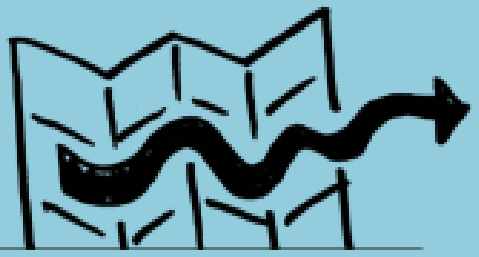
Efficiency:

- Lead time of donor search & find

Usability:

- 100% of all registry to registry transactions performed via new services of API tooling

Registry Communications Workshop



Ideation & feasibility

1. Rapid ideation

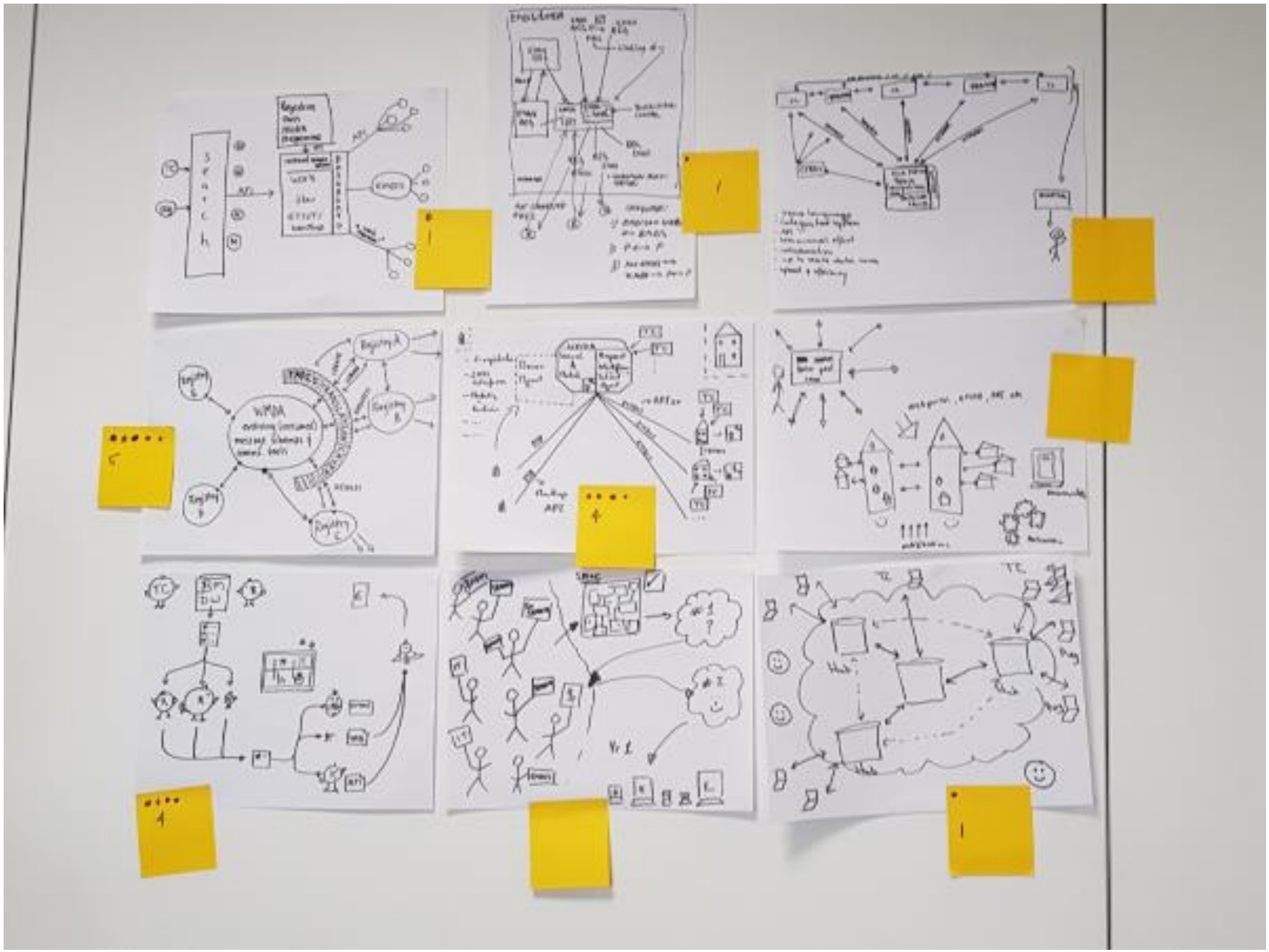
- Sketching workshop, outputs prioritised, feasibility discussed

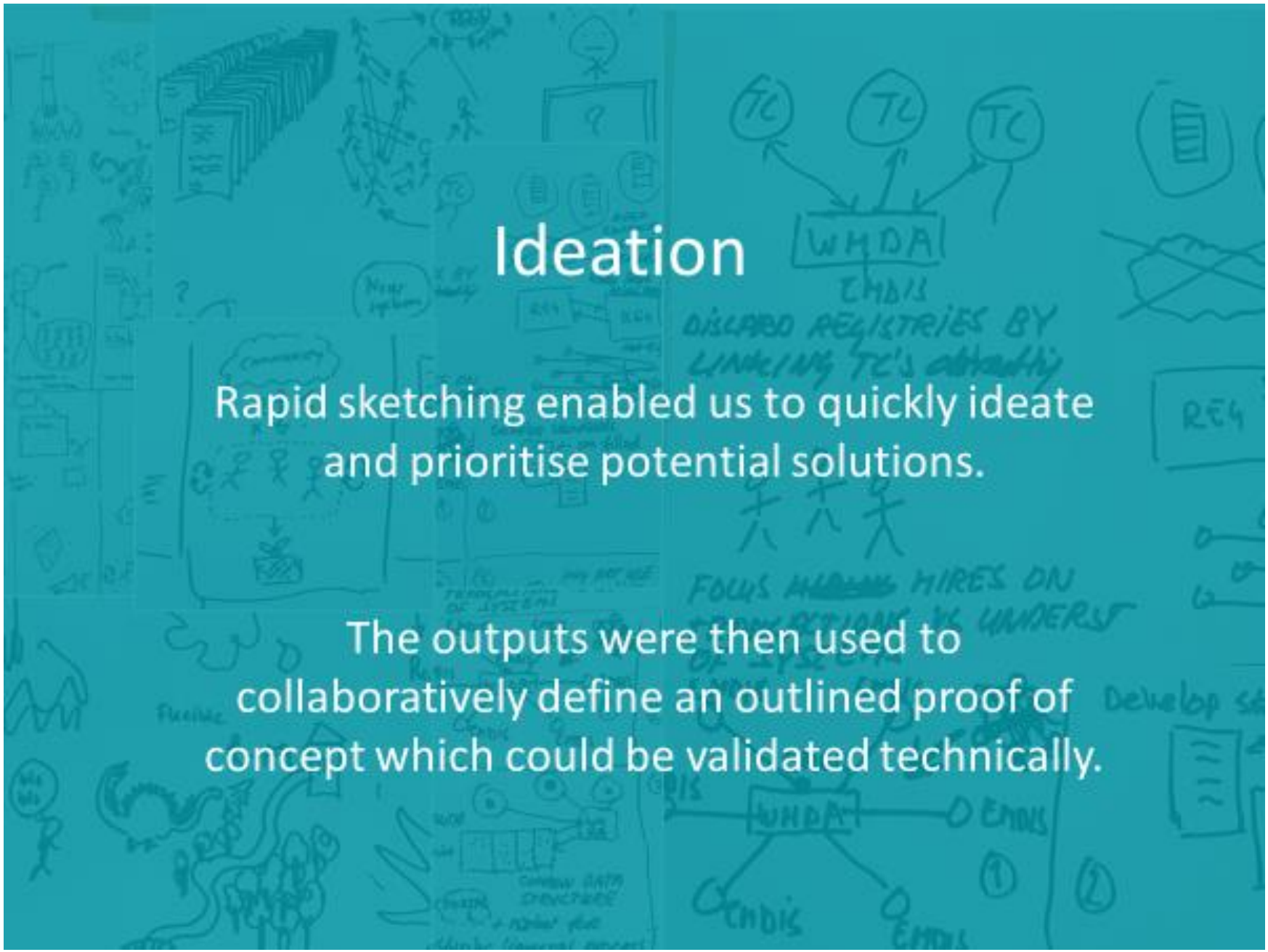
2. Prioritisation

- Options voted on based on meeting key requirements, top options merged to ensure coverage

3. Collaboratively defined a PoC

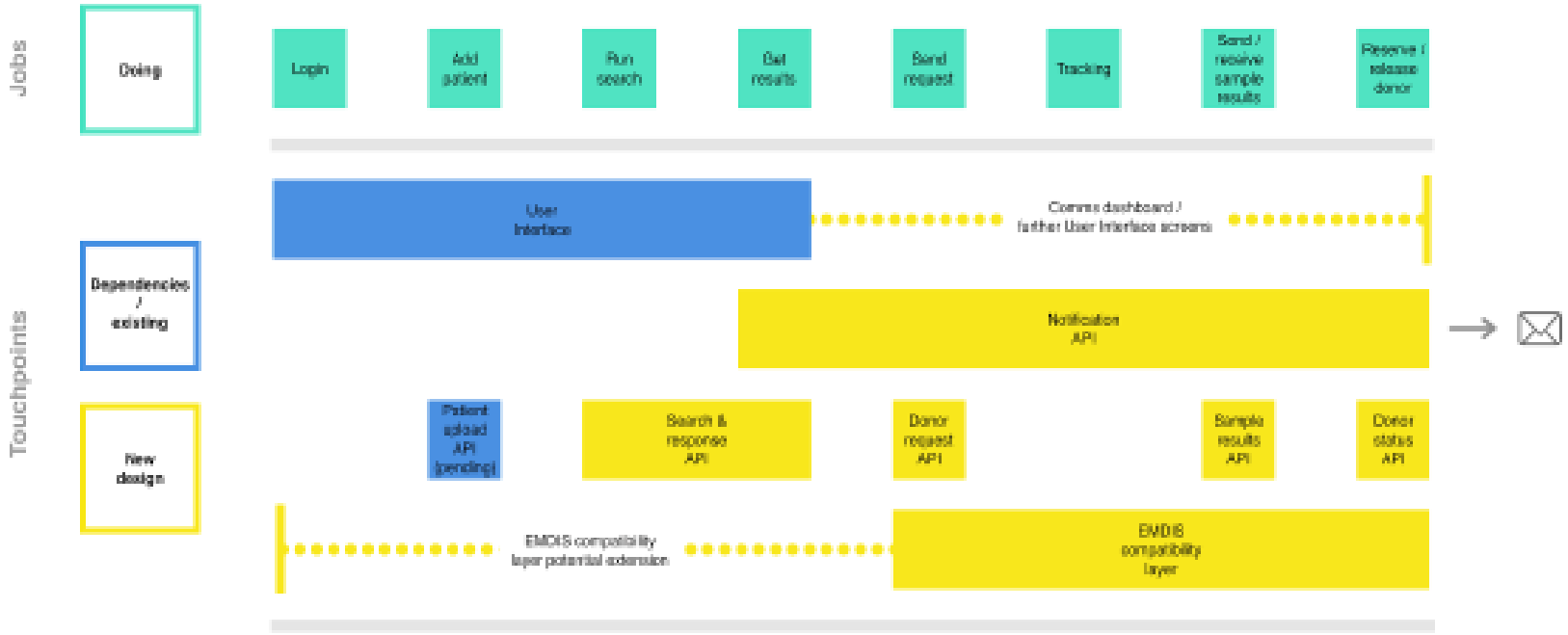
- High level epics (user jobs) created as part of to-be service map and n
  functional requirements identified



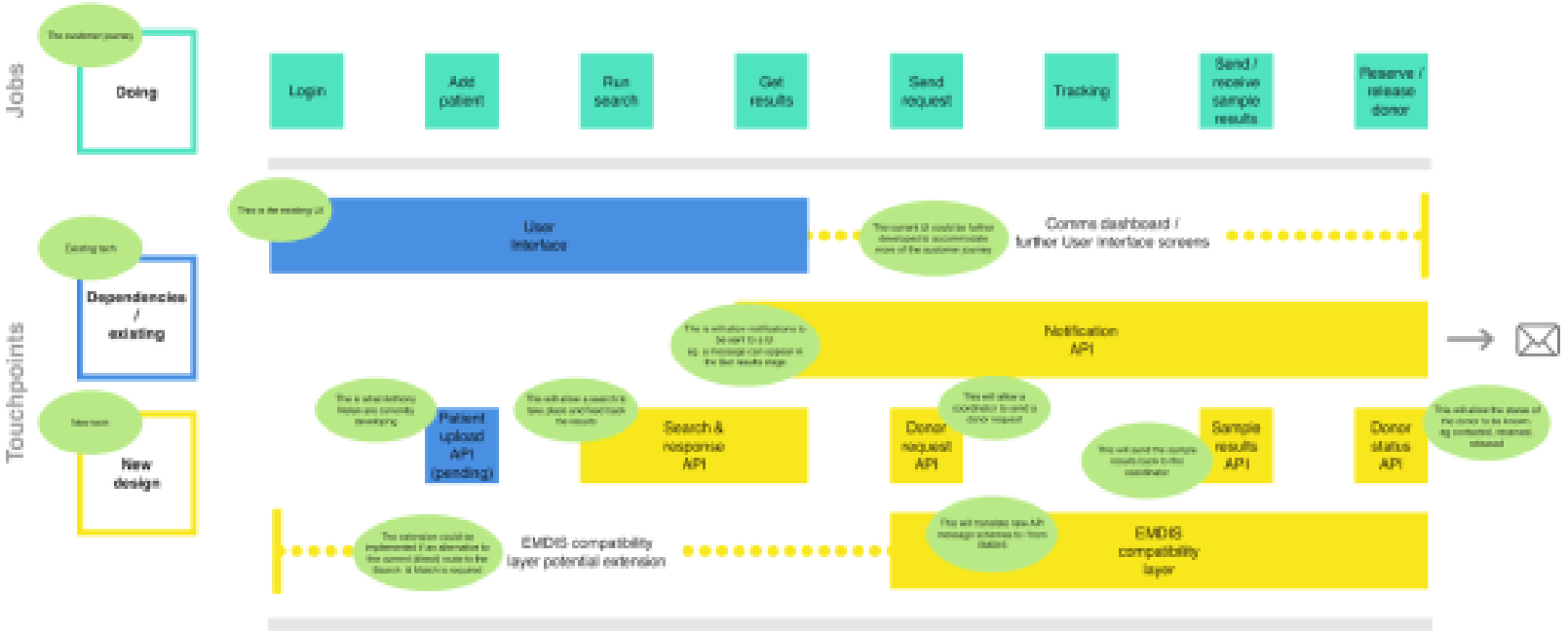

Ideation
Rapid sketching enabled us to quickly ideate and prioritise potential solutions.
The outputs were then used to collaboratively define an outlined proof of concept which could be validated technically.

To-be service map

This is an outline of what a new solution could look like, as well as how it integrates with existing services. By starting with the user jobs and then moving into technical touchpoints we remained focused on user value and experience.



To-be service map - explained

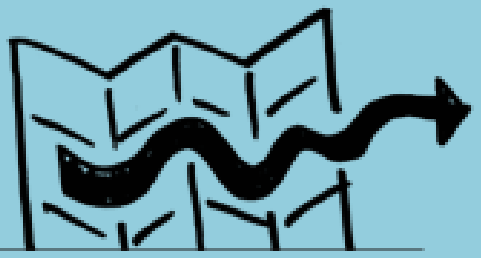


Requirements

We listed all the features and attributes that we take for granted. Whether they are non-feature specific, or non-functional, they would be key to realising the vision.

1. Data quality and timeliness
2. Performance capability and response time
3. Security and authentication
4. Privacy & compliance (e.g. GDPR)
5. Agility and flexibility
6. Governance
7. Browser/device support
8. National requirements consolidated
9. Support and maintenance provision

Registry Communications Workshop



Engaging the community

1. Communication recommendations

2. User Survey suggestion

Recommendations

In order to gain good engagement and alignment around the proposal, we suggest you position it carefully in order to minimise any potential alienation from within the community.

Initial framing

- We listened to user feedback from previous meetings about this challenge, that is why we decided to escalate the issue
- Participants which represent both the business and search coordinator roles, from a range small and large registries, have fed into this idea

Presentation

- In order to ensure that no single registry is seen as 'owning' this solution, we recommend that multiple registries should jointly present it

Engagement

- Attendees should be asked for input, this could involve a survey or vote (see next slide)

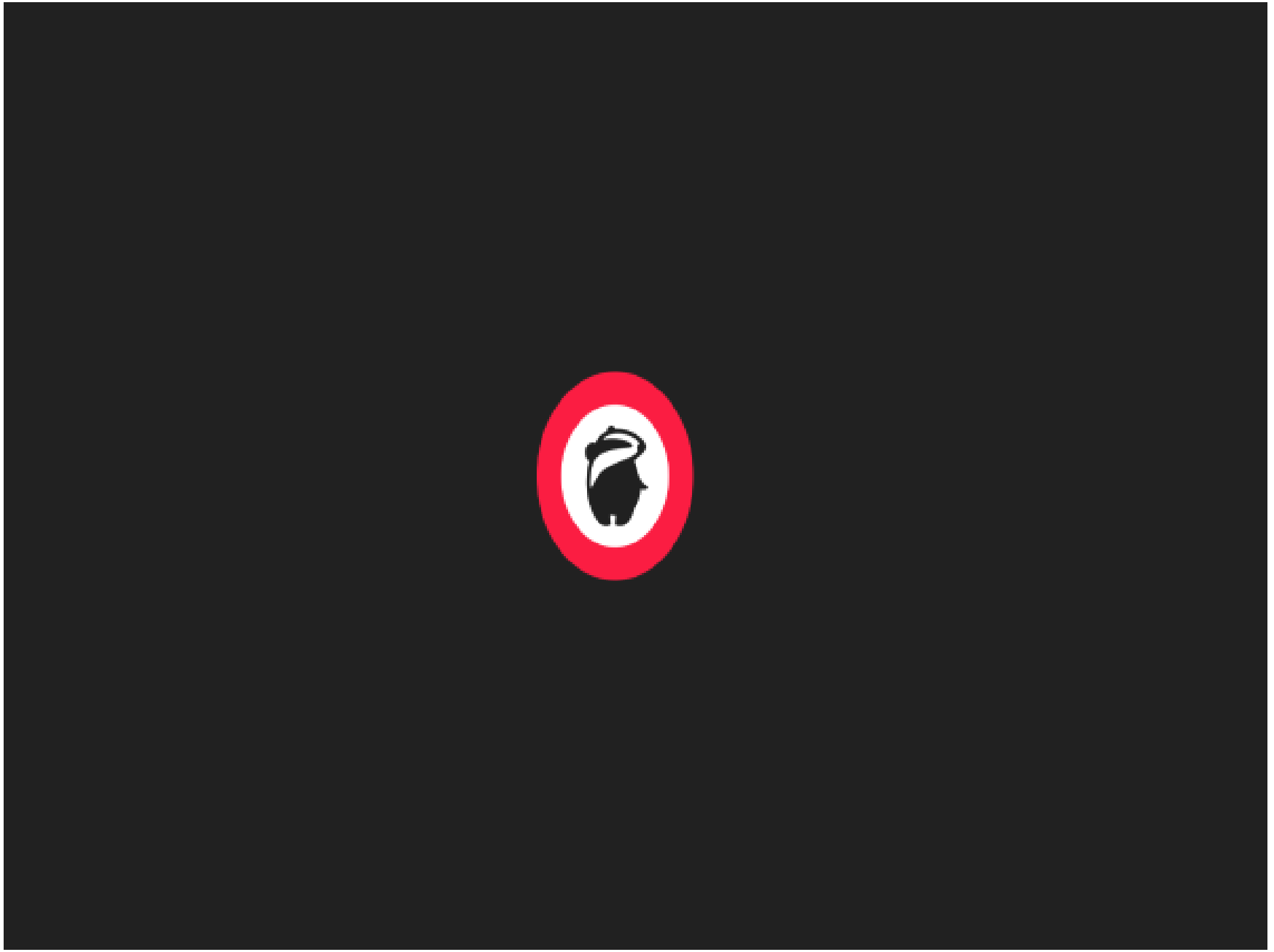
Engaging the community

A survey to involve attendees and drive priority.

Engagement can be leveraged for recruitment of a User Group to input into requirements, product development and usability testing.

Thanks for taking the time to complete our quick survey, we'll use your responses to help direct which areas to start working on first. You will also have the opportunity to input into development and help with testing.

- Think about your registry-to-registry communication experience. Rank the following in order of where you experience the most pain (in relation to registry-to-registry communication).
  - No single source of truth
  - Response time slow
  - Data quality (not up to date)
  - Too many communication paths
  - ---
- Become more involved in improving your registry-to-registry communication experience. Enter your email address and we'll contact you about how you can get involved



2. Improved Search Algorithm

matching donors • serving patients

D2.1 Search Algorithm

An expanded search algorithm that accomodates searches of non-North-Western descent donors

Allogeneic haematological stem cell donors for:

- Patients with haematological malignancies (leukaemia)
- Patients with non-malignant haematological diseases (aplastic anaemia, immunodeficiency syndromes, inborn errors)

What is the chance to find one …

Photo by CEphoto, Uwe Aranas, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=38131336

Donor selection

- Donors are primarily selected on compatibility for Human Leukocyte Antigens (HLA).
- HLA matching is done on HLA-A,-B,-C,-DRB1 and often also on -DQB1.
- HLA is very polymorphic, this means that there are many different HLA types.
- The frequency of the various HLA types differs in the different ethnic groups.

To select the best stem cell source

- Patients in need of an allogeneic hematopoietic stem cell transplantation need a suitable stem cell source.
- The first choice is an HLA identical family donor.
- Only about 25-30% of patients have such a donor.
- For 70-75% of patients an unrelated donor needs to be found.

Unrelated donors

- HLA data of over 30 million donors can be found in the Search & Match Service of WMDA.
- These donors originate from 54 countries.
- **Q: How do you find the HLA matched donor amongst the over 30 million donors?**

Unrelated donors

- HLA data of over 30 million donors can be found in the Search & Match Service of WMDA.
- These donors originate from 54 countries.
- **Q: How do you find the HLA matched donor amongst the over 30 million donors?**

- *A: You need a matching algorithm*

Which one shall I take?

Matching Algorithm

There are two main matching algorithms:

- Standard matching algorithm (original algorithm)
- Probability matching algorithm (new expanded algorithm)

HLA typing level of donors in Search & Match Service

- Not all donors are fully typed for HLA-A,-B,-C,--DRB1 and -DQB1.
- Not all donors are typed at the desired high resolution level. Many are at low resolution level or at intermediate resolution level.
- This makes the selection of the donor(s) that will turn out to be HLA compatible with the patient very difficult.

Standard matching algorithm

A		B		C		DRB1		DQB1		Reg	#
HLA patient:											
11:01	24:02	44:03	55:01	03:03	16:01	04:01	15:01	03:02	06:02		
Potential (HLA allele) matched donors:											
11:01	24:02	44:XX	55:01			04:01	15:01	03:02	06:02	NL	1
11:XX	24:XX	44:XX	55:XX	03:XX	16:XX	04:01	15:01	03:02	06:02	GB	1

XX= low resolution typing: Could potentially be any allele

Standard matching algorithm

Which donor to choose?

A		B		C		DRB1		DQB1		Reg	#
HLA patient:											
11:01	24:02	44:03	55:01	03:03	16:01	04:01	15:01	03:02	06:02		
Potential (HLA allele) matched donors:											
11:01	24:02	44:XX	55:01			04:01	15:01	03:02	06:02	NL	1
11:XX	24:XX	44:XX	55:XX	03:XX	16:XX	04:01	15:01	03:02	06:02	GB	1

NL donor or the GB donor?

XX= low resolution typing: Could potentially be any allele

Standard matching algorithm

Which donor to choose?

A		B		C		DRB1		DQB1		Reg	#
HLA patient:											
11:01	24:02	44:03	55:01	03:03	16:01	04:01	15:01	03:02	06:02		
Potential (HLA allele) matched donors:											
11:01	24:02	44:XX	55:01			04:01	15:01	03:02	06:02	NL	1
11:XX	24:XX	44:XX	55:XX	03:XX	16:XX	04:01	15:01	03:02	06:02	GB	1

NL donor is typed at higher level than the GB donor therefore choose the NL donor?

XX= low resolution typing: Could potentially be any allele

Standard matching algorithm

Which donor to choose?

	A		B		C		DRB1		DQB1		Reg	#
	HLA patient:											
	11:01	24:02	44:03	55:01	03:03	16:01	04:01	15:01	03:02	06:02		
Probability	Potential (HLA allele) matched donors:											
3%	11:01	24:02	44:XX	55:01			04:01	15:01	03:02	06:02	NL	1
90%	11:XX	24:XX	44:XX	55:XX	03:XX	16:XX	04:01	15:01	03:02	06:02	GB	1

Probability matching tells you that the GB donor has a matching probability of 90% and the NL donor only 3%.

XX= low resolution typing: Could potentially be any allele

14

3:42 / 8:32

Standard matching algorithm

Which donor to choose?

	A		B		C		DRB1		DQB1		Reg	#
	HLA patient:											
	11:01	24:02	44:03	55:01	03:03	16:01	04:01	15:01	03:02	06:02		
Probability	**Potential (HLA allele) matched donors:**											
3%	11:01	24:02	44:XX	55:01			04:01	15:01	03:02	06:02	NL	1
90%	11:XX	24:XX	44:XX	55:XX	03:XX	16:XX	04:01	15:01	03:02	06:02	GB	1

Probability matching tells you that the GB donor has a matching probability of 90% and the NL donor only 3%.

Answer: Choose the GB donor instead of the NL donor

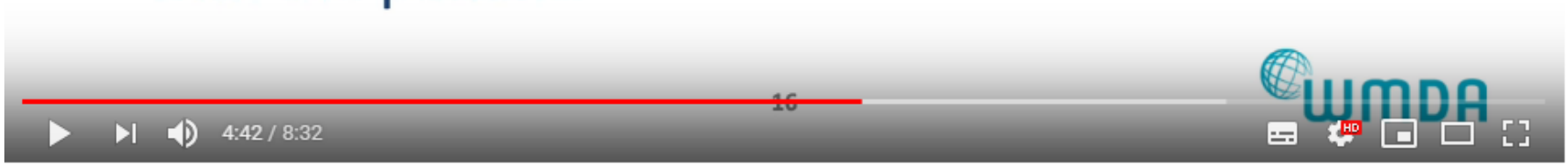
XX= low resolution typing: Could potentially be any allele

Probability matching

- Many donors are not typed at the resolution desired on all loci.
- This leads to ambiguity and uncertainty about the real match grades of such donors.
- “Standard” matching relies on the knowledge and skills of search coordinators to predict which donors might be matches once they are typed more extensively.

Probability matching – The reason

- Huge and still increasing polymorphism of HLA makes it difficult for search coordinators to predict matches. Extensive knowledge on HLA is needed.
- Difficult to choose donors with highest chance to match with the patient when there are many incompletely typed donors.
- Probability matching uses knowledge to predict which donors have the highest chance to match with the patient.



Probability matching – How it works

- Estimates HLA-A, -B, -C, -DRB1 (and -DQB1) haplotype frequencies based on the phenotype data of a population.
- Calculates the likelihood that patient and donor match using the estimated haplotype frequencies.
- Considers allele associations over complete haplotypes.

Probability matching – Example

Probability of mismatches 0, 1, 2	A 24:02 11:01	B 44:02 13:02	C 05:01 06:02	DRB1 04:01 07:01	DQB1 03:01 02:02	DPB1	DRB3/4/5
Donor details:	Registry Name: Israel-Ezer Mizion Bone Marrow Donor Registry				Donor ID:		
31 P P P P A	11 24	13 44		04:XX 07:XX	02:02 03:01		
82% , 12% , 5%	99%	95%	86%	96%	100%		
Donor details:	Registry Name: Brazil-REDOME - Registro Nacional de Doadores Voluntarios de Medula Ossea				Donor ID:		
32 P P P P P	24:AGVE 11:YPP	44:AJAZ 13:BC		07:APA 04:NJV			
78% , 19% , 3%	99%	99%	90%	99%	85%		
Donor details:	Registry Name: USA-National Marrow Donor Program				Donor ID:		
33 P P P P P	24:CNFM 11:ANRC	44:AJAZ 13:BC		07:APA 04:NJV			
78% , 19% , 3%	99%	99%	90%	99%	85%		
Donor details:	Registry Name: USA-National Marrow Donor Program				Donor ID:		
34 P A P A P	24:DRZW 11:CRGR	44:02:01G 13:02		07:01 04:01			
78% , 19% , 3%	99%	100%	90%	100%	85%		

First improvement to the matching algorithm

- Implemention of the probability matching algorithm based on HLA haplotype frequencies of all donors in the Search & Match file.

- As the majority of donors are of North-Western descent it worked well for these donors but not for non-North-Western descent donors.

Second improvement to the matching algorithm

- The HLA haplotype frequencies were calculated based on the different countries rather than on the entire file of donors.

- This has improved the matching results for non-North-Western descent donors significantly.

Especially helpful for searches of non-North-Western descent donors

- Donors from non-North-Western descent are often not fully typed.
- Donors from non-North-Western descent originate from many different countries.
- The HLA types in the different countries vary enormously.
- To have knowledge of all these different HLA types is practically impossible for one person.
- This makes selection of donors from non-North-Western descent with the standard matching algorithm especially difficult.

The probability matching algorithm can predict the likelyhood that a donor will be matched based on phenotype frequencies of the particular ethnic group.

3. Search & Match Service Data submission information

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.

2. 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	Type	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.
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.
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.
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	Type	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: XXXXXXXXXXXXXXXXXXX. Only upper case letter and numbers are allowed. The first 4 characters include the ION were the donor is registered. The following 13 characters contain the donor specific ID and the last two characters are check characters that are calculated from the 17 previous characters.
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 1968-06-28, 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 AFSS = African: Sub-Sahara Africa ASSW = Asian: Southwest Asia (Middle East, Turkey) ASSO = Asian: Southern Asia (India, Pakistan, Bangladesh, Sri Lanka, Bhutan, Nepal) ASCE = Asian: Central Asia (Eastern Russia, Kazakhstan, Uzbekistan, Kyrgyzstan, Tajikistan) ASSE = Asian: Southeast Asia (China, Mongolia, Burma, Laos, Cambodia, Thailand, Vietnam, Taiwan) ASNE = Asian: North and Northeast Asia (Japan, North Korea, South Korea) ASOC = Asian: Oceania (Pacific Islands, excluding Japan, Australia, Taiwan, Sakhalin, Aleutian Islands) CAEU = Caucasian: Mainland Europe, Greenland, Iceland, Western Russia CAER = Caucasian: Eastern Russia CANA = Caucasian: North America (USA, Canada, Mexico) CAAU = Caucasian: Australia (Australia, New Zealand) HICA = Hispanic: Central America, Caribbean HISA = Hispanic: South America MX = Mixed / multiple OT = Other (e.g. Australian Aborigine) UK = Unknown
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
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></A> ...
or
<DQB1><DNA><FIELD1>05:02:01G</FIELD1><FIELD2 /></DNA></DQB1>

Field Identifier	Required	Description	Type	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
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
B	Yes	HLA-B values	hlaLocusType		Both SER and DNA possible; either SER or DNA values required
C	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	Type	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	Type	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.
TOXO	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

5 donItemType elements

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

Field Identifier	Required	Description	Type	Length	Comment
STAT_REASON	No	Additional information relevant to the donor 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 = "-"
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	Short		Number between 1 and 999, no decimals
HEIGHT	No	Height in 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 cbuItemType elements

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

Field Identifier	Required	Description	Type	Length	Comment
STAT_REASON	No	Additional information relevant to the CBU 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	
BANK_MANUF_ID	No	Unique identifier of the CBB that manufactured the CBU. ID shown in table in tab Cord blood bank IDs	String	10	PLEASE NOTE: For the upload the fields BANK_MANUF_ID and BANK_DISTRIB_ID should be fulfilled with a new ID for the corresponding cord blood banks 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	No	Unique identifier of the CBB distributing the CBU. ID shown in table in tab Cord blood bank IDs	String	10	PLEASE NOTE: For the upload to WMDA the fields BANK_MANUF_ID and BANK_DISTRIB_ID should be fulfilled with a new ID for the corresponding cord blood banks 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.
COLL_DATE	No	Date that the CBU was collected	bareDateType	10	Date without timezone information, example 1968-06-28, Date Delimiter = "-"
PROC_DATE	No	Date that the processing started	bareDateType	10	Date without timezone information, example 1968-06-28, Date Delimiter = "-"
PROC_METH	No	Processing method used	procMethType	3	procMethType: "HES","DGS","CEN","FIL","FIC","PER","OTH" 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 = "")

PROC_METH_TYPE	No	Processing method type used	procMethTypeType	3	procMethTypeType: "MAN","SPX","OTP","AXP","OTH" MAN = Manual SPX = Sepax OTP = Optipress II AXP = AXP OTH = Other
FREEZE_DATE	No	Date that the CBU was frozen	bareDateType	10	Date without timezone information, example 1968-06-28, Date Delimiter = "-"
FREEZE_METH	No	Freezing method used	freezeMethType	1	freezeMethType: "C","M" C = Controlled Rate M = Manual
PROD_MOD	No	Product modifications made	prodModType	3	prodModType: "BCE","DNE","PLR","PRR","RBR","NOT","OTH" 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
BAG_TYPE	No	Type of bag used (bag fractions / split unit)	bbagTypeType	5	bagTypeType: "80/20","50/50","40/60","NS" (no split)
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
VIABILITY	No	Viability as percentage value	Short		Number between 0 and 100, no decimals
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 cells 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	Type	Length	Comment
ID	No	Identification used to identify the maternal donor as assigned by the registry	String	15	
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