ASSESSING THE RELIABILITY OF COMPUTER-PROCESSED DATA
• 6.66 Auditors should assess the sufficiency and appropriateness of computer-processed information regardless of whether this information is provided to auditors or auditors independently extract it. The nature, timing, and extent of audit procedures to assess sufficiency and appropriateness is affected by the effectiveness of the audited entity’s internal controls over the information, including information systems controls, and the significance of the information and the level of detail presented in the auditors’ findings and conclusions in light of the audit objectives. The assessment of the sufficiency and appropriateness of computer-processed information includes considerations regarding the completeness and accuracy of the data for the intended purposes.
KEY DEFINITIONS

• Data reliability refers to the accuracy and completeness of computer-processed data

• Computer-processed data may be data
  • entered into a computer system or
  • resulting from computer processing
SOME SPECIFIC EXAMPLES OF COMPUTER-PROCESSED DATA

- Data extracts from databases, data warehouses, or data repositories;
- Data maintained in Microsoft excel or access or similar commercial products;
- Data extracts from enterprise software applications supported by information technology departments or contractors;
- Public use data or other replicated detail or summary-level databases accessible through an application other than the original source system;
- Data collected from forms and surveys on web portals; and
- Data summarized in a report or copied from a table in a document.
COMPLETE AND ACCURATE

- Completeness refers to the extent that relevant records are present and the fields in each record are populated appropriately.
- Accuracy refers to the extent that recorded data reflect the actual underlying information.
- Consistency, a subcategory of accuracy, refers to the need to obtain and use data that are clear and well defined enough to yield similar results in similar analyses. For example, if data are entered at multiple sites, inconsistent interpretation of data entry rules can lead to data that, taken as a whole, are unreliable.
What is the type of engagement?

Financial and financial-related audits


All other engagements

Do you anticipate that data will materially support findings, conclusions, or recommendations?

No

If primarily background or contextual information that does not materially affect findings, determine if from best available source

Yes

Do the research question require a determination of the reliability of an information system?

Yes

Conduct a computer system review and disclose in the section on objectives, scope, and methodology the work done, results, and any limitations found

No

Will the data be used in multiple future engagements?

Yes

Should you do a computer system review?

No

Not at this time

Continue with a data reliability assessment
CONDITIONS REQUIRING DATA RELIABILITY ASSESSMENT

• You should assess reliability if the data to be analyzed are intended to materially support your findings, conclusions, or recommendations.
CONDITIONS NOT REQUIRING DATA RELIABILITY ASSESSMENT

• You do not need to assess the reliability of data if their use in the report does not materially affect findings, conclusions, or recommendations.
QUESTION

- You have a finding that includes the number of uninsured Americans. To put that number in context, you report the overall U.S. population.
- Is an assessment needed for:
  - The estimate of the number of Americans who are uninsured
    Yes/No
  - The estimate of the U.S. population as determined by the U.S. Census
    Yes/No
FINANCIAL AUDITS

- You should not follow this guidance in assessing data reliability.
- For financial audits, which include financial statements and financial-related audits, you should follow the Financial Audit Manual and the Federal Information System Controls Audit Manual.
• Can you use the data to answer the research questions
  • Depends on
    • Expected importance of the data to the final report,
    • Strength or weakness of any corroborating evidence, and
    • Anticipated level of risk in using the data.
EXPECTED IMPORTANCE OF THE DATA TO THE FINAL REPORT

• Will the project team depend on the data alone to answer a research question?
• Will the data be summarized or will detailed information be reported?
• Is it important to have precise data?
CORROBORATING EVIDENCE

- Consistent with yellow book standards of evidence - sufficiency and appropriateness;
- Able to provide crucial support;
- Drawn from multiple sources;
- Drawn from multiple types of evidence, such as testimonial, documentary, and physical; and
- Independent of other sources.
RISK LEVEL IN USING THE DATA

- Could be used to inform legislation, policy, or a program that could have substantial effect;
- Could be used to inform important decisions by individuals or organizations with an interest in the subject;
- Will be the basis for numbers that are likely to be widely quoted
- Are relevant to a sensitive or controversial subject;
- Have been judged for their quality by experts or external stakeholders who have taken positions on the information
All phases of assessment are influenced by:
- importance of data to message,
- strength of corroborating evidence, and
- risk of using data

Plan the assessment

Perform data assessment with appropriate mix of work

Make determination

- Sufficiently reliable
- Not sufficiently reliable
- Undetermined reliability
PLANNING A DATA RELIABILITY ASSESSMENT

• Timing an Assessment
  • Early as possible
• Level of Detail of the Data
  • Record-level data
  • Summary-level data
• Documenting the Assessment
STEPS IN THE ASSESSMENT

- Review existing information
- Electronic testing
- Trace sample
- Review selected system controls
REVIEWING EXISTING INFORMATION

• Helps you determine what is already known about the data and the computer processing
• Can indicate both the accuracy and completeness of the entry and processing of the data, as well as how data integrity is maintained
• Can be in the form of reports, studies, or interviews with individuals who are knowledgeable about the data
REVIEWING EXISTING INFORMATION

• GAO, IG, State Auditors, Single Audit Reports
• Agency under review
• Other sources
PERFORMING DATA TESTS

• Data testing can be done by applying logical tests to electronic data files or paper copies of reports. For record-level electronic data, you can use computer programs to test all entries of key data elements in an entire data file.

• For paper copy or summarized data—provided by the agency or retrieved from the Internet—ask for the electronic data file that was used to create them. If you are unable to obtain electronic data, use the paper copy or summarized data and, to the extent possible, manually apply the tests to all instances of key data elements or, if the report or summary is voluminous, to a sample of them.
PERFORMING DATA TESTS

• Checking total number of records provided against agency totals
• Testing for missing data, either entire missing records or missing values in key data elements
• Looking for duplicate records
• Looking for invalid or duplicate identifiers
PERFORMING DATA TESTS

- Testing for values outside a designated range
- Looking for dates outside valid time periods or in an illogical progression
- Following up on troubling aspects of the data—such as extremely high values associated with a certain geographic location—found while analyzing the data
- Testing relationships between data elements (sometimes by merely doing a cross tabulation), such as whether data elements follow a skip pattern from a questionnaire
- Verifying that computer processing is accurate and complete, such as testing a formula used in generating specific data elements, or testing to ensure that edit checks are working correctly.
• You obtained a spreadsheet from the auditee. You confirm record counts and dollar amounts to information provided by the auditee. Are you done testing?
QUESTION

• You obtained a spreadsheet from the auditee. One of the key fields is Receipt Date. What sort of test would you perform on this field?

• After performing your test, you note that there are dates outside the scope of your audit. What do you do?
TRACING TO AND FROM SOURCE DOCUMENTS

• Tracing a sample of data records to source documents helps you determine whether the computer data accurately and completely reflect these documents.
• Consider the relative risks of overstating or understating conclusions drawn from the data.
If you are particularly concerned that questionable cases might not have been entered into the computer system and that, as a result, the degree of compliance may be overstated, consider tracing from source documents to the database.

If you are more concerned that ineligible cases have been included in the database and that, as a result, the potential problems may be understated, consider tracing from the database back to source documents.
• Trace only a sample to save time and cost
• Random sample
• Large enough to estimate the error rate within reasonable levels of precision
TRACING TO SOURCE DOCUMENTS

- Consider tracing to source documents when
  - They are available relatively easily or
  - The possible magnitude of error is especially critical.
• Same spreadsheet. You performed all tests possible on data. You take a sample of transactions and trace them to source documents.
  • You can’t find the source document for a couple transactions. Continue/Stop?
  • Some of the transactions have been amended? Continue/Stop?
REVIEWING SELECTED SYSTEM CONTROLS

• Access Controls
• Application Controls
What is the final determination of reliability?

- Importance of data to message
- Strength of corroborating evidence
- Risk of using data
- Review of existing information (documentation, interviews)
- Results of electronic testing
- Results of tracing to or from source documents
- Results of review of selected system controls

- Sufficiently reliable
  - Use data and disclose limitations
- Not sufficiently reliable
  - Take alternative actions
MAKING THE DATA RELIABILITY DETERMINATION

- Considering the Results of Your Assessment Work
- Outcomes to Consider in the Assessment
  - Sufficiently reliable
  - Not sufficiently reliable
  - Undetermined reliability
  - Professional judgement
• Sufficiently reliable
  • The likelihood of significant errors or incompleteness is minimal and
  • The use of the data would not lead to an incorrect or unintentional message
MAKING THE DATA RELIABILITY DETERMINATION

• Not sufficiently reliable
  • Significant errors or incompleteness in some of or all the key data elements and
  • That using the data would probably lead to an incorrect or unintentional message, given the research questions and intended use of the data
MAKING THE DATA RELIABILITY DETERMINATION

• The source is the only source.
• Should include in the report’s methodology section a statement about having conformed to generally accepted government auditing standards.

• These standards include the appropriateness of the data being used.

• You conform to GAGAS by discussing in the report what you did to assess the data, disclose any data concerns, and make a judgment about the reliability of the data used in the report.
INCLUDING APPROPRIATE LANGUAGE IN THE REPORT

• Sufficiently Reliable Data
  • noting the kind of assessment you relied on,
  • explaining the steps in the assessment,
  • describing any corrections made to the data, and
  • disclosing any data limitations.
INCLUDING APPROPRIATE LANGUAGE IN THE REPORT

- Not Sufficiently Reliable Data
  - describe the problems with the data, as well as why using them would probably lead to an incorrect or unintentional message, and
  - state that the data problems are significant or potentially significant
• Data of Undetermined Reliability
  • Include such factors as the deletion of original computer files,
  • data limitations that prevent an adequate assessment, short time periods,
  • and the lack of access to the data source or to needed documents.
  • Explain the reasonableness of using the data
INCLUDING APPROPRIATE LANGUAGE IN THE REPORT

• We assessed the reliability of _______ data by (1) performing electronic testing of required data elements, (2) reviewing existing information about the data and the system that produced them, and (3) interviewing agency officials knowledgeable about the data. We determined that the data were sufficiently reliable for the purposes of this report.

• We assessed the reliability of _______ data by (1) performing electronic testing of required data elements, (2) reviewing existing information about the data and the system that produced them, and (3) interviewing agency officials knowledgeable about the data. In addition, we traced a statistically random sample of data to source documents (see appendix x for details). We determined that the data were sufficiently reliable for the purposes of this report.
To assess the reliability of the data elements needed to answer the engagement objectives, we (1) performed electronic testing of required data elements, (2) reviewed related documentation, and (3) interviewed agency officials knowledgeable about the data. The results of our electronic testing showed that data elements key to our review contained high percentages of missing data. (See appendix x for further details.) Therefore, we determined that the data were not sufficiently reliable for the purposes of this report.
SSA MASTER DEATH FILE
• Contains the complete and official SSA database extract, as well as updates to the full file of persons reported to SSA as being deceased.

• The Social Security Death Master file contains the following information:
  • Social Security Number, Name, Date of Birth, Date of Death, Proof/Verified Code, State (or country) of last known residence, and ZIP code of last lump sum payment.
SSA MASTER DEATH FILE

- SSA authorizes the use of this database as an identity verification tool, but notes that the Death Master File may contain inaccuracies.
- SSA cannot guarantee the accuracy of the Death Master File. Therefore, the absence of a particular person on this file is not proof that the individual is alive. Further, in rare instances, it is possible for the records of a person who is not deceased to be included erroneously in the Death Master File.
IMPORTANCE

• Sources of Date of Death
  • Match to claims with:
    • Discharge Status Code = Death
    • Medicare Enrollment Database
    • SSA Master Death File
**Corroborating Evidence**

- Per SSA DOD = 07/15/2007

<table>
<thead>
<tr>
<th>HOSH</th>
<th>HOSPICE CLAIM HISTORY</th>
<th>PAGE 1 OF 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIC</td>
<td>xxxxxxxxxxxxx</td>
<td>NORTHEAST</td>
</tr>
<tr>
<td>CORR</td>
<td>xxxxxxxxxxxxx</td>
<td>NAME xxxxxxxxxxxxx</td>
</tr>
<tr>
<td>BOI</td>
<td>1 CMN 3 HBB 1 HHE 1 HOS 1 MSP 1 REP 2 SUR 1</td>
<td></td>
</tr>
<tr>
<td>INF</td>
<td>0008 OUT 0000 HOS 0002 PTB 0045 DME 0001 HHA 0026 BLK 00</td>
<td></td>
</tr>
<tr>
<td>ICN</td>
<td>207219000053002</td>
<td>PROV 331529</td>
</tr>
<tr>
<td>CANCEL ADJ CODE</td>
<td>BILL 834</td>
<td>PAT STAT 41</td>
</tr>
<tr>
<td>FMT DIST</td>
<td>PROV 8,048.91</td>
<td>PEIMB AMT 8,048.91</td>
</tr>
<tr>
<td>RECEIVED DTE</td>
<td>08/06/07</td>
<td>APPROVED DTE 10/24/07</td>
</tr>
<tr>
<td>NOPAY CD:</td>
<td>DOA 10/25/07</td>
<td>FORM: UB92</td>
</tr>
<tr>
<td>DEMO NUMBER</td>
<td>UNSOL IND</td>
<td>PRICER CD 00</td>
</tr>
<tr>
<td>ADJ:</td>
<td>MASS CLM</td>
<td>LIAB IND</td>
</tr>
</tbody>
</table>
RISK

- Risk in using it
- Risk of not using it
  - Overpayments not identified
  - Potential fraud continues
QUICK RESPONSE EVALUATION

Sources of Erroneous Death Entries Input into the Death Master File

A-06-09-29095

January 2009
## Existing Evidence

### Sources of 7,597 Erroneous Death Reports

<table>
<thead>
<tr>
<th>Description</th>
<th>Transactions</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSA Staff Death Entry Input</td>
<td>6,754</td>
<td>88.9</td>
</tr>
<tr>
<td>First-Party Death Report to SSA</td>
<td>311</td>
<td>4.1</td>
</tr>
<tr>
<td>Funeral Home Death Report to SSA</td>
<td>257</td>
<td>3.4</td>
</tr>
<tr>
<td>EDR State</td>
<td>113</td>
<td>1.5</td>
</tr>
<tr>
<td>Non-EDR State</td>
<td>53</td>
<td>0.7</td>
</tr>
<tr>
<td>Veterans Affairs</td>
<td>54</td>
<td>0.7</td>
</tr>
<tr>
<td>Other</td>
<td>55</td>
<td>0.7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>7,597</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

OFFICE OF
THE INSPECTOR GENERAL
SOCIAL SECURITY ADMINISTRATION

PAYMENTS TO INDIVIDUALS
WHOSE NUMIDENT RECORD
CONTAINS A DEATH ENTRY

June 2009    A-06-08-18995

AUDIT REPORT

OFFICE OF THE INSPECTOR GENERAL
SOCIAL SECURITY ADMINISTRATION

USA

SOCIAL SECURITY ADMINISTRATION

OFFICE OF THE INSPECTOR GENERAL

USA
SOCIAL SECURITY
DEATH DATA

Additional Action
Needed to Address
Data Errors and
Federal Agency
Access
EXISTING EVIDENCE

305 Beneficiaries in Current Pay Status
Whose Numident Contained a Death Entry
As of January 2008

Vital Status

- Unverified (77)
- Verified (228)
- 88 Deceased (39%)
- 140 Alive (61%)
ELECTRONIC TESTING
FEDERAL INFORMATION SECURITY MANAGEMENT ACT REPORT

Fiscal Year 2008 Evaluation of the Social Security Administration's Compliance with the Federal Information Security Management Act

September 2008 A-14-08-18063

Patrick P. O'Carroll, Jr. – Inspector General
MEDICARE PART A AND B DEATH MATCH

• Objective:
  • To determine the amount of Medicare fee-for-service payments where
    • From DOS > Date of Death
MEDICARE PART A AND B DEATH MATCH

- Matched Medicare claims data to CMS Enrollment database
  - $67 million
- Matched Medicare claims data to SSA Master Death File
  - $172 million
<table>
<thead>
<tr>
<th>Service</th>
<th>Amount</th>
<th>Claims</th>
</tr>
</thead>
<tbody>
<tr>
<td>DME</td>
<td>$10.1 million</td>
<td>67,789</td>
</tr>
<tr>
<td>HHA</td>
<td>$3.3 million</td>
<td>1,441</td>
</tr>
<tr>
<td>Hospice</td>
<td>$8.1 million</td>
<td>3,949</td>
</tr>
<tr>
<td>Inpatient</td>
<td>$96.3 million</td>
<td>8,052</td>
</tr>
<tr>
<td>Outpatient</td>
<td>$9.1 million</td>
<td>23,906</td>
</tr>
<tr>
<td>Physician Supplier</td>
<td>$32.9 million</td>
<td>241,535</td>
</tr>
<tr>
<td>SNF</td>
<td>$12.1 million</td>
<td>2,633</td>
</tr>
</tbody>
</table>