Data Quality using MicroStrategy Reports

Excel version

V1.0 | August, 2025

Introduction

To carry out the activities described in this document, you will need to have access to the EBMT Registry. Access to the EBMT Registry is only available for EBMT member centers. For more information, click here.

This PDF document is created to provide users of the EBMT Registry with guidance on how to perform data quality checks using the downloaded reports from the MicroStrategy library in Excel. Prior knowledge on how to access MicroStrategy, navigate the library and downloading tables is required to follow the next pages.

If at any point you want to reset the filters you have applied in the MicroStrategy reports, use the ‘Reset’ button at the top of the report:

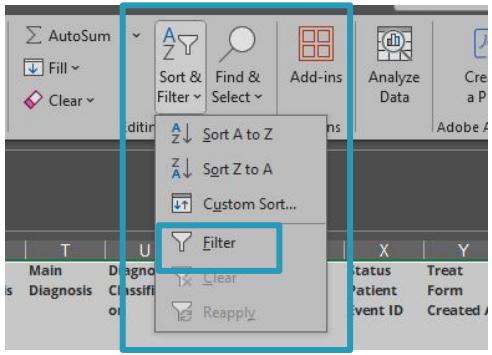
For more information about MicroStrategy and the EBMT Registry, click here.

Excel tools used

- Sorting
- Filtering

To add easy sorting and filtering in your Excel document:

- select all data in the sheet by pressing Ctrl+A on the keyboard, then select ‘Filter’ from ‘Sort & Filter’ in the Excel menu;

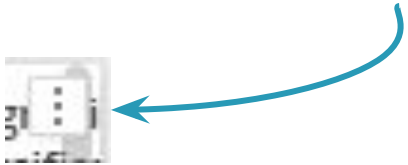


- or press ‘Ctrl + Shift + L’ on the keyboard.

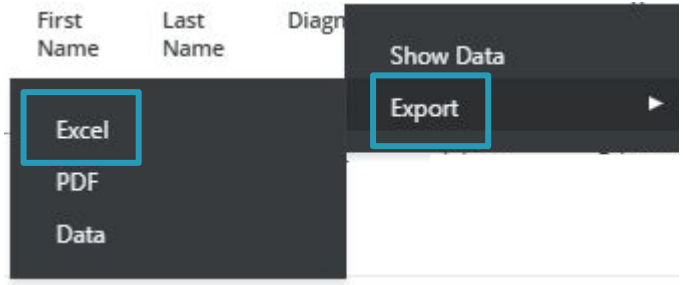
Once the filter tool is added, an 🔽 icon will appear to cells in row 1.

Downloading tables from MicroStrategy

After opening a report on the data table chapter in MicroStrategy, use the options button in the right upper corner of the data table:



Select ‘Export’ and ‘Excel’ in the appeared menu to download the report in excel format:



EBMT Registry

Illogical dates

The first data quality check that can be performed using the treatment overview is checking for **illogical dates**. By weird dates we mean dates in the future, or very far in the past.

For example, diagnosis dates in the year 107, or HCTs in 3031.

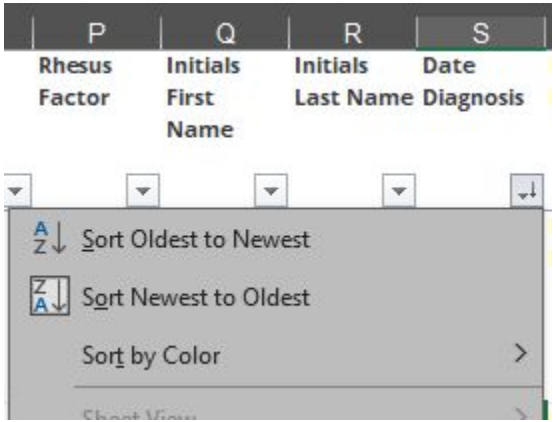
Steps to follow

Download the treatment overview from the MicroStrategy library.

EBMT Registry

Illogical dates

To check the dates that may have illogical values, select any date column, and first ‘Sort Oldest to Newest’. This way you can see the oldest dates, which might give insight in typos or entry mistakes.

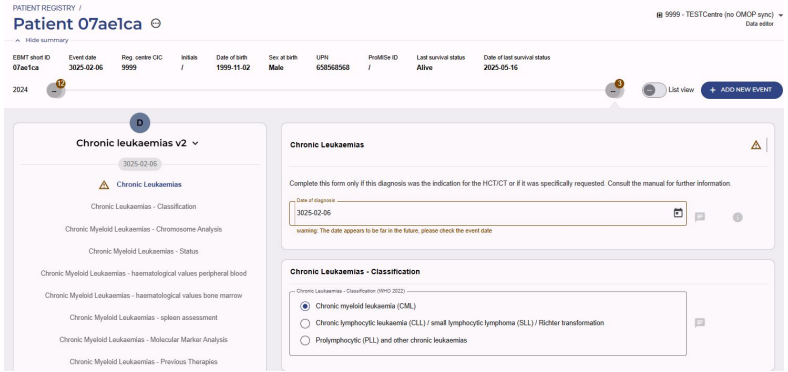


After, you can sort ‘Sort Newest to Oldest’, to see if there are any dates that appear to be in the future.

Illogical dates - example

Initials First Name	Initials Last Name	Date Diagnosis	Main Diagnosis
		06-02-3025	Chronic leukaemia

The unrealistic date in the report: a diagnosis in the year 3025. Possibly the center made a typo, meaning to write 2025 instead



The patient in the EBMT Registry, with a warning on the diagnosis date

Alive patients with death date

In the EBMT Registry, it is currently not possible for the system to automatically check data between different events.

In this exercise we will check for patients whose last survival status is alive, while they have been reported dead on a previous assessment.

Steps to follow

Download the treatment overview from the MicroStrategy library.

Alive patient with death date

Date Last Follow Up No Censor	Status Last Follow Up No Censor	Death Date	Death Cause
04-03-2025	Alive	15-02-2024	
03-06-2021	Alive	03-06-2021	Other
04-04-2024	Alive	03-09-2023	Unknown
04-04-2024	Alive	03-09-2023	Unknown
25-08-2022	Alive	13-06-2022	Relapse/progression

On the left you see a screenshot of patients who are marked as alive, yet the ‘Death Date’ column is filled in. To better understand this situation, let’s take a look at the columns:

- Date last follow up no censor: the date of the last follow-up for this patient in the EBMT Registry, without censoring subsequent follow-ups;
- Status last follow up no censor: last known survival status for the patient (from follow-up or status event), without censoring subsequent follow-ups;
- Death date: date of assessment where patient was reported dead;
- Death cause: reported cause of death.

How can alive patients have a death date?

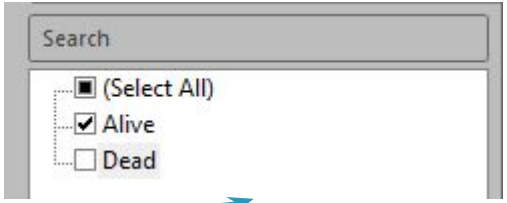
In the EBMT Registry, it is possible to mark a patient as dead and create an event on a later date where the patient is alive.

Alive patient with death date

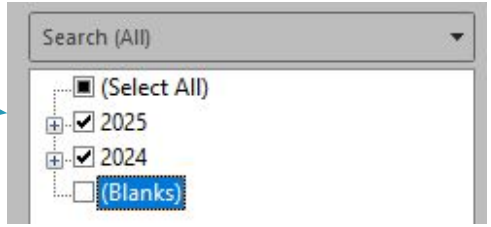
How to find these patients:

1. Navigate to the relevant columns (“status last follow up no censor” and “death date”)
2. Use the filter button in the column to filter down:

   ➢ In the filter dropdown of the column “status last follow up no censor”, filter on ‘Alive’



   ➢ In the ‘Death Date’ column, unselect the empty cells



3. You have now made a filter for patients who are alive and the death date is not empty

If you cannot carry out this step, congratulations! It means there are no alive patients with a death date in your context.

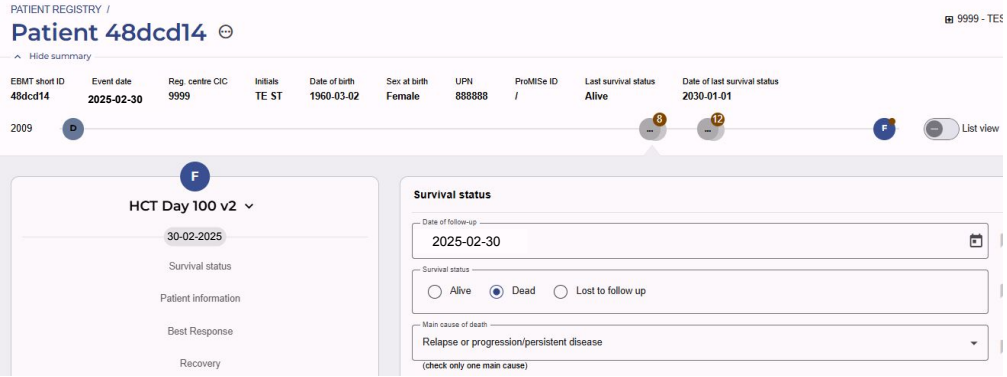
If you do see results, check the patient’s long or short ID to open them in the EBMT Registry. The event with a date identical to the ‘Death Date’ is the one that has an issue.

EBMT Registry

Alive patient with death date - example

Status Last Follow Up No Censor	Death Date	Death Cause
Alive	30-2-2025	Relapse/progression

The patient's last survival status is alive, yet the patient is reported to have died in 2025.



The patient in the EBMT Registry, where we see the patient being reported as dead in 2025. However, as seen in the 'Last survival status' above the timeline, the patient is reported to be alive in 2030. We also see more events after the follow-up where the patient is reported to have passed away.

Patient’s age at treatment

The patient’s age at treatment is automatically calculated using the patient date of birth and the date of the treatment. The age at treatment can be a good indication for a data entry mistake, either in the patient details or in the treatment date.

Steps to follow

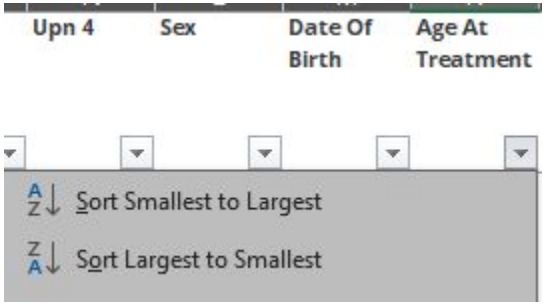
Download the treatment overview from the MicroStrategy library.

EBMT Registry

Patient’s age at treatment

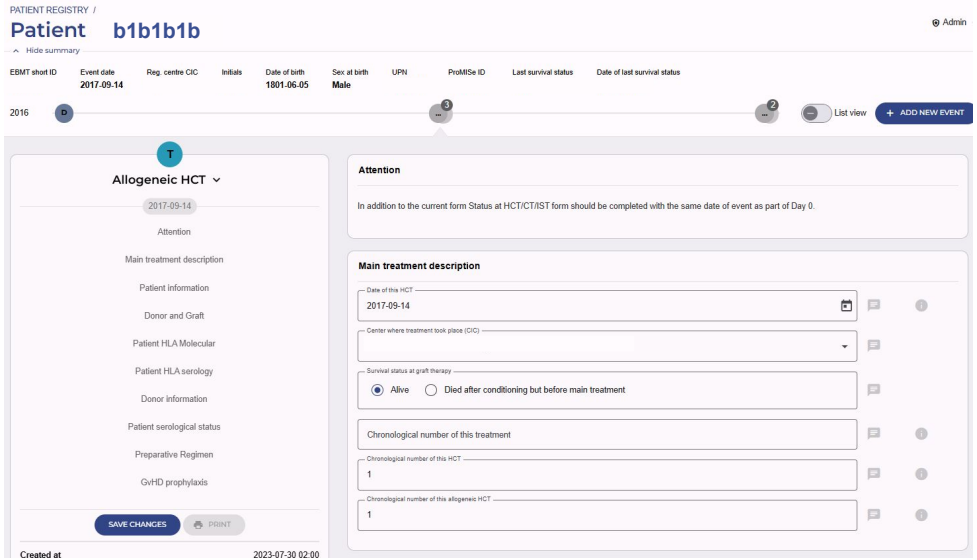
In the treatment overview there is a column named ‘Age at Treatment’. This column can be sorted from largest to smallest or vice versa by clicking on the filter 🔽 button

If there is an odd age (e.g. negative number, or >100), in the column on the left of the age at treatment you can see the date of birth, and several columns to the right the date of treatment. This can help you identifying the cause of the unexpected age.



Patient’s age at treatment - example

The odd age in the treatment overview: the patient was 216 years old at HCT



The patient in the EBMT Registry, with a treatment in 2017 and a date of birth in 1801. It is unlikely this is the real year of birth of the patient. Ideally, the real year should be entered.