

Massachusetts Data Analysis Center – Mass-DAC

Cardiac Surgery Data Quality Report Summary

[STS Version 2.73]

March 9, 2012

Introduction

The Mass-DAC version 2.73 Cardiac Surgery Data Quality Report (DQR) contains structural changes due to changes in the data submission process and the new data elements required by the STS version 2.73 data specifications. The Data Quality Report consists of two files, a PDF formatted report and an Excel workbook that contains the details of issues found in the submission data using the Mass-DAC and STS data cleaning edits. The PDF report contains five different sections, beginning after the title page and the section introduction, that summarize the information contained within the report.

PDF Report Sections

1. **Data Submission Verification Summary:** Summarizes the total number of submitted records by Procedure Type, total operative mortalities and the total number of 30-day mortalities. The Hospital column should reflect the expected number of records in each submission for each category, where as the Mass-DAC column shows the total records processed by Mass-DAC. Any discrepancies between what the hospital reported and what Mass-DAC found should be corrected or explained in the next data submission or via e-mail to the Mass-DAC data manager.
2. **Submitted File Record Summary:** Provides details on the number of records found in each text delimited data file. If a separate file is used for Mass-DAC data elements, the total number of records in each file should agree. If only a single file is utilized, then the Mass-DAC value will be “Not Applicable”.

Surgeon volumes by Procedure Type should be checked with every submission. Please use an independent source, something other than the database used to create the harvest extract, to verify that the volumes are correct.

3. **Data Quality Issue Summary:** Summarizes the issues found using two tables. The first table provides the total number of each issue type found, and if a second or later submission splits the issues by New/Old based on whether or not it was in the prior submission data quality report. The second table provides additional details on each issue, breaking out the issues by variable. The columns are:
 - a. **Source** – Either STS defined edit or a Mass-DAC defined edit.
 - b. **Variable Name** – The primary variable name assigned by Mass-DAC to the issue. If there is more than one variable involved, it may be noted in the Issue Explanation and can be found in the Detailed Issue Listing and the Excel workbook.

- c. **Label (STS Seq#)** – A short descriptive label assigned to the Variable Name, and if it is an STS variable, the STS sequence number is in parentheses.
 - d. **Issue Explanation** – Provides additional information about the edit that flagged the issue.
 - e. **N** – Indicates the total number of records that had the same issue.
4. **Detailed Issue Listing:** Provides the details at the procedure level for each issue.
 - a. **Status** – New means this patient’s procedure issue is new to this DQR and was not on the prior DQR;
Old = this was listed on the prior DQR.
 - b. **PatId** – The STS patient id that is generated by the vendor software.
 - c. **MedRecN** – The patient medical record number submitted to STS.
 - d. **Admit Date** – The STS Admission Date value in YYYY-MM-DD format.
 - e. **Surgery Date** – The STS Surgery Date value in YYYY-MM-DD format.
 - f. **Issue Type** – Mass-DAC defined issue type,
 - 1=Missing,
 - 2=Out of Range,
 - 3=Out of Usual Range,
 - 4=Parent/Child Inconsistency,
 - 5=Date Inconsistency,
 - 6=Other multi-variable relationship check.
 - g. **Variable Name** – The Mass-DAC defined variable name for the primary variable used in the edit code. The names are usually the same as the STS short names, with a few exceptions.
 - h. **All Values** – If only one variable is used in the edit code, than this contains information pertaining only to the variable listed in the Variable Name column. If more than one variable is used in the edit code, then the first line corresponds to the primary variable, and the ones below are for all other associated variables. Each line has the structure where the text explanation of the value along with the STS sequence number is to the left of the = and the text equivalent of the stored value is to the right of the =.
5. **Lists for Patient Mortality, Non-Cardiac Procedures, and Frequency Distributions:** The first several pages of this section prints out the key fields for patients identified as a mortality and the list of patients that were identified as a non-cardiac case using the STS procedure identification logic. All variables used to determine the procedure type are included in the Non-Cardiac Procedure list. The remaining pages list the frequencies for most variables submitted, including Mass-DAC computed variables or useful cross-tabulations. NA in the variable value columns denotes not applicable. (For example, cause of death is NA for those discharged alive.) This part of the report is divided into several sections that correspond to the STS section letters on the data collection form and specifications.

Excel Workbook Columns

The Excel file provides the same records as the Detailed Issue Listing in the PDF report, but provides additional columns to help identify records flagged as an issue.

1. **Issue Type:** Mass-DAC defined issue type;
 - 1=Missing,
 - 2=Out of Range,
 - 3=Out of Usual Range,
 - 4=Parent/Child Inconsistency,
 - 5=Date Inconsistency,
 - 6=Other multi-variable relationship check.

2. **Issue Status-New or Old (in prior):**
 - New means this patient's procedure issue is new to this DQR and was not on the prior DQR;
 - Old = this was listed on the prior DQR.

3. **Total # of Associated Variables:** If 0, then only the primary variable was involved in the edit, otherwise there are one or more associated variables that may have caused the issue to be flagged and could be the possible place a correction is needed.

4. **Site-Hospital Abbr:** This is the short 10 character hospital site abbreviation used by Mass-DAC to identify each hospital.

5. **Record ID (#50):** This is the STS Record id, sequence #50, that is generated by the vendor software.

6. **Patient ID (#80):** This is the STS patient id, sequence #80, that is generated by the vendor software.

7. **Medical Record Number (#170):** This is the patient medical record number provided to STS, sequence #170.

8. **Date of Admission (#570):** The STS Admission Date, sequence #570, value in YYYY-MM-DD format.

9. **Date of Surgery (v610):** The STS Surgery Date, sequence #610, value in YYYY-MM-DD format.

10. **Check Number for Edit Code:** This is the Mass-DAC assigned number for the edit program code and may be used to lookup more details on the actual edit. The first digit represents the type of issue
 - 1=Missing,
 - 2=Out of Range,
 - 3=Out of Usual Range,
 - 4=Parent/Child Inconsistency,

- 5=Date Inconsistency,
- 6=Other multi-variable relationship check.

The remaining digits before the decimal are either Mass-DAC variable number or the STS sequence # preceded by two zeros.

The decimal digits are a subcategory value used by Mass-DAC to distinguish different types of edits for the same issue type.

11. **Checked Edit Name with Suffix:** The edit code Mass-DAC variable name for the primary variable. This is often the same name as the STS Short Name for the variable, but it may also be a Mass-DAC specific variable name that excludes the numeric suffix for repeatable field names.
12. **Checked Variable Name:** This is the actual name of the variable in the Mass-DAC data set. This is often the same name as the STS Short Name for the variable, but it may also be a Mass-DAC specific variable name.
13. **Priority Status of Edit Code:** Defines the importance level of the issue. If this value is 1, then the issue is critical to the Mass-DAC models and must either be corrected or resolved If possible, otherwise the record may cause problems when the models are run. A value of 2 indicates an issue that may need to be corrected, but often times the value is fine as is, it just needs to be checked.
14. **Source of Edit Code:** 1 indicates this is a standard STS edit, and a 2 indicates this is a Mass-DAC specific edit.
15. **Checked Variable Label:** This is the detailed label of the primary variable involved in the flagged issue and is the label for the value located in the Checked Variable Name column.
16. **Checked Value Type:** If N then the variable value is Numeric, if C, then the value is character.
17. **Checked Value formatted:** Provides the actual value of the variable, followed by a dash and the text equivalent for the value of the primary variable found in the Checked Variable Name column.
18. **Assoc. Var # Name:** For each associated variable used in the edit code, provides the STS or Mass-DAC variable Name.
19. **Assoc. Var # Label:** For each associated variable used in the edit code, provides the detailed label for the variable, including the STS sequence number if it is an STS variable.

20. **Assoc. Var # Label:** For each associated variable used in the edit code, provides the value of the associated variable for the Assoc. Var # Name column.