

**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 Progress report on the accuracy, quality of data in the global database.

**Grant Agreement number:** 101015514

**Project acronym:** SAVDON

**Work Package number:** WP1

**Periodic report:** 1st  2<sup>nd</sup>  3<sup>rd</sup>  4<sup>th</sup>

**Period covered:** from 01<sup>st</sup> January to 31<sup>st</sup> December 2021

**Organisation:** **World Marrow Donor Association (WMDA)**

**LEAR:** Lydia Foeken

**Project coordinator:** Lydia Foeken

**Tel:** 0031 88 505 7900

**E-mail:** lydia.foeken@wmda.info

**Organisation website address:** [www.wmda.info](http://www.wmda.info)



Co-funded by  
the Health Programme  
of the European Union

*"This Deliverable D1.1 of an activity received funding under an operating grant from the European Union's Health Programme (2014-2020)."*

## Table of Contents

<b>ABBREVIATIONS .....</b>	<b>3</b>
<b>1. INTRODUCTION.....</b>	<b>4</b>
1.1 BACKGROUND.....	4
1.2 REQUIREMENTS.....	4
<b>2. PROGRESS AND MAINTENANCE OF DATA QUALITY .....</b>	<b>6</b>
2.1 UPDATING THE DATA DICTIONARY .....	6
2.2 DATA DICTIONARY RFC AND XSD 2.4 .....	6
<i>Data Dictionary RFCs in 2021.....</i>	<i>6</i>
<i>XSD 2.4 .....</i>	<i>7</i>
2.3 DATA QUALITY WEBINAR .....	8
2.4 IMPACT OF ALLDIFF-DIFFERENTIAL UPLOAD FEATURE .....	8
2.5 CONTINUED MONITORING TOWARDS OPTIMIZATION AND IMPROVED DATA QUALITY.....	10
<i>Density Improvement in 2021.....</i>	<i>10</i>
<i>Density report per registry Improvement .....</i>	<i>12</i>
<i>Deprecated Code.....</i>	<i>14</i>
<i>XSD 2.2, 2.3 End User Usage Status.....</i>	<i>14</i>
<b>3. BEYOND 2021.....</b>	<b>17</b>
3.1 GAP ANALYSIS – COMPARING LIVE REGISTRY DATA WITH UPLOADED DATA .....	17
3.2 HLACORE .....	17
3.3 LEVERAGING MACHINE LEARNING TO ADVANCE GLOBAL DATABASE .....	18
<b>APPENDIX 1.....</b>	<b>19</b>
<b>APPENDIX 2.....</b>	<b>26</b>
<b>APPENDIX 3.....</b>	<b>30</b>

## Abbreviations

ADCU = Adult Donor Cryopreserved Unit

CBB = Cord Blood Bank

CBU = Cord Blood Unit

CI/CD = Continues Integration/ Continues Development

DD = Data Dictionary

DQ = Data Quality

HLA = Human Leucocyte Antigen

QA = Quality Assurance

R2S = Ready-to-ship

NMDP = National Marrow Donor Programme

Organisations = Donor registries or Cord Blood Banks

RFC = Request For Change

TNC = Total Nucleated Cell

WMDA = World Marrow Donor Association

XML = Extensible Mark-up Language

XSD = XML Schema Definition

## 1. Introduction

In the field of hematopoietic stem cell transplantation (HSCT), with an increasingly growing number of voluntary unrelated donors globally, it is key that the data of all these donors has the highest security, accuracy and quality possible to ensure patients can be transplanted promptly and safely.

This progress report on data quality in EU Member States (D1.1 – 2021) details the process and achievements WMDA and each EU member state has made towards meeting the defined data set requirements in 2021. This report provides progress information to allow comparative analysis of progress by EU member states and will enable them to focus their efforts on areas of identified weaknesses.

### 1.1 Background

In April 2018, WMDA successfully upgraded the data upload system for WMDA global donor and cord blood database, which is the global service for all European transplant centres and search coordination units to find the best suitable stem cell source. The new system replaced the outdated DOT20 format with the modern XML in the upload file. After this system upgrade, over the course of 2018 and 2019, the vast majority of organisations gradually transitioned to the XML file format for data uploads.

As the industry grows and develops, new data requirements emerge that the WMDA database needs to be able to support. Therefore, an updated version of the XML file format is introduced annually with backwards compatibility of at least two older versions. Already most EU records is uploaded using the WMDA XSD 2.2 standard, however a significant amount of non-EU members is on older XSD versions. WMDA has initiated several steps to assist member organisations move to XSD 2.2 or even the latest 2.3.

The WMDA Accreditation Standard 5.10 states that: “The registry is expected to be an active WMDA member, demonstrated by submitting to the global trends report of its activities and by uploading the donor and cord blood data on a regular base to the Search & Match Service (at least monthly) and by storing its database electronically.”. With this addition to the 2020 Accreditation Standards, donor and CBU data is uploaded to the global database more frequently, so to facilitate this, WMDA has introduced the ALLDIFF differential upload programme. This programme tracks the changes made in the new data set compared to the old file.

To ensure future data processing security and speed, WMDA always observes the recent developments on data processing. Recently WMDA has opted to move from on-premises server hosting at a third-party hosting company to cloud hosting namely Microsoft’s Azure. The move to cloud computing gives similar to better performance compared to the old configuration and promises benefits for WMDA’s services in the future. For example, cost savings on server capacity.

### 1.2 Requirements

The XML file format significantly increases the number of data elements (fields) that can be handled and stores while also providing the possibility to expand the definition by adding more data elements in the future. However, more data may lead to more issues during processing or the data transfer from member organisations to WMDA Search & Match Service. To guarantee the accuracy and quality of the data,

lessening any uncertainty for search coordinators when using Search & Match Service, WMDA defined a requirement plan to improve data quality based on 3 main data characteristics: Completeness, Timeliness and Reliability (Consistency and Accuracy).

The data quality plan has two main objectives:

- Improve, optimize and enhance current data upload solution
- Investigate possible new solutions or replacement certain parts in the data process flow for Search & Match

In 2019 and 2020, data quality projects mainly focused on improving or automating the process for quality assurance checks and provided reports or dashboards to help organisations or WMDA to monitor their data quality. These projects helped WMDA to quickly check and confirm the possible data upload issues in order to make the data available for search coordinators and transplant centres in the most efficient way possible.

In 2021, WMDA continued working on improvements to optimize the data upload and maintained strategies already put in place. Investigations into possible new solutions or replacement of certain parts in the data process flow for Search & Match has also started, addressing the second objective of the data quality plan.

Small organisations seldom have a data quality management system within their organisation, the data quality plan helps these small organisations to improve their data internally. Chapter 2, Progress and Maintenance of Data Quality, details how WMDA handles these set requirements.

## 2. Progress and Maintenance of Data Quality

### 2.1 Updating the Data Dictionary

The central data dictionary is managed and maintained by a community driven committee with active participation and leadership from WMDA staff. However, as committee members and data projects come and go, it gets harder to sustain continuity. The DD therefore set out to do a full review of the current dictionary of data fields in order to identify areas for improvement and schedule an update of all fields as a stand-alone project. Key activities included removing duplicate fields, updating field descriptions, adding new API related fields and syncing up the data from another merged source.

The effort put in by the DD was incredible and a true testament to their community spirit. It is an arduous task with no real end as new fields are constantly being added or redefined, but this group of hardworking volunteers have proved that they are up to the task.

### 2.2 Data Dictionary RFC and XSD 2.4

The WMDA Data Dictionary (DD) Committee was established in 2018 to create a sustainable strategy ensuring effective implementation of community Requests for Change (RFCs) by identifying best practices globally. Furthermore, the committee is responsible for the data consistency across all platforms, including: EMDIS, WMDA Search & Match, WMDA forms and APIs. Recommendations made by the committee are open to public review for a month. After DD makes recommendations, the technical XSD group proceeds to update the XSD schema file with the new RFCs, and annually publish a new XSD version for community implementation.

#### Data Dictionary RFCs in 2021

**Error! Reference source not found.**1 is a summary of all RFCs the DD group worked on thus far.

Figure 1. Status Overview of All DD RFCs

Title	Document Owner	Document status	RFC #	Request Date	Request Registry	Target release
RFC-001S: Update fields to allow for reporting of both antibody & NAT testing for CMV and HIV	@ user-a8583	APPROVED	RFC-001S	12 Jul 2017	ION-9341	XSD 2.2 (2019-11)
RFC-002S Add optional value "NA"(not available)	@ Zhihong He	REJECTED	RFC-002S	20 Mar 2019	WMDA	N/A
RFC-003S Process and display "HB"(Hybrid) CBU	@ Zhihong He	CLOSED	RFC-003S	03 May 2019	ION-9431	N/A
RFC-004: HLA-E	@ Riny Heijndael	APPROVED	RFC-004S	11 Jul 2019	ION-6939	XSD 2.3
RFC-005: MICA/MICB	@ Riny Heijndael	APPROVED	RFC-005S	11 Jul 2019	ION-6939	XSD 2.3
RFC-006: KIR on allele level	@ Riny Heijndael	IN PROCESS	RFC-006S	11 Jul 2019	ION-6939	
RFC-007: more fields to distinguish IDM-tests (HBV, HCV, HEV, West Nile Virus and Chagas)	@ Ying Li	PUBLIC REVIEW	RFC-007	27 Nov 2019	ION-9341	XSD 2.4
RFC-008: Introduction of Adult Donor Cryopreserved Unit (ADCU)	@ Jürgen Sauter	IN PROCESS	RFC-008	28 May 2020	ION-4596 ION-3553	XSD 2.4
RFC-009 R2sh-CBU fields request	@ Zhihong He	PUBLIC REVIEW	RFC-009	20 Jul 2020	CBWG	XSD 2.4
RFC-010 CBU Viability for Multiple Cell Types	@ Zhihong He	PUBLIC REVIEW	RFC-010	26 Nov 2020	ION-3553 WO-1372	XSD 2.4
RFC-011 Fields for DRS	@ Ying Li	DRAFT	RFC-011	18 Oct 2021	ION-	XSD 2.5

#### XSD 2.4

The technical sub-group of DD, the XSD group, has continued their work on the improvement plan set out in 2019 by implementing RFCs approved by the DD Committee, and developing the new XSD version 2.4. This will be used as new schema for the WMDA data upload service. The updated schema was processed, finalized and published on 26 November 2021 for public review and feedback from WMDA members for one month. The final version will be released in January 2022. Detailed information about the new schema, along with a comparison to older versions, is made available on WMDA Share.

Due to the refactoring work on WMDA internal infrastructure (see D1.2 Progress report on the implementation of a secure registry-to-registry communication system) and the slow transition period experienced from member organisations, the release of this XSD schema will be delayed compared to other years. XSD group adjusted WMDA's initial release plan made in 2019 for XSD 2.2, 2.3 and 2.4, and introduced new guidelines for the development of future XSD versions.

XSD 2.4 will include the features as below:

RFC-007 will introduce a range of new fields related to infectious disease marker testing of donors at confirmatory typing stage as well as medical work-up stage. There is a standard panel of tests performed globally to ensure the safety of donor blood and blood products prior to cross-border transport. Some member organisations have started to include additional tests to the standard panel over the years which will now be added to communication messages and the option to upload the test results to WMDA.

RFC-008 will introduce a new product type to the global community in the form of Adult Donor Cryopreserved Units (ADCUs). This is a landmark event for our community as peripheral blood stem cells, bone marrow stem cells and cord blood derived stem cells have been the only options since the development of this lifesaving treatment. ADCUs present a slight challenge in the sense that they are quantified in a similar way as with CBUs (using TNC values) but that they are collected from adult donors and should therefore pass through the matching algorithm to be ranked amongst other adult donor records. No new data fields were created, instead a smart combination of donor and CBU fields were combined to define this new product in the global database.

RFC-009, or commonly described as the "ready-to-ship" (R2S) flag, will introduce an icon in CBU search result to that all release testing for the unit has been completed. Since this level of typing is not common for all units it will give R2S units a distinct advantage and patients waiting to be transplanted an shorter waiting time. Several new fields were added to report their status and quality as release ready.

Currently, some CBBs have more than one viability result depending on the cell type. However, they were only able to provide WMDA viability results for one cell type. RFC010 provides the flexibility to add

multiple viability results (i.e. per cell type: TNC, CD34PC, CD45PC) and the ability to provide additional details to the searching patient, search coordinator, or requesting transplant centre. This change would allow for each of the following per cell type TNC, CD34PC, CD45PC:

- VIABILITY
- VIABILITY\_DATE
- VIABILITY\_METHOD

## 2.3 Data Quality Webinar

The DQ project implemented key updates and features since in 2019 and continued to do so in 2021 many of which rely heavily on actions from our member organisations. In order to promote these novel features and to guide our members to implement them, WMDA hosted an educational webinar. The webinar provided a brief refresher on WMDA's DQ management cycle with a six-step roadmap that could also be applied by WMDA member organisations to improve their own DQ management.

The six-step WMDA DQ Roadmap is as following:

**Step 1, Definition:** The DD committee defines data definitions for data exchange and communication.

**Step 2, Control:** Tools like the XSD scheme, HLA-core and data validation business rules improve data quality across all registries.

**Step 3, Implementation:** The data upload for Search & Match Service validates global data and make it available to the global community

**Step 4, Improvement:** Continuous data management is a key project within WMDA to improve data available for search coordinators every day

**Step 5, Analysis:** WMDA is offering tools to registries to help them improving their internal data like DQ density report and the WMDA donor statistic report.

**Step 6, Assessment:** Both search coordinators and WMDA have constant attention as new data get introduced.

The roadmap refresher was followed by an overview of the impact that differential uploads have had on the global data quality and density of data. The webinar concluded with a guest speaker from a member organisation describing how they improved their own data quality by applying the proposed roadmap and through consistent monitoring of the monthly data density reports they receive from WMDA.

The presentation of the webinar - Data Dictionary & Differential Uploads is provided in Appendix 1.

## 2.4 Impact of ALLDIFF-Differential upload feature

In an effort to increase the data upload efficiency, WMDA has deployed and released the Differential Upload Service in 2020. This service enables organisations to only process changes in their records, instead of uploading a complete dataset of records. Differential uploads have significant impact in terms



of data processing for especially bigger organisations as they do not have to upload their entire dataset when they only send a file containing the updates to their records.

WMDA released the ALLDIFF in production in October 2020, and at the time of release only 2.44% of the global records experienced monthly change demonstrating the need for and value of the ALLDIFF feature. While the ALLDIFF data upload feature was not implemented by all listing organisations, a marked improvement in data density and quality is observed after only one year in production. Figure 2 shows the frequency of data uploads over time while Figure 3 provides a breakdown of the global database by upload frequency.

Figure 2. Monthly data upload track chart for CBU, Donor and Total from 2018-2021

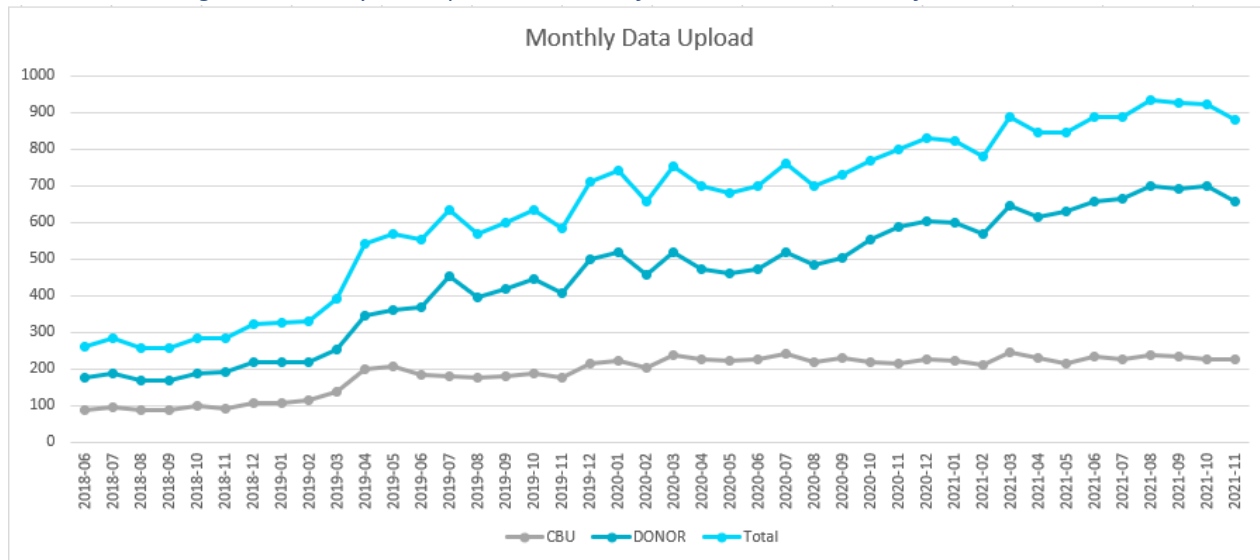
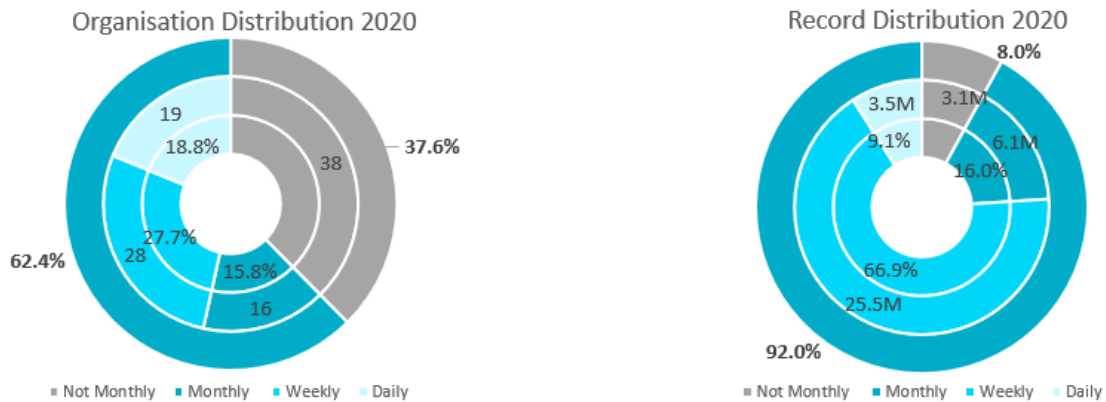
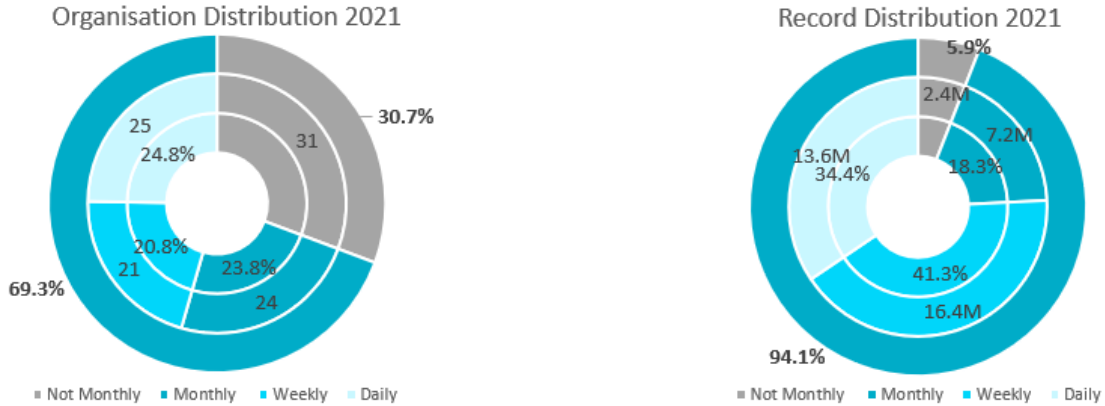


Figure 3. The Upload Frequency Distribution by Organisation and Record (DONOR and CBU) in 2020 vs. 2021





Since WMDA introduced an accreditation guidance in 2020, strongly encouraging all member organisations to have monthly uploads, almost 70% of WMDA member organisations moved to monthly, weekly, or even daily uploads, increasing data uploads significantly. In 2021, these organizations kept up this higher update frequency, with even more registries managing to improve their update frequency thanks to the ALLDIFF feature. At the time of preparing this report, there are 9 registries that have implemented differential upload with daily update. Based on the data upload dashboard, the frequency by which donor and CBU records are being uploaded, are gradually increasing and currently reach over 94% uploads monthly and 34.4% uploads daily as shown in Figure 3. The overall effect of this change in membership behaviour inches the community closer and closer to a near real-time donor database.

## 2.5 Continued monitoring towards optimization and improved data quality

### Density Improvement in 2021

At the end of 2018, WMDA introduced the Density Improvement Plan. For this plan WMDA identified a first set of data elements with high clinical impact. During 2020 and 2021 we saw continuous improvement of the data. Figure 4 shows the average density of these fields at the end of 2018, 2019, 2020 and 2021 up until the moment of writing. Some fields show significant improvement, for example ETHN, ABO, ANTI\_CMV and CBU\_SEX.

Sadly, some significant fields did not show this same change because of known difficulties. DONOR ETHN (race/ethnicity) has a strong cultural connotation and is subject to self-reporting. Donors of mixed heritage will self-report their ethnicity based on which heritage they more closely associate with which might not be reflected in their DNA. Similarly, DONOR SEX and BIRTH\_DATE are not provided by some listing organisations because of national data protection laws. WMDA is working with them to investigate any possibilities to overcome this obstacle.

Figure 4. Density improvement in 2021

Attention Time	XML Field	Data Type	XML Field Description	Percentage 2018	Percentage 2019	Percentage 2020	Percentage 2021
2019-02	ETHN	DONOR	Ethnicity	30%	56%	56%	58%
2019-02	SEX	DONOR	Sex	95%	96%	96%	96%
2019-02	ABO	DONOR	ABO	38%	43%	46%	47%
2019-02	BIRTH_DATE	DONOR	Age	95%	96%	96%	96%
2019-06	CONTACT_DATE	DONOR	Last contact date	33%	36%	37%	38%
2019-10	ANTI_CMV	DONOR	CMV status	9%	12%	14%	17%
2019-04	NMBR_TRANS	DONOR	No. of blood transfusions	6%	4%	8%	8%
2019-02	BANK_MANUF_ID	CBU	Accreditation Status	67%	76%	83%	83%
2019-02	COLL_DATE	CBU	Collection date	89%	57%	58%	61%
2019-02	SEX	CBU	Sex	78%	80%	86%	86%
2019-02	ABO	CBU	ABO	73%	73%	74%	76%
2019-04	TNC_FRZN	CBU	post processing TNC	92%	98%	99%	99%
2019-04	CD34PC_FRZN	CBU	Post processing CD34+	69%	73%	75%	75%
2019-02	ETHN	CBU	Ethnicity	37%	40%	46%	46%
2019-05	VOL_FRZN	CBU	post processing volume	82%	81%	87%	87%
2019-09	VIABILITY	CBU	Viability	36%	38%	39%	39%
2019-10	MNC_FRZN	CBU	post processing MNC	22%	21%	24%	23%
2019-11	ATT_SEG	CBU	Attached segments	33%	32%	33%	35%
2020-07	CFU_FRZN	CBU	Total colony forming units	33%	33%	36%	36%

For CBUs, both non-EU and EU members have improved BANK\_MANUF\_ID and . EU members were able to improve CD34PC\_FRZN significantly compared to non-EU members.

For Donors, both non-EU and EU members did an extraordinary job increasing the density. For CONTACT\_DATE, a good indicator of donor availability, EU members managed to significantly increase from 24.0% to 70.3% in 2020 and continue to improve to 71.1% in 2021. We see the ANTI\_CMV increased significantly after XSD 2.2 update from 2020, and data keeps well maintained in 2021.

Table 1. Density report of EU members VS non-EU members in focused fields for 2019 to 2021

Density (data available) in %												
	2019		2020		2021		2019		2020		2021	
	Non-EU	EU members	Non-EU	EU members	Non-EU	EU members	Non-EU	EU member	Non-EU	EU member	Non-EU	EU member
Data element	CBU	CBU	CBU	CBU	CBU	CBU	Donors	Donors	Donors	Donors	Donors	Donors
ABO	69.2	73.0	71.5	80.0	74.3	80.1	44.6	51.5	41.0	54.4	41.6	56.5
ATT_SEG	15.6	15.0	36.4	24.5	40.0	24.5						
BIRTH_DATE	81.3	100.0	99.1	100.0	99.2	100.0	91.3	100.0	93.4	100.0	93.4	100.0
BANK_MANUF_ID	60.7	75.0	85.2	78.8	85.2	78.8						
CD34PC_FRZN	68.2	80.6	68.4	90.7	68.9	90.4						
ANTI_CMV	12.0	11.5	4.7	13.9	4.2	15.4	4.6	16.2	7.2	25.3	9.4	28.7
COLL_DATE	36.7	33.0	60.2	51,1	64.7	51,1						

CONTACT_DATE							7.1	24.0	16.1	70.3	17.5	71.1
ETHN	20.8	29.7	52.8	29.8	52.6	30.8	35.4	34.0	54.7	59.1	54.6	64.6
GRID							89.0	94.3	91.7	99.8	92.0	99.8
NMBR_TRANS							14.6	35.3	8.4	6.5	8.3	7.5
SEX	60.8	90.1	83.2	93.0	83.3	93.0	93.5	100.0	93.4	100.0	93.4	100.0
TNC_FRZN	98.1	94.3	99.6	99.2	99.6	98.9						
VIABILITY	16.7	19.7	41.3	32.7	41.7	33.0						
VOL_FRZN	80.4	84.6	84.7	91.0	85.1	90.7						
CFU_FRZN	32.9	34.6	35.3	37.3	35.5	38.0						

### Density report per registry Improvement

To gain a better insight into the data quality of individual donor registries and cord blood banks, WMDA provides monthly Data Quality (DQ) reports for both internal and public evaluation and review. This DQ Programme is an initial version of an individual DQ report per organisation provided and distributed by WMDA. The DQ programme was evaluated with the feedback from registries and is officially distributed on 15<sup>th</sup> of each month starting from April 15<sup>th</sup>, 2020. Detail information can be found in the public [DQ report User Guide](#).

By creating the DQ report, WMDA can get a more accurate overview on the Qualitative Distribution report. As a result of this report, CBBs have better overview on the quality of their databases and can initiate corrections quicker and easier if necessary. This report is now available for more useful and important fields listed in Table 2.

*Table 2. List of fields that are qualitatively monitored for CBU in DQ density report*

Field name	Explanation	Normal Value Range
TNC_FRZN	Total Nucleated Cells in CBU post processing/prior to cryopreservation	$50 \times 10^7$ to $300 \times 10^7$
CD34PC_FRZN	Total number of CD34+ cells (post processing, prior to cryopreservation)	$1 \times 10^6$ to $20 \times 10^6$
RED_BC_FRZN	Total number of nucleated red blood cells (post processing, prior to cryopreservation)	$1 \times 10^7$ to $100 \times 10^7$
CFU_FRZN	Total count of colony forming units (post processing, prior to cryopreservation)	$1 \times 10^5$ to $70 \times 10^5$
VOL_FRZN	Total volume frozen (post processing, prior to cryopreservation) in ml	About 25 or 50
VIABILITY	A calculated score based on specific test in % for TNC_FRZN, CD34PC_FRZN or CD45PC_FRZN	80-100

The significant importance of the Data Quality Qualitative Distribution report is demonstrated using Figure 5. This report shows registry/CBB upload data for several registries and CBBs. As a result of the report, registries are now able to identify the issues listed below:

1. There are 10 records still missing TNC-FRZN from CBB with WO-ID 1111.
2. There are 2 records for CBB with WO-ID 2222 with really high TNC-FRZN in the range of [700-3000], so there might be a mistake in the data.
3. There are CBBs that are missing a WO-ID to identify if the cord blood units are listed in an accredited cord blood bank.
4. For the CBUs without a WO-ID, the quality may be bad with as many as 634 records with a TNC\_FRZN value less than 90.

Figure 5. The example of TNC\_FRZN Qualitative Distribution report

## TNC\_FRZN

Organisation	[1-49]	[50-89]	[90-124]	[125-149]	[150-199]	[200-249]	[250-299]	[300-699]	[700-3000]	Missing
1111	0	2	2329	3114	3515	1067	281	119	0	10
2222	15	799	3368	2111	2126	599	132	28	2	0
3333	54	1385	3022	1272	981	192	32	8	1	1
Missing	304	330	207	609	351	31	3	1	1	2

To demonstrate the value of the DQ report to listing organisations, one the largest Southern European registries presented how they implemented their own DQ Programme based on these monthly reports.

Figure 7. Data Upload processing report of ION 7358 in February 2021

```
File: ION-7358-D.zip.pgp 2021-02-12 12:51:37
Pool(s): 7358
Content Type: D
Update Mode: FULL
Start processing: 2021-02-12 13:00:00
Schema version: 2.1
Total records processed: 395886
Total records with warnings: 157420
Total records rejected: 0
Total valid records: 395886
```

Figure 8. Data Upload processing report of ION 7358 in March 2021

```
File: ION-7358-D.zip.pgp 2021-03-09 18:05:32
Pool(s): 7358
Content Type: D
Update Mode: FULL
Start processing: 2021-03-09 18:20:00
Schema version: 2.1
Total records processed: 401218
Total records with warnings: 2
Total records rejected: 0
Total valid records: 401218
```

### Deprecated Code

HLA coding used in the data upload is validated by an HLA validation engine. Since the nature of some of these codes is volatile, codes may get deprecated quarterly after a new release of the HLA nomenclature.

In 2019, the data upload service deployed the Deprecated Code handling to provide suggested replacement code of the deprecated code. In 2021, WMDA sees this feature works as expected and is extremely helpful quarterly when there is a new nomenclature release that introduces new deprecated code. In the data upload report, the replacement suggestion will be provided in time. As in the case of the Israeli registry, Bone Marrow and Cord Blood Bank Registry Hadassah University Hospital. This saves time and effort for the organisation to further check the replacement by their own.

Deprecate codes must be substituted by an appropriate replacement, but obviously this cannot come into effect worldwide at the same point in time. Hence, grace periods of 1 year have been defined in the guidelines to avoid the frequently observed immediate rejection of renamed alleles and deprecated codes in data exchange.

### XSD 2.2, 2.3 End User Usage Status

XSD 2.2 was released on 16 December 2019. After the XSD 2.2 release, WMDA updated the data dictionary and data upload application with this latest version.

XSD 2.3 was released on 18 Jan 2021. After the XSD 2.3 release, WMDA updated the data dictionary and data upload application with this latest version.

Figure 9. Handling of multiple concurrent versions of the XSD schemas

Handling :

1. Change XSD and change delivered XML
2. Create XSL that arranges version changes (courtesy)

Preparation: dd=data dictionary committee xsd=technical xsd group	Valid: october	Grace period: move to the new version	Invalid: not supported anymore
---	----------------	---------------------------------------	--------------------------------

Ver	Proposed Changes	2019				2020				2021				2022				2023			
		Q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4
2.1	RFC xxx																				
2.2	<ul style="list-style-type: none"> <li>RFC 1</li> <li>clean up description</li> </ul>	dd	dd	dd	Oct																
2.3	<ul style="list-style-type: none"> <li>GRID/ID</li> <li>RFC 4, 5</li> </ul>				dd	dd	dd	dd	Oct												
2.4	<ul style="list-style-type: none"> <li>RFC 7, 8, 9, 10</li> </ul>								dd	dd	dd	dd	Oct								
2.5	<ul style="list-style-type: none"> <li>new RFC XX</li> </ul>																				

Figures 10 and 11 provides a summary of WMDA members that use XSD schema 2.2 compared to the members who have switched to XSD schema 2.3 to generate the XML file for data uploads. We can see that 12.9% organizations in WMDA have currently implemented XSD 2.3, compared to 18.8% for EU members. Currently 75.7% of all records in WMDA are uploaded using XSD 2.2 or 2.3. As shown in Figure 9, XSD schema 2.1 became invalid in October 2021 and is therefore no longer being supported, necessitating listing organisations to update to the newer 2.2 or even the latest 2.3 schema.

Figure 10. XSD 2.2 usage Distribution by Organisation and Records till November 2020 (DONOR and CBU)

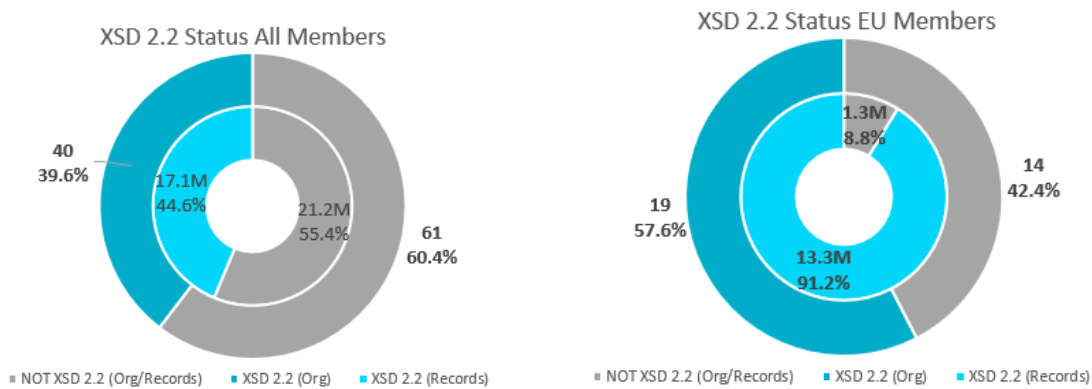
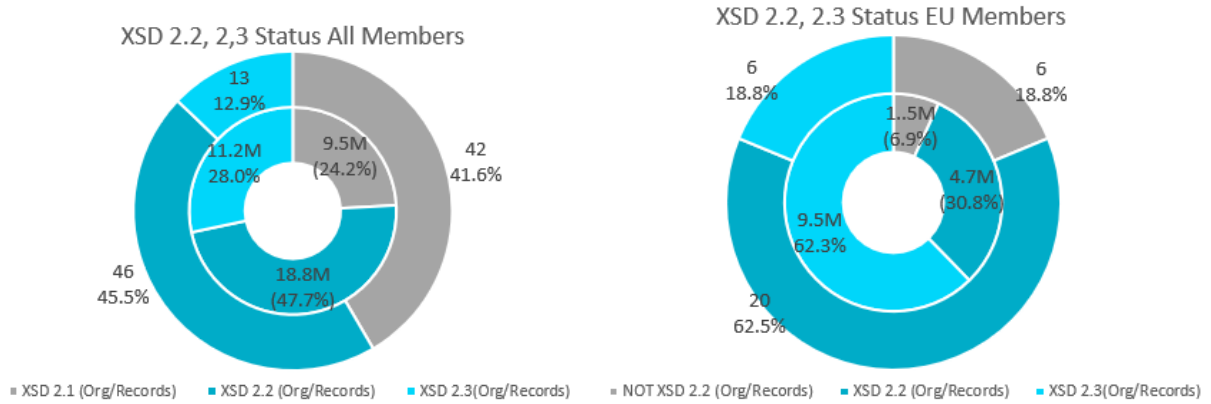


Figure 11. XSD 2.3 usage Distribution by Organisation and Records till December 2021 (DONOR and CBU)





### 3. Beyond 2021

As with most other things, what you get out is only ever as good as what you've put in, and data is no exception. For a long time, there has been a drive towards increasing the quantity of records uploaded to the WMDA global database, however what value does a record hold for a patient when the data is of such poor quality that a Search Coordinator cannot recommend it for transplantation. The DQ initiatives taken by WMDA since 2019 has seen that drive redirected towards quality instead. Not just of the individual records, but also in how the data is stored and processed.

#### 3.1 Gap Analysis – Comparing live registry data with uploaded data

Through collaboration with one of our largest member organisations, National Marrow Donor Programme (NMDP), we noticed a variation of 1% - 10% in the data uploaded to WMDA and live registry data being mirrored to NMDP. Having multiple sets of the same data is problematic enough without it being out of sync too, so WMDA set out to get to the bottom of it.

NMDP receives mirrored data from two-thirds of the global database and thus serves as a good partner for this gap analysis. This mirroring of data has its origins in EMDIScord, six member organisations that share real-time CBU data with one another, enabling upfront search of that subset of records offering some business value.

Initial findings showed that the CBU data uploaded to WMDA from CBBs are more complete than what they send to NMDP. Needless to say, these findings negate the value of CBU data mirroring when more complete data can be retrieved from WMDA Search & Match Service. However, it appears that the opposite is true of donor data where initial findings show that mirrored real-time data is often more complete and more frequently updated than what is available in WMDA.

In 2022 this gap analysis will continue by specifically approaching listing organisations that infrequently uploads incomplete data to WMDA, working with them to implement the ALLDIFF feature and improving the quality of data they share with the community.

#### 3.2 HLAcore

An international agreement about validation and interpretation of HLA-typing data is indispensable for their electronic exchange via computer networks and their automated processing. The WMDA HLA Nomenclature Guidelines are the international consensus used by the worldwide registry community comprising the WMDA membership. The application currently used by WMDA, and the membership as single HLA validation repository was developed and is maintain by the German National Bone Marrow Donor Registry (ZKRD) and is called HLAcore. This application derives its core HLA data from <http://hla.alleles.org/wmda/> and MAC code data from <https://hml.nmdp.org/MacUI/> where data is updated frequently.

Unfortunately, this application is not open source, difficult to implement into modern software applications and considered a bit of "black box" in terms of its operation. The WMDA therefore set-out to develop an open-source version available for its own internal use, for implementation by WMDA members and ultimately publicly available for all industry related uses.

This open-source repository will have to be updated, maintained and supported indefinitely by the WMDA team as novel HLA typing techniques continue to expand on the nuances of HLA antigens and their alleles.

### 3.3 Leveraging machine learning to advance global database

The stem cell transplant industry is no stranger to advanced algorithms and the WMDA has both implemented and developed many over the years in an effort to reduce the timeline to transplantation and improve donor match prediction. However, algorithms are somewhat static in nature which is in stark contrast to the dynamic genetic fields we operate in.

WMDA will be working with a student from the Delft University of Technology to investigate machine learning solutions based on the vast quantity of data we have access to as part of his master's dissertation. Two projects are being considered:

1. Using machine learning to improve the matching of stem cell donors to patients. The matching is currently done based on a matching score that is calculated by a classical algorithm. The machine learning algorithm could be used for improving the matching score or as secondary advice for matching.
2. Using a machine learning algorithm to calculate a risk factor for a stem cell donor. There is always risk involved in donating stem cells and this can differ based on a lot of factors like the age, location, and BMI of the donor. This risk factor can then be used to choose which donor to use when there are multiple similar options for donating.

Like any other research project based on data from our membership, WMDA will seek approval from its members prior to the start of the project. Depending on which project is selected, data will be used from the donor records, CBU records, ADCU records, global trends report and/or S(P)ear reporting data. Ideally the end product will pave the way for continued advancement in this direction while also providing the community with immediate benefit.

## WMDA Data Quality

March 22nd, 2021  
Zhihong He, WMDA

WMDA Virtual Meetings Week 2021

### Outline

- Why is data quality important?
- What is WMDA role in improving data quality?
- How can you improve data quality in your registry?
- What do our members do in the area of data quality?
  - Portugal
- What does WMDA do in the area of data quality?

WMDA Virtual Meetings Week 2021

## Why is data quality important?

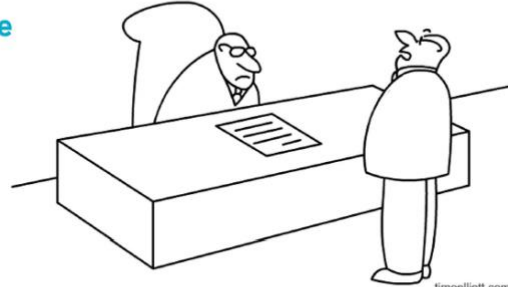
Accurate, reliable data is important to identify a potential match

### Key figures of WMDA inventory database

- 2+ years extended data fields
- 2+ years validation control
- 92% records monthly upload

### Current situation

- Check on completeness of data
- Check on accuracy of data
- Strive to complete match lists
- Data report is to help you checking the data



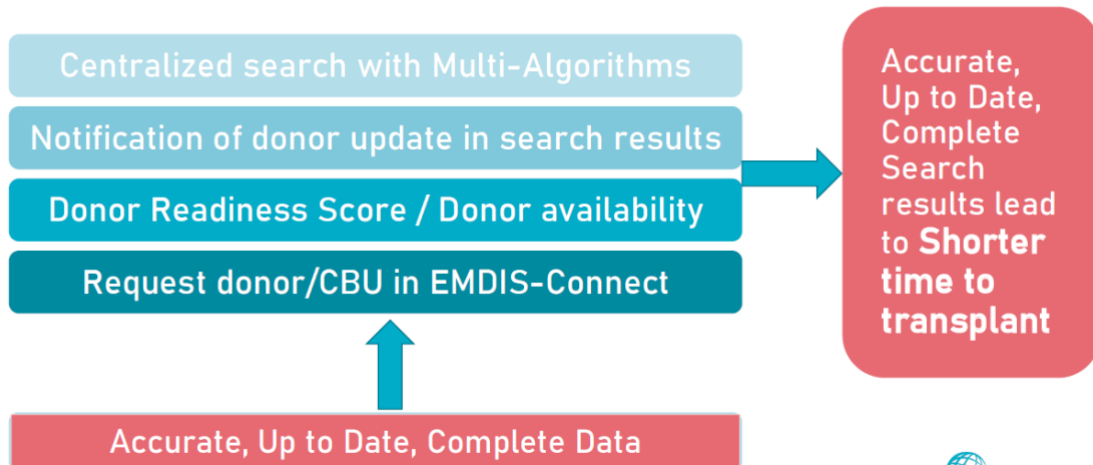
*"Yes sir, you can absolutely trust those numbers"*

➔ Data is not good enough

WMDA Virtual Meetings Week 2021

## Why Data Quality?

### Vision of Search, Match & Connect/EMDIS



WMDA Virtual Meetings Week 2021

## Data quality achievements WMDA Data Improvement life cycle



WMDA Virtual Meetings Week 2021



## DQ achievements Data Dictionary RFCs to XSD schema



- RFC08 Adult Donor Cryopreserved Unit
- RFC09 Ready to ship cord blood unit
- RFC10 CBU Viability for Multiple Cell Types



- Single source of truth of data validation for different domains

### • XSD update schedule

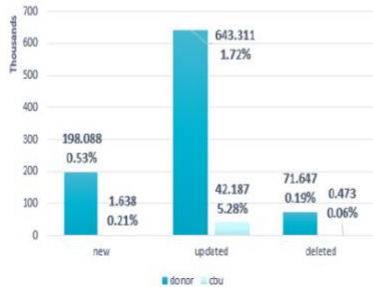
Preparation Valid Grace period Invalid

Ver	Proposed Changes	2019				2020				2021				2022				2023			
		Q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4	q1	q2	q3	q4
2.1	RFC xxx																				
2.2	<ul style="list-style-type: none"> <li>• RFC 1</li> <li>• clez</li> </ul>	dd	dd	dd	Oct																
<p><b>Reminder</b> Registries that use XSD 2.1 please plan to upgrade to XSD 2.2 or XSD 2.3 before October 1</p>																					
2.3	<ul style="list-style-type: none"> <li>• GRI</li> <li>• RFC</li> </ul>						xso	xso													
2.4	<ul style="list-style-type: none"> <li>• RFC 7, 8, 9, 10</li> </ul>									dd	dd	dd	dd	Oct							
2.5	<ul style="list-style-type: none"> <li>• new RFC XX</li> </ul>											xsd	xsd								

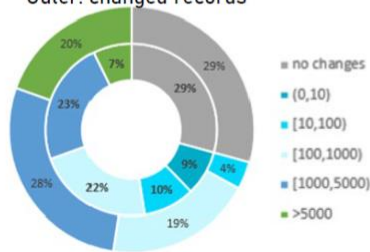
## DQ achievements Differential upload & data differential check

- Differential upload (DIFF mode) (>100K or >100 monthly changes)
- Differential data check (internal validation by WMDA)

Number of donor/cbu records that change



Percentage of Organisations with Records Changes Monthly  
Inner: new records,  
Outer: changed records



WMDA Virtual Meetings Week 2021

Date Last File  
Latest records update date

DNA Class I	DNA Class II	Date Last File
14,542	14,898	2019-03-06
360,034	368,470	2021-03-21
119	119	2020-08-13
50,450	50,244	2021-03-16
77,161	77,161	2021-03-21
13,318	14,170	2020-02-26
381	382	2016-11-18
3663	4301	2021-03-19
4689	4689	2021-03-15
68,396	69,037	2021-03-18

**WMDA**  
matching donors • serving patients

## DQ achievements Automated monthly density report

- Poll: Have you ever used the report?
- Let us go through a [real report from NMDP](#) as an example

Density in new fields

Last 12 months density trend

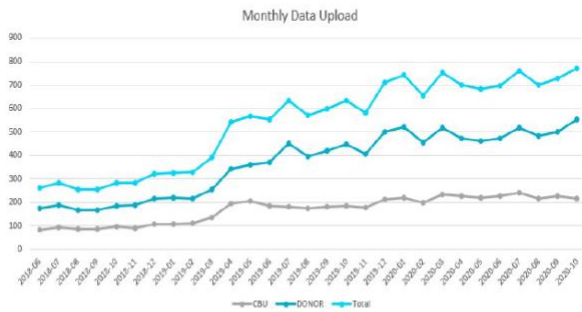
Distribution of important fields

WMDA Virtual Meetings Week 2021

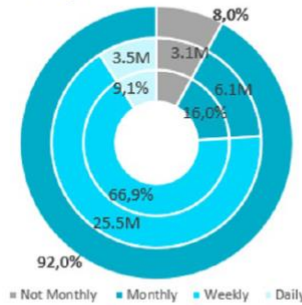
**WMDA**  
matching donors • serving patients

## Data quality achievements - update frequency

Number of data uploads are increasing over time



70% of the donors are updated weekly



WMDA Virtual Meetings Week 2021

## WMDA is working actively on the following data fields

Attention Time	XML Field	Data Type	XML Field Description	Percentage 2018	Percentage 2019	Percentage 2020
2019-02	ETHN	DONOR	Ethnicity	30%	56%	56%
2019-02	SEX	DONOR	Sex	95%	96%	96%
2019-02	ABO	DONOR	ABO	38%	43%	46%
2019-02	BIRTH_DATE	DONOR	Age	95%	96%	97%
2019-06	CONTACT_DATE	DONOR	Last contact date	33%	36%	37%
2019-10	ANTI_CMV	DONOR	CMV status	9%	12%	14%
2019-04	NMBR_TRANS	DONOR	No. of blood transfusions	6%	4%	8%
2019-02	BANK_MANUF_ID	CBU	Accreditation Status	67%	76%	83%
2019-02	COLL_DATE	CBU	Collection date	89%	57%	58%
2019-02	SEX	CBU	Sex	78%	80%	86%
2019-02	ABO	CBU	ABO	73%	73%	74%
2019-04	TNC_FRZN	CBU	post processing TNC	92%	98%	99%
2019-04	CD34PC_FRZN	CBU	Post processing CD34+	69%	73%	75%
2019-02	ETHN	CBU	Ethnicity	37%	40%	46%
2019-05	VOL_FRZN	CBU	post processing volume	82%	81%	87%
2019-09	VIABILITY	CBU	Viability	36%	38%	39%
2019-10	MNC_FRZN	CBU	post processing MNC	22%	21%	24%
2019-11	ATT_SEG	CBU	Attached segments	33%	32%	33%

Fields commonly raise issues:

- TNC\_FRZN
- CD34PC\_FRZN
- GRID
- HLA

## Data quality achievements – What can our members do?

### Data uploader

- Implement an automated upload
- Implement the differential upload – send only updated/new records
- Check your upload report – solve the issues
- Check the update date <https://statistics.wmda.info/>
- Check your own density report

### Search coordinator

- Report any data issue to WMDA
  - HLA discrepancy
  - HLA deprecated code
  - Suspicious value of CBU or DONOR records  
e.g. TNC\_FRZN, CD34PC\_FRZN, Duplicate GRID, donor age, file upload date

WMDA Virtual Meetings Week 2021



## Member experience in Data Quality – Portugal

### Search Factors Submitted by CEDACE

#### Up To December 2020

##### Demographics:

- Donor GRID
- Birth Date
- Gender
- Status

##### Genotype for:

- HLA – A
- HLA – B
- HLA – C
- HLA – DRB1

#### From January 2021

##### Demographics:

- Donor GRID
- Birth Date
- Gender
- Status

##### Genotype for:

- HLA – A, B, C
- HLA – DRB1/3/5
- HLA – DQA1/B1
- HLA – DPA1/B1
- ABO Rh

##### Allosensitization events:

- Pregnancies
- Transfusions
- Anti-CMV Status
- **#Previous Donations?**

WMDA Virtual Meetings Week 2021





## Member experience in Data Quality – Portugal Improve the data according to upload report

```

23712_JON-7358-D.zio.pgp-pickedup-20210212-12h00m-processed-20210212...
File: IDN-7358-D.zio.pgp 2021-02-12 12:51:37
Pool(s): 7358
Content Type: D
Update Mode: FULL
Start processing: 2021-02-12 13:00:00
Schema version: 2.1
Total records processed: 395866
Total records with warnings: 157420
Total records rejected: 0
Total valid records: 395866

W | 7358 | C3050041415 | 735807C305004141525 | NBR_PREG (Warning) Number of pregnancies
must be empty/blank for male donors.
W | 7358 | C3050041419 | 735807C305004141926 | HLA: #Error at locus DRB3=: invalid HLA
code DRB3=52:XX. locus blanked.
W | 7358 | C3050041423 | 735807C305004142327 | NBR_PREG (Warning) Number of pregnancies
must be empty/blank for male donors.
W | 7358 | C3050027775 | 735807C305002777521 | IDN_OBV_DATE (Warning) OBV date is not
provided
W | 7358 | C3050027775 | 735807C305002777521 | IDN_OBV_DATE (Warning) OBV result can not be
provided without a OBV date
W | 7358 | C3050041404 | 735807C305004140400 | HLA: #Error at locus DRB4=: invalid HLA
code DRB4=43:XX. locus blanked.
W | 7358 | C3050041414 | 735807C305004141400 | HLA: #Error at locus DRB4=: invalid
code DRB4=52:XX. locus blanked. #error a
25467_JON-7358-D.zio.pgp-pickedup-20210300-18h20m-processed-20210300-18h20...
File: IDN-7358-D.zio.pgp 2021-03-09 18:05:32
Pool(s): 7358
Content Type: D
Update Mode: FULL
Start processing: 2021-03-09 18:20:00
Schema version: 2.1
Total records processed: 401218
Total records with warnings: 2
Total records rejected: 0
Total valid records: 401218

W | 7358 | C306000223 | 735807C30600022322 | HLA: #Error at locus DRB3=: invalid
HLA code DRB3=53:XX. locus blanked.
W | 7358 | C306133851 | 735807C30613385113 | HLA: #Error at locus DPB1=: invalid
HLA code DPB1=04:DDK5. locus blanked.

Processing finished at: 2021-03-09 18:26:08
Total processing time: 6 minutes.

24738_JON-7358-D.zio.pgp-pickedup-20210226...
File: IDN-7358-D.zio.pgp 2021-02-26 12:42:06
Pool(s): 7358
Content Type: D
Update Mode: FULL
Start processing: 2021-02-26 12:50:00
Schema version: 2.1
Total records processed: 395853
Total records with warnings: 0
Total records rejected: 0
Total valid records: 395853

Processing finished at: 2021-02-26 18:56:27
Total processing time: 365 minutes.
    
```

WMDA Virtual Meetings Week 2021




## Exercise example to fix a simple Data Quality issue

Donor Age	Count
61	19,257
62	7,459
63	5,939
64	4,927
65	887
66	376
67	200
68	44
69	16
70	14
71	10
72	5
73	7
74	3
76	2
Total	39,146


- Report by search coordinator to WMDA
  - Donor age range issue check
- Defined validation: 18 - 60
- What WMDA implemented: 16 - 80
- WMDA is checking with registries sending donors >60 years old
- WMDA will update business rule based on feedback of registries
- Implement new rule

WMDA Virtual Meetings Week 2021





WMDA Virtual Meetings Week 2021



## Appendix 2

Full donor and CBU data density report comparing EU states with non-EU states up until November 2021. For the 2019 and 2020 data, please refer to the 2019 D1.2 and 2020 D1.1 Progress report on the accuracy, quality of data in the global database report. (<https://share.wmda.info/x/gRrkEw>).

Density (data available) in %				
Data element	Non-EU		EU member states	
	Cord Blood Units	Cord Blood Units	Donors	Donors
A1	24.1	52.2	10.4	18.3
A2	21.1	46.0	9.0	15.7
ABO	74.3	80.1	41.6	56.5
AL_PLA	2.7	15.7		
AL_RED_BC	0.5	4.1		
AL_SER	0.1	8.8		
ALT	0.0	0.2	0.0	0.8
ANTI_CMV	4.2	15.4	9.5	28.7
ANTI_CMV_DATE	4.1	15.4	9.5	28.7
ANTI_HBC	2.3	22.5	0.1	3.3
ANTI_HBS	0.1	1.7	0.0	0.3
ANTI_HCV	0.6	29.5	0.1	4.1
ANTI_HIV_12	0.6	22.0	0.0	3.8
ANTI_HTLV	2.6	17.4		1.2
ATT_SEG	40.0	24.5		
B1	24.0	52.1	10.4	18.3
B2	22.9	48.7	9.6	16.9
BACT_CULT	57.1	49.9		
BAG_ID	17.6	16.4		
BAG_TYPE	15.2	47.2		
BAGS	19.0	22.3		
BANK_DISTRIB_ID	42.2	78.8		

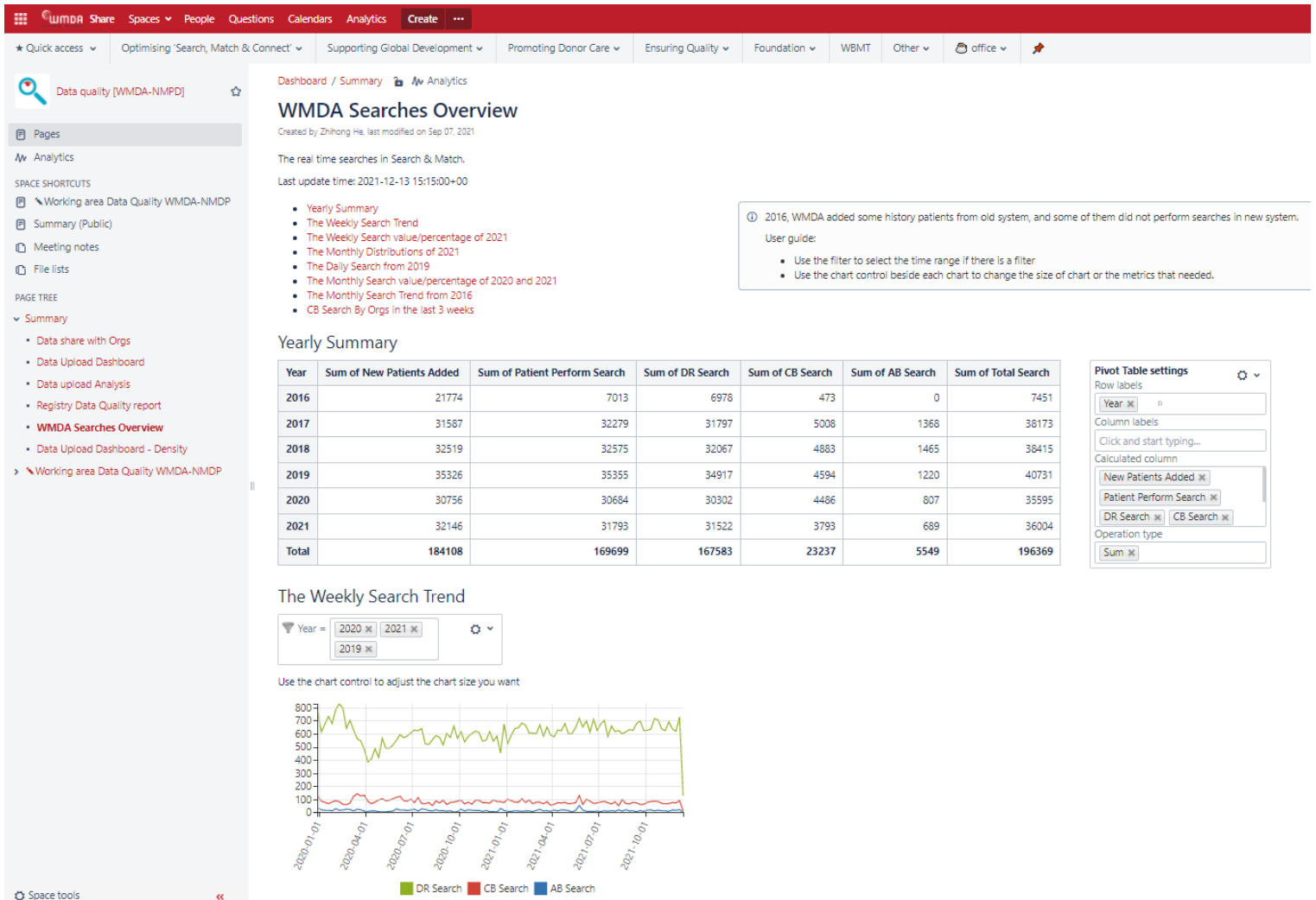
BANK_DISTRIB_ID_EM				
DIS	5.2	22.1		
BANK_DISTRIB_ID_W				
MDA	42.2	78.8		
BANK_MANUF_ID	85.2	78.8		
BANK_MANUF_ID_EM				
DIS	11.8	22.1		
BANK_MANUF_ID_W				
MDA	85.2	78.8		
BANK_MAT_ID	0.1	1.0		
BIRTH_DATE	99.2	100.0	100.0	100.0
C1	0.8	13.5	0.7	7.8
C2	0.4	7.4	0.3	3.9
CCR5	0.1	10.1	5.9	38.4
CD34PC	0.1	12.6		
CD34PC_FRZN	68.9	90.4		
CFU_FRZN	35.6	38.0		
CHAGAS	0.1	4.1		0.0
CHECKUP_DATE			0.8	9.1
CMV_NAT	0.0	8.6	0.0	0.2
CMV_NAT_DATE	0.0	8.6	0.0	0.2
COLL_DATE	64.7	51.0		
COLL_TYPE			3.2	5.6
CONTACT_DATE			17.5	71.1
CT_COMPLETE_DATE	0.2	9.5		
CT_SMPL_TYPE	0.2	9.5		
DNA_A1	84.9	86.4	92.0	91.6
DNA_A2	80.2	76.9	88.3	82.2
DNA_B1	85.0	86.5	92.0	91.6
DNA_B2	82.6	82.0	90.3	87.5
DNA_C1	40.6	43.7	55.0	82.1
DNA_C2	39.2	40.5	53.5	76.6
DNA_E1	0.0	0.0		
DNA_E2	0.0	0.0		
DNA_SMPL	3.7	24.9		
DON_ATTR	34.2	7.6	38.1	63.9
DON_POOL	100.0	100.0	100.0	100.0
DPA11	1.7	4.0	8.7	8.6
DPA12	1.2	2.3	8.2	8.2
DPB11	9.4	9.7	34.5	62.3
DPB12	8.6	8.1	32.8	52.7
DQ1	0.2	7.2	1.0	3.2
DQ2	0.2	6.2	0.9	2.6
DQA11	2.8	4.7	9.0	10.2
DQA12	2.6	4.2	8.8	9.6
DQB11	16.6	33.0	46.4	75.1
DQB12	15.8	28.5	44.8	69.4
DR1	19.5	42.3	4.7	11.5
DR2	18.2	38.4	4.3	10.3
DRB11	90.1	98.6	95.4	95.2
DRB12	86.7	91.5	92.9	89.1
DRB31	10.9	6.4	23.2	17.4
DRB32	2.2	1.5	7.2	4.0
DRB41	9.1	4.4	18.3	11.6
DRB42	1.1	0.9	5.2	0.8
DRB51	6.5	2.8	15.4	8.3
DRB52	1.0	0.7	4.7	0.7
EBV	0.0	9.7	0.0	0.3
ETHN	52.6	30.8	54.6	64.6
FREEZE_DATE	54.9	45.6		
FREEZE_METH	15.8	39.1		
FUNG_CULT	57.1	50.6		
GRID			92.0	99.8
HBS_AG	1.4	25.5	0.1	4.1
HBV_NAT	0.3	12.9	0.0	0.7
HCV_NAT	2.0	21.3	0.0	0.8

HEIGHT			15.2	14.0
HEMO_STATUS	48.0	26.9		
HIV_1_NAT	2.0	19.2	0.0	0.7
HIV_P24	0.5	15.0	0.0	1.3
ID	100.0	100.0	82.3	21.4
KIR2DL1	0.0	0.0	4.0	30.6
KIR2DL2	0.0	0.0	4.0	30.6
KIR2DL3	0.0	0.0	4.0	30.6
KIR2DL4	0.0	0.0	4.0	30.6
KIR2DL5A	0.0	0.0	4.0	30.6
KIR2DL5B	0.0	0.0	4.0	30.6
KIR2DP1	0.0	0.0	4.0	30.6
KIR2DS1	0.0	0.0	4.0	30.6
KIR2DS2	0.0	0.0	4.0	30.6
KIR2DS3	0.0	0.0	4.0	30.6
KIR2DS4	0.0	0.0	4.0	30.6
KIR2DS5	0.0	0.0	4.0	30.6
KIR3DL1	0.0	0.0	4.0	30.6
KIR3DL2	0.0	0.0	4.0	30.6
KIR3DL3	0.0	0.0	4.0	30.6
KIR3DP1	0.0	0.0	4.0	30.6
KIR3DS1	0.0	0.0	4.0	30.6
KIR_GLS_URI	0.0			
LOCAL_ID	23.8	44.7		
MAT_A1	0.0	2.9		
MAT_A2	0.0	2.5		
MAT_AL_PLA	2.6	12.7		
MAT_AL_SER	0.1	14.1		
MAT_ALT	0.0	1.1		
MAT_ANTI_CMV	30.8	38.3		
MAT_ANTI_CMV_DATE	13.8	5.2		
MAT_ANTI_HBC	16.7	40.8		
MAT_ANTI_HBS	0.0	2.7		
MAT_ANTI_HCV	47.1	53.6		
MAT_ANTI_HIV_12	19.5	50.2		
MAT_ANTI_HTLV	45.4	30.3		
MAT_B1	0.0	2.9		
MAT_B2	0.0	2.7		
MAT_C1	0.0	0.8		
MAT_C2	0.0	0.6		
MAT_CHAGAS	25.1	4.5		
MAT_CMV	31.1	38.3		
MAT_CMV_DATE	13.8	5.2		
MAT_CMV_NAT	0.0	0.2		
MAT_CMV_NAT_DATE	0.0	0.0		
MAT_DNA_A1	2.6	8.4		
MAT_DNA_A2	2.5	7.8		
MAT_DNA_B1	2.6	8.4		
MAT_DNA_B2	2.5	8.0		
MAT_DNA_C1	1.3	2.2		
MAT_DNA_C2	1.2	2.0		
MAT_DNA_E1	0.0	0.0		
MAT_DNA_E2	0.0	0.0		
MAT_DPA11	0.1	0.0		
MAT_DPA12	0.1	0.0		
MAT_DPB11	0.1	0.8		
MAT_DPB12	0.1	0.6		
MAT_DQ1	0.0	1.0		
MAT_DQ2	0.0	0.9		
MAT_DQA11	0.1	0.0		
MAT_DQA12	0.1	0.0		
MAT_DQB11	0.3	2.1		
MAT_DQB12	0.2	1.9		
MAT_DR1	0.0	2.8		
MAT_DR2	0.0	2.4		
MAT_DRB11	2.5	4.7		

MAT_DRB12	2.5	4.2		
MAT_DRB31	0.1	0.0		
MAT_DRB32	0.0	0.0		
MAT_DRB41	0.0	0.0		
MAT_DRB42	0.0	0.0		
MAT_DRB51	0.0	0.0		
MAT_DRB52	0.0	0.0		
MAT_EBV	0.0	17.6		
MAT_HBS_AG	17.3	60.0		
MAT_HBV_NAT	9.5	31.5		
MAT_HCV_NAT	17.1	37.4		
MAT_HIV_1_NAT	16.6	35.7		
MAT_HIV_P24	6.0	11.6		
MAT_ID	39.5	6.6		
MAT_MICA		0.0		
MAT_MICB		0.0		
MAT_PB19_NAT	0.0	0.0		
MAT_PLA_QUANT	2.6	11.8		
MAT_SER_QUANT	0.4	12.4		
MAT_SYPHILIS	29.6	56.2		
MAT_TOXO	1.1	23.0		
MAT_WNV	26.3	0.7		
MICA	0.0	0.0		
MICB	0.0	0.0		
MNC_FRZN	16.8	39.5		
NMBR_MARR			17.8	82.8
NMBR_PBSC			17.8	82.8
NMBR_PREG			0.5	4.6
NMBR_TRANS			8.3	7.5
OTH_SIMPL	2.2	18.8		
PB19_NAT	0.0	12.6	0.0	0.5
PLA_QUANT	2.6	11.5		
PROC_DATE	51.0	50.2		
PROC_METH	13.3	63.2		
PROC_METH_TYPE	26.5	48.3		
PROD_MOD	47.5	42.4		
RED_BC_FRZN	30.7	10.5		
RHESUS	74.3	80.1	40.9	55.9
RSV_PAT	0.0	0.0	0.0	0.0
SER_QUANT	0.0	8.9		
SEX	83.3	93.0	93.5	100.0
STAT_END_DATE	0.0	0.3	0.1	0.6
STAT_REASON	0.0	0.5	0.0	0.3
STATUS	100.0	100.0	100.0	100.0
SYPHILIS	2.2	23.6	0.1	4.0
TNC	13.5	34.1		
TNC_FRZN	99.6	98.9		
TOXO	0.0	12.4	0.0	0.3
VIABILITY	41.7	32.9		
VIABILITY_CELLS	15.8	9.2		
VIABILITY_DATE	12.8	24.9		
VIABILITY_METHOD	42.6	9.1		
VOL	42.1	81.1		
VOL_FRZN	85.1	90.7		
WEIGHT			19.1	17.1
WNV	0.0	0.0	0.0	0.2

## Appendix 3

The Search & Match dashboard to monitor the impact of COVID-19. It is available in the member access share page <https://share.wmda.info/x/SkuOF>.



**WMDA Searches Overview**  
Created by Zhihong He, last modified on Sep 07, 2021

The real time searches in Search & Match.  
Last update time: 2021-12-13 15:15:00+00

- Yearly Summary
- The Weekly Search Trend
- The Weekly Search value/percentage of 2021
- The Monthly Distributions of 2021
- The Daily Search from 2019
- The Monthly Search value/percentage of 2020 and 2021
- The Monthly Search Trend from 2016
- CB Search By Orgs in the last 3 weeks

2016, WMDA added some history patients from old system, and some of them did not perform searches in new system.  
User guide:

- Use the filter to select the time range if there is a filter
- Use the chart control beside each chart to change the size of chart or the metrics that needed.

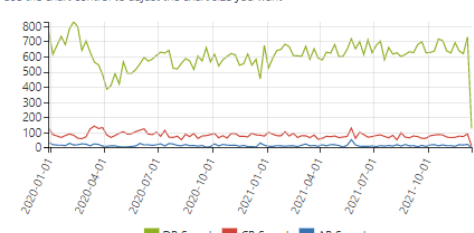
**Yearly Summary**

Year	Sum of New Patients Added	Sum of Patient Perform Search	Sum of DR Search	Sum of CB Search	Sum of AB Search	Sum of Total Search
2016	21774	7013	6978	473	0	7451
2017	31587	32279	31797	5008	1368	38173
2018	32519	32575	32067	4883	1465	38415
2019	35326	35355	34917	4594	1220	40731
2020	30756	30684	30302	4486	807	35595
2021	32146	31793	31522	3793	689	36004
<b>Total</b>	<b>184108</b>	<b>169699</b>	<b>167583</b>	<b>23237</b>	<b>5549</b>	<b>196369</b>

**The Weekly Search Trend**

Year = 2020 | 2021 | 2019

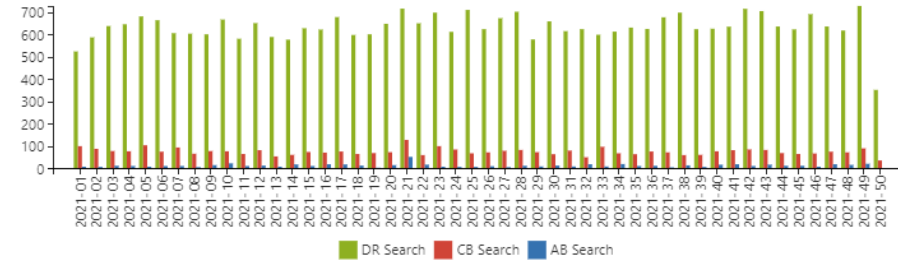
Use the chart control to adjust the chart size you want



Data quality [WMDA-NMPD]

- Pages
- Analytics
- SPACE SHORTCUTS
- Working area Data Quality WMDA-NMPD
- Summary (Public)
- Meeting notes
- File lists
- PAGE TREE
- Summary
  - Data share with Orgs
  - Data Upload Dashboard
  - Data upload Analysis
  - Registry Data Quality report
  - WMDA Searches Overview
  - Data Upload Dashboard - Density
- Working area Data Quality WMDA-NMPD

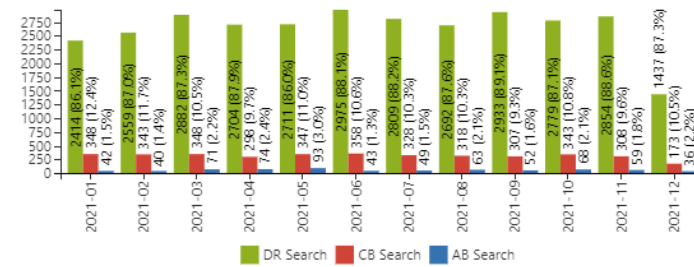
### The Weekly Search value/percentage of 2021



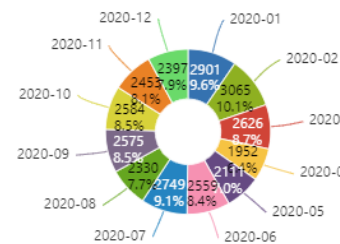
### The Monthly Distributions of 2021

Year Month =

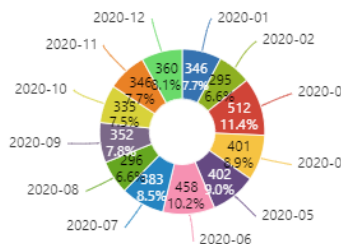
Select the months to compare.



#### DR Search



#### CB Search



Data quality [WMDA-NMPD]

Pages

Analytics

SPACE SHORTCUTS

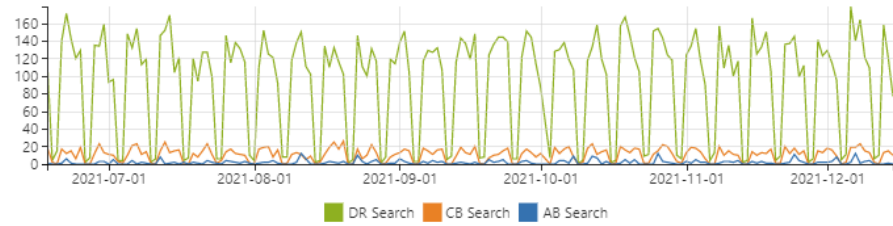
- Working area Data Quality WMDA-NMPD
- Summary (Public)
- Meeting notes
- File lists

PAGE TREE

- Summary
  - Data share with Orgs
  - Data Upload Dashboard
  - Data upload Analysis
  - Registry Data Quality report
  - WMDA Searches Overview**
  - Data Upload Dashboard - Density
- Working area Data Quality WMDA-NMPD

### The Daily Search from 2021

Date =



### The Monthly Search value/percentage of 2020 and 2021

