

Data Management Sessions: EBMT 2014 REVIEW



In Milan earlier this month we enjoyed meeting 150 Data Managers from over 30 countries in Europe and also outside including Australia, Brazil, Canada & China

Thank you to everyone who attended. We have had some excellent feedback so far and a full report will be published on www.ebmt.org – [Data Management] – Helpdesk. Your comments are very important to us in determining future educational programmes on Data Management. Pamela Welson has been coordinating a great educational programme and will be all set for the next one. Suggestions for future topics can be emailed to registryhelpdesk@ebmt.org

The Educational sessions included clinical background and data management specific talks. For the clinical talks on Aplastic Anaemia; Cell Therapy; Cytogenetics & Markers; Graft manipulation; Graft Loss; GvHD definitions; Haplo transplants; Multiple Myeloma and Viral Serology we would like to express our gratitude to **Jakob Passweg; Willem Fibbe; Richard Schlenk; Alvaro Urbano; Attilio Bondanza; Hildegard Greinix, Daphna Hutt; Curly Morris & Jan Styczynski.**

Data Management specific sessions were presented from the Central Registry: Common data reporting problems; Role of data management in EBMT. We are extremely grateful to **Henk Jan van der Wijk** for his talk on Data Confidentiality and to **Marja Pekkanen** of Helsinki University Hospital who presented methods of collecting long term follow up data on behalf of a centre. And thank you also to **Tuula Rintala** at Kings College Hospital, for her introduction to HSCT & MED-A which combined both data management and clinical aspects

Slides from these educational sessions are being made available on www.ebmt.org – [Data Management] – Helpdesk once we have permission from the speakers.

Previews: New Registry System

During the congress, Remedy Informatics, provider of the new EBMT Registry System, gave a series of demos on Tuesday 1st April to 120 participants. This is where many of us were able to find out more and see for the first time how the new system will look before it goes live later in the year.

The EBMT Registry Database (ProMISe) training courses were also well attended. Prior to the data migration, it is important for users to run a series of reports to check for errors that may prevent records from being transferred. These Data Quality reports were demonstrated by **Carmen** from the Registry Office in London. Click [here](#) to view the slides from her **Tidy Up Your Data** session.

Please see overleaf for further important info on Data cleaning prior to migration

We would like to thank **Emmanuelle Polge & Babatunde Oyenuga** for presenting the ProMISe Data Retrieval session - and all the computer room helpers: **Leila & Nicole** from the French registry, **Beate** from the Austrian Registry and **Mor Seny** from the Study Office in Paris.

Finally - **SAVE THE DATE** - We really hope to see as many of our Data Managers again next year at [EBMT 2015](#) in Istanbul (22nd to 25th March)

A Very Special Mention: Ronald & Henk Jan



Professor Ronald Brand (Leiden University Medical Centre) was presented the Outstanding Achievement Award at the opening ceremony. He has been involved in EBMT projects for over 30 years as a statistician and designer of Project Manager and then ProMISe, the centralised system in use for more than a decade. Ronald in turn thanked all of his friends at EBMT.



Henk Jan van der Wijk (also from LUMC) was presented with a thank you gift at the Data Managers reception. He has spent many years providing ProMISe technical support, setting up the training environment and maintaining a range of Excel macros to help data managers interpret the data for analysis. And all this in the most calm, friendly and efficient manner!

Data Cleaning Prior to Migration to New System

Some of the current data needs to be cleaned in order for the records to be successfully transferred to the new system. At the moment we are cleaning **Patients that have 2 different UPN (hospital numbers) within the same centre.**

For these cases we will use the following procedure:

1. the UPN in the patient table (field UPN) will be considered the best. The reasoning is that if users detect a UPN problem, they are likely to correct it here without realising that there is another UPN lurking in the treatment table.
2. the UPN in the patient table will be copied to the UPN in the treatment table (field UPN2) so that there is only one UPN per patient
3. the UPN that exists in the treatment table will be copied to the "Dossier number" (field VDOSSIER). This will surprise users that do not use this field. These users can ask us to delete all the numbers here after they have checked that the UPN in the patient table is indeed the one they want to retain. (The dossier number is only an optional field and we can always hide it from users later on).
4. The differences between the UPN in the patient and treatment table may be minimal; a typical example is that the number of the transplant has been used as an add-on, so one UPN says 20874.1 and the other one says 20874.2. It is impossible to check this one by one, so sometimes the entry in the "Dossier number" will be practically identical to the entry under UPN.

More Data Checks before Migration

In relation to the Registry Upgrade mentioned above, we would like to remind centres that this is the perfect opportunity to look again at their data and to check the Data Quality before it is transferred into the new system. Some records may not be transferable if the data is not correctly recorded.

Data Quality reports

Many national registries are in regular contact with centres regarding missing items, data discrepancies and follow up updates. However not all centres have this contact, so this appeal is to those centres in particular.

For users who currently have access to ProMISe, a good place to start is to run the Data Quality Reports which are available to all centres. (Even if you do not use ProMISe to submit data, you can still apply for access in order to check your centre's data registration status and quality)

These reports can help you to identify some key issues such as invalid data, missing data, duplicate information, incorrect diagnosis, error messages and conflicting information. A guide detailing how to run these reports is available in the Data Management section of our web site: [Guide to producing Data Quality Reports in the EBMT Registry Database Using ProMISe](#)

The reports are straightforward to run – they make it easy to see where data is missing or wrongly entered and then you will have the opportunity to correct the errors.

You can also have a look at Carmen's [Tidy Up Your Data](#) presentation mentioned on the previous page.

We are happy to help you so if you need any help please contact us on registryhelpdesk@ebmt.org

Data Collection Forms

Please click [here](#) to download the current data collection forms.

Forms are reviewed on a regular basis and the content is decided by the EBMT Working Parties. A data collection review process is currently underway. A special Registry Working Group was set up, led by Per Ljungman, with representatives from the Scientific Council, Statisticians, Definitions Group and EBMT Study/Data Offices. The aims of this process are to review and improve the EBMT data collection structure and procedures; use the forthcoming Remedy system as efficiently as possible, and improve data quality.

Suggestions for changes to the forms – the clinical content - can be referred to the [Scientific Council](#)

It is still under review but it is likely that the MED-A will be expanded to include items on relapse between HSCT; chromosome analysis and molecular markers at diagnosis and pre-transplant treatment for some diseases.

Error prevention

Follow up events: Please note that while carrying out our data checks we have found some cases where the date of chimaerism or the date of a complication or gvhd is after the date last seen. This is not possible. If you are adding these dates in ProMISe, please ensure that there is always a follow up assessment on that date showing the last patient status

MED-A general: The documentation from the congress session on **Common Data Reporting Problems** is available online [here](#). You can view some MED-A forms with the most common discrepancies highlighted.

Data Confidentiality

Please remember to password protect / anonymise patient data sent by email. Henk Jan van der Wijk's presentation on protecting privacy in medical research is available [here](#)