Top Five Data Quality Issues

Andréa Benoit
Tuesday, February 20th, 2018

There are no conflicts of interest to disclose.
More than a Completed Form…

• One completed form is only part of the patient’s story…

GVHD previously reported

Recipient is in CR
Each Patient Has a Story
Issues to be Covered

• “Previously reported” issue
• Reporting the date of death
• Status at transplant
• Donor identification across forms
• Patient Basics
Previously Reported

• What is ‘Previously Reported’?
  – Option on commonly asked questions to eliminate duplicate reporting
  – Most often reported for initial dates
    • Initial ANC & platelet recovery date
    • Initial GVHD onset
    • Original date of first CR

• What is the issue?
  – Using the option of ‘previously reported’ when the question *was not* previously answered
Previously Reported Scenario

- Why might issues happen?
  - Ex: +100 d F2450: patient did not have initial ANC recovery

<table>
<thead>
<tr>
<th>Initial ANC Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>14. Was there evidence of initial hematopoietic recovery?</td>
</tr>
<tr>
<td>☐ Yes (ANC ≥ 500/mm³ achieved and sustained for 3 lab values) - Go to question 15</td>
</tr>
<tr>
<td>☐ No (ANC ≥ 500/mm³ was not achieved) - Go to question 16</td>
</tr>
<tr>
<td>☐ Not applicable (ANC never dropped below 500/mm³ at any time after the start of the preparative regimen)</td>
</tr>
<tr>
<td>☐ Previously reported (Recipient’s initial hematopoietic recovery was recorded on a previous report) - Go to</td>
</tr>
</tbody>
</table>

| 15. Date ANC ≥ 500/mm³ (first of 3 lab values): [YYYY] |
Previously Reported Scenario

+6 month (and on) F2450: patient must have recovered during 100 day timeframe

<table>
<thead>
<tr>
<th>Initial ANC Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>14. Was there evidence of initial hematopoietic recovery?</td>
</tr>
<tr>
<td>□ Yes (ANC ≥ 500/mm³ achieved and sustained for 3 lab values) - Go to question 15</td>
</tr>
<tr>
<td>□ No (ANC ≥ 500/mm³ was not achieved) - Go to question 16</td>
</tr>
<tr>
<td>□ Not applicable (ANC never dropped below 500/mm³ at any time after the start of the preparative regimen)</td>
</tr>
<tr>
<td>● Previously reported (Recipient’s initial hematopoietic recovery was recorded on a previous report) - Go to</td>
</tr>
</tbody>
</table>

15. Date ANC ≥ 500/mm³ (first of 3 lab values): ___ ___ ___ ___
Previously Reported Scenario

- Why should we care?
  - Missing data!

<table>
<thead>
<tr>
<th>CRID</th>
<th>Transplant Date</th>
<th>Report</th>
<th>ANC Recovery?</th>
<th>ANC Date?</th>
</tr>
</thead>
<tbody>
<tr>
<td>56789</td>
<td>8/15/2015</td>
<td>F2450 – 100 d</td>
<td>No</td>
<td>n/a</td>
</tr>
<tr>
<td>56789</td>
<td>8/15/2015</td>
<td>F2450 – 6 mo.</td>
<td>Previously reported*</td>
<td>n/a</td>
</tr>
<tr>
<td>56789</td>
<td>8/15/2015</td>
<td>F2450 – 1 yr.</td>
<td>Previously reported*</td>
<td>n/a</td>
</tr>
<tr>
<td>56789</td>
<td>8/15/2015</td>
<td>F2450 – 2 yr.</td>
<td>Previously reported*</td>
<td>n/a</td>
</tr>
</tbody>
</table>

*Remember, it is important to note that the answer of 'previously reported' is synonymous with YES, the patient recovered
Previously Reported

• How do we fix it?
  – Usually just one form within the appropriate time frame has to be updated

• How do we prevent this from happening?
  – Make a note to yourself in your shadow chart
  – Open the prior form(s)
  – FormsNet validations prevent ‘previously reported’ from being answered on the 100d form
Previously Reported Scenario

• How do we fix it?

<table>
<thead>
<tr>
<th>CRID</th>
<th>Transplant Date</th>
<th>Report</th>
<th>ANC Recovery?</th>
<th>ANC Date?</th>
</tr>
</thead>
<tbody>
<tr>
<td>56789</td>
<td>8/15/2015</td>
<td>2450 – 100 d</td>
<td>No</td>
<td>n/a</td>
</tr>
<tr>
<td>56789</td>
<td>8/15/2015</td>
<td>2450 – 6 mo.</td>
<td>Yes</td>
<td>1/15/2016</td>
</tr>
<tr>
<td>56789</td>
<td>8/15/2015</td>
<td>2450 – 1 yr.</td>
<td>Previously reported</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Reporting Date of Death

• What and where to report date of death?
  – Recipient death is reported on both the TED (F2450) and CRF (F2100) track
    • CRF recipients also require a F2900

• What are the issues?
  a) ‘Last contact date’ (LCD) reported on F2100 and ‘date of death’ (DOD) on F2900 not matching
  b) Overriding the date of death (F2450)
Date of Death Scenario

• Why might it happen?
  – “Date of *actual* contact” can be misleading when the recipient dies
  – Exact date of death unknown

• Why should we care?
  – Different answers lead to confusion on when the patient actually died (i.e. TCSA)
  – The system will remove future due forms when they match
  – Prevents additional, unnecessary form completion
## Date of Death Scenario (CRF)

<table>
<thead>
<tr>
<th>Status</th>
<th>Center</th>
<th>Event</th>
<th>Form</th>
<th>Visit</th>
<th>Group</th>
<th>Sequence</th>
<th>Earliest</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2400</td>
<td>Pre-TED</td>
<td>5864230</td>
<td>2015-05-30</td>
<td></td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2402</td>
<td>Pre-TED</td>
<td>45687423</td>
<td>2015-06-13</td>
<td></td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2400</td>
<td>Baseline</td>
<td>7895213</td>
<td>2015-06-13</td>
<td></td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2400</td>
<td>Baseline</td>
<td>5864137</td>
<td>2015-06-13</td>
<td></td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2400</td>
<td>Indication</td>
<td>44458631</td>
<td>2015-09-21</td>
<td></td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2400</td>
<td>Preliminary</td>
<td>6321019</td>
<td>2015-09-21</td>
<td></td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2400</td>
<td>Preliminary</td>
<td>9041337</td>
<td>2015-09-21</td>
<td></td>
</tr>
<tr>
<td>DUE</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2100</td>
<td>Follow-up</td>
<td>2015-09-21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUE</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2100</td>
<td>Follow-up</td>
<td>2015-12-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUE</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2100</td>
<td>Follow-up</td>
<td>2016-06-13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUE</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2100</td>
<td>Follow-up</td>
<td>2016-06-13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUE</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2100</td>
<td>Follow-up</td>
<td>2017-06-13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUE</td>
<td>99999</td>
<td>2015-06-13</td>
<td>2100</td>
<td>Follow-up</td>
<td>2017-06-13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: CMP stands for Control, CMP stands for Treatment Group*
Date of Death Scenario (CRF)

F2100

Vital Status

Information should come from an actual examination by the Transplant Center provider or the local provider who is following post-HCT.

1. Date of actual contact with the recipient to determine medical status for this follow-up report: \( \frac{2014}{YYYY} / \frac{08}{MM} / \frac{01}{DD} \)

2. Specify the recipient’s survival status at the date of last contact

☐ Alive – Answers to subsequent questions should reflect clinical status since the date of last report.

☒ Dead – Answers to subsequent questions should reflect clinical status between the date of last report and immediate death. Complete a Form 2900 – Recipient Death Data.

F2900

Recipient Death

1. Date of death: \( \frac{2014}{YYYY} / \frac{08}{MM} / \frac{15}{DD} \) ☐ Date estimated

CIBMTR

CENTER FOR INTERNATIONAL BLOOD & MARROW TRANSPLANT RESEARCH
Date of Death Scenario (CRF)

- How do we fix it?

<table>
<thead>
<tr>
<th>Query Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question #1 - Date of actual contact with the recipient to determine medical status for this follow-up report: (Insert Patient's Full Name)</td>
</tr>
<tr>
<td>Comment:</td>
</tr>
<tr>
<td>4000 Character</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comment History</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>2018-01-15 10:28:51 AM</td>
</tr>
</tbody>
</table>
## Date of Death Scenario (CRF)

### F2100

**Vital Status**

Information should come from an actual examination by the Transplant Center provider or the local provider who is following post-HCT.

1. Date of actual contact with the recipient to determine medical status for this follow-up report: \( \frac{2014}{YYYY} \) / \( \frac{08}{MM} \) / \( \frac{15}{DD} \)

2. Specify the recipient’s survival status at the date of last contact
   - Alive – Answers to subsequent questions should reflect clinical status since the date of last report.
   - Dead – Answers to subsequent questions should reflect clinical status between the date of last report and immediate death. Complete a Form 2900 – Recipient Death Data.

### F2900

**Recipient Death**

1. Date of death: \( \frac{2014}{YYYY} \) / \( \frac{08}{MM} \) / \( \frac{15}{DD} \) 

   - \( \square \) Date estimated
Date of Death Scenario (CRF)

<table>
<thead>
<tr>
<th>Status</th>
<th>Center</th>
<th>Event</th>
<th>Sequence</th>
<th>Earliest</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>5864230</td>
<td>2015-05-30</td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>45687423</td>
<td>2015-06-13</td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>7895213</td>
<td>2015-06-13</td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>5864137</td>
<td>2015-06-13</td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>44456931</td>
<td>2015-09-21</td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>6321019</td>
<td>2015-09-21</td>
</tr>
<tr>
<td>CMP</td>
<td>99999</td>
<td>2015-06-13</td>
<td>0641337</td>
<td>2015-09-21</td>
</tr>
</tbody>
</table>
Date of Death Scenario (TED)

- Similar scenario, except there are no two forms that need to match
- F2450 should have a date of death, even if it is estimated
  - Estimating Dates
- Future forms will be removed!
Date of Death Scenario (TED)
Disease Status

• What is the disease status?
  – Assessment of disease prior to and after transplant
  – How did the patient respond to transplant?

• Where is disease status reported?
  – F2402 for all patients with malignant disease
  – F2450 (TED)
  – F20XX and F21XX (CRF)
Disease Status

• What are the issues?
  a) Reported disease status conflicts with actual lab values/hematological results
  b) Difference between the disease status at transplant reported on disease specific baseline and F2402 (and F2400 in prior years)
Disease Status Scenario (a)

• Do the hematological results/disease parameters I am reporting at the last assessment match up with the disease status I am reporting?
  – Use the CIBMTR response criteria*

• Conflicting results
  – Ex: recipient with AML in ‘CR’ with 9% blasts
  – Ex: recipient with MM in ‘CR’ with IFEs ‘not done’ or ‘unknown’

*We do acknowledge that there are rare exceptions to the criteria based on the patient’s whole clinical workup and physician’s judgement.
Disease Status Scenario (a)

127. WBC:
- [ ] known
- [ ] Unknown

128. ___ ___ ___ ___ ___ 6 6

129. Date sample collected: ___ ___ ___ YYY

130. Blasts in blood:
- [ ] known
- [ ] Unknown

131. ___ ___ 3 %

132. Date sample collected: ___ ___ ___ YYY

193. What was the disease status (based on hematoxylin and eosin-stained bone marrow aspirate)?
- [ ] Primary induction failure
- [ ] 1st complete remission
- [ ] 2nd complete remission
- [ ] ≥3rd complete remission
- [ ] 1st relapse
- [ ] 2nd relapse
- [ ] ≥3rd relapse
- [ ] No treatment
Disease Status Scenario (b)

Status at transplantation:

57. What was the disease status (based on hematological test results)?
   - Primary induction failure - Go to question 63
   - 1st complete remission (no previous bone marrow or extramedullary relapse) - Go to question 58
   - 2nd complete remission - Go to question 58
   - ≥ 3rd complete remission - Go to question 58
   - 1st relapse - Go to question 62
   - 2nd relapse - Go to question 62
   - ≥ 3rd relapse - Go to question 62
   - No treatment - Go to question 63

This is the same question for the same time-point, and they should match

F2402

F2010

Disease Status at the Last Evaluation Prior to the Start of the Preparative Regimen (Conditioning)

193. What was the disease status (based on hematological test results)?
   - Primary induction failure
   - 1st complete remission
   - 2nd complete remission
   - ≥ 3rd complete remission
   - 1st relapse
   - 2nd relapse
   - ≥ 3rd relapse
   - No treatment

Specify which of the following show start of the preparative regimen:
194. Blood
195. Bone marrow
196. Central nervous system
197. Skin
198. Testes
199. Other site
Disease Status Scenario

• How might discrepancies occur?
  – Disease inserts collect hematological results/disease parameters
  – F2400/2402 often only collects disease status and does not collect hematological results
• Easier to consider the disease status when reviewing all the assessments

• Bonus Fun Fact: Mass scale clean-up for mismatches are on their way 😊
Disease Status

• Why do we care?
  – Status at transplant is an important study criteria
    • CVDR, TCSA
  – Accurate transplant status at baseline will help with future form completion
    – Auditors love this field!

• How do we fix (prevent) it?
  – Disease tracker
  – Check what was reported on the F2400/2402
  – Form revision to remove duplicate questions
Donor Identification

• Where is this data reported?
  – F2400: Donor Information

• What is the issue?
  – Multiple IDs for each donor, product and recipient (external and internal) is confusing!
  – Once an ID is captured in FN3, this ID must be consistently reported on all subsequent forms
  – Only way to link all relevant data
Donor Identification Scenario

F2400 – first form that is filled out – make this the ‘source of truth’

40. Specify the related donor type:
   □ Syngeneic (monozygotic twin)
   □ HLA-identical sibling (may include non-monozygotic twin)
   □ HLA-matched other relative
   □ HLA-mismatched relative

41. Date of birth: (donor / infant)
   □ Known
   □ Unknown

42. Date of birth: (donor / infant): __2010___ / __08___ / __15___

43. Age: (donor/infant)
   □ Known
   □ Unknown

   □ Months (use only if less than 1 year old)
   □ Years
Donor Identification Scenario

Donor 1

**DOB= 8/15/2010**

F2004
Donor DOB = 8/15/2010

F2005
Donor DOB = 9/2/1966

F2006
Donor DOB = 8/15/2010

F2400
Donor DOB = 8/15/2010
Donor Identification Scenario

F2400 – first form that is filled out – make this the ‘source of truth’

31. Specify donor:
   - □ Autologous - Go to question 46
   - □ Autologous cord blood unit - Go to question 35
   - □ NMDP unrelated cord blood unit - Go to question 32
   - □ NMDP unrelated donor - Go to question 33
   - □ Related donor - Go to question 40
   - □ Related cord blood unit - Go to question 35
   - □ Non-NMDP unrelated donor - Go to question 34
   - □ Non-NMDP unrelated cord blood unit - Go to question 35

34. Non-NMDP unrelated donor ID: (not applicable for related donors)

   IGGY1045

   _____________________________ _____________________________ _____________________________ – Go to question 38
Donor Identification Scenario

Donor 1
ID = IGGY1045

F2004
Donor ID = IGGY1045

F2005
Donor ID = IGGY1045

F2400
Donor ID = IGGY1045

F2006
Donor ID = IGGY1045
Donor Identification Scenario

- Why might it happen?
  - Looking at recipient DOB instead of donor DOB
  - Zero vs O - I vs 1 vs i - etc.
  - We know they are the same donor, but at the database level, it has to look at what is actually reported

- Why do we care?
  - Missing, unlinking data!
  - The donor has a story too
Patient Basics

- Demographics
- Primary disease for transplant
- Donor/Products
Patient Basics - Demographics

• What demographics are collected?
  – Age, race
  – Ethnicity, zip code (US patients)

• Where is the data collected?
  – F2804, F2400, F2000
  – Previous versions of F2450

• Why do we care?
  – ALWAYS reviewed: data used in studies, HRSA reports

• What are the issues?
  – Discrepancies between forms, data reported
Patient Basics - Demographics

**F2400**

Recipient Data

1. Date of birth: ___/___/___
   
2. Sex: □ Male □ Female

3. Ethnicity: □ Hispanic or Latino □ Not Hispanic or Latino □ Not applicable (not a resident of the USA) □ Unknown

4. Race: □ White □ Black or African American □ Asian □ American Indian or Alaska Native □ Native Hawaiian or Other Pacific Islander □ Not reported □ Unknown

**F2000**

Race:

- □ White
- □ Black or African American
- □ Asian
- □ American Indian or Alaska Native
- □ Native Hawaiian or Other Pacific Islander

5. Race detail
   - □ Eastern European
   - □ Mediterranean
   - □ Middle Eastern
   - □ North Coast of Africa
   - □ North American
Patient Basics - Demographics

• Why does it happen?
  – Asked slightly differently on forms
  – Asked for each F2400/F2000 completed

• How do we fix (prevent) it?
  – Check twice, enter once
  – Review previously submitted forms
  – Ask your CRC for guidance
Patient Basics – Primary Disease

- Where is the primary disease collected?
  - F2402 (current)
  - Previously F2400, F2000, F20XX

- Why do we care?
  - Arguably the most important data point!

- What are the issues?
  - FormsNet uses primary disease information to generate disease specific forms (CRF)
    - Completing forms that are not required, re-doing work
  - FormsNet uses primary disease information to enable/disable questions
    - Missing or unnecessary data
Patient Basics – Primary Disease

• Ex: Enabling/disabling questions on F2450 depend on whether or not the disease is malignant or non-malignant

Malignant Diseases Only
Only complete questions 75-97 if the HCT being reported was given to treat a malignant disease. If the HCT being reported was given to treat a non-malignant disease, leave questions 75-97 blank. FormsNet should enable / disable this section based on the primary disease reported on the Pre-TED Disease Classification Form (Form 2402). Contact your CRC if you believe FormsNet is incorrectly enabling / disabling these fields.

• Updating the primary disease may require you to review the data reported on the F2450
Patient Basics – Donor/Products

- Where are the donor and products received collected?
  - F2400

- Why do we care?
  - Forms (e.g. F2005, F2006) come due based on how these questions are answered
Patient Basics – Donor/Products

• What are the issues?
  – Misreporting of donor type/products may lead to the wrong forms coming/not coming due
  – Ex: IDM (F2004) and HLA (F2005) information not required for NMDP products
    • Data already available via NMDP

• Why does it happen?
  – Multiple IDs for each donor, product and recipient (external and internal) is confusing!
  – NMDP has started facilitating related transplants
You’re not in this alone!

• Form validations are reviewed and added periodically
  – Working on introducing more cross form validations (e.g. disease status at transplant)

• Form revision
  – Removing duplicate questions on different forms

• Data quality checks that look at more recent data
Additional Resources

- CIBMTR Form Instruction Manual
- Disease Tracker Example
- General Guidelines for Completing Forms
- How Forms Come Due
- Your assigned CRC
- anbenoit@mcw.edu