Guidance on
Testing Data Reliability

January 2004

Office of the City Auditor **Austin, Texas**

City Auditor
Stephen L. Morgan, CIA, CFE, CGAP, CGFM

Deputy City Auditor
Colleen G. Waring, CIA, CGAP, CGFM

Please send any questions or comments to: *taylor.dudley@ci.austin.tx.us*

DATA RELIABILITY TESTING

What is data reliability?

Data reliability is a state that exists when data is sufficiently complete and error free to be convincing for its purpose and context. In addition to being reliable, data must also meet other tests for evidence.

Computer-processed data must meet evidence standards before it can support a finding. For all types of evidence, various tests are used—sufficiency, competence, and relevance—to assess whether the GAGAS standard for evidence is met. Per GAGAS, evidence is:

- *relevant* if it has a logical, sensible relationship to the finding it supports. What data is relevant to answering an audit objective is usually self-evident, presuming a precise objective written as a question. Timeliness (the age of the evidence) must be considered, as outdated data is considered irrelevant. As a result, relevance is closely tied to the scope of the audit work, which establishes what time period will be covered. Data is relevant if they have a logical, sensible relationship to the overall audit objective in terms of:
    - the audit subject
    - the aspect of performance being examined
    - the finding element to which the evidence pertains, and
    - the time period of the issue being audited
- *sufficient* if there is enough of it to support the finding. Sufficiency establishes that evidence or data provided has not been overstated or inappropriately generalized. Like relevance, sufficiency must be judged in relationship to the finding element to which the data pertains, and is closely tied to the audit scope. The audit scope establishes what portion of the universe is covered (important for sufficiency) through 3 choices:
    1. obtain data on (mine) the entire universe
    2. sample the universe
    3. limit findings to that portion or segment of the universe they examine
- *competent* if it both *valid* and *reliable*. In assessing computer-processed data, the focus is usually on one test in the evidence standard—competence—which includes both validity and reliability. Per GAGAS, "Auditors should determine if other auditors have worked to establish the validity and reliability of the data or the effectiveness of the controls over the system that produced it. If they have, auditors may be able to use that work. If not, auditors can obtain evidence about the competence of computer-processed data by direct tests of the data (through or around the computer, or a combination of both.) Auditors can reduce the direct tests of the data if they test the effectiveness of general and application controls over computer-processed data, and these tests support the conclusion that controls are effective." The fundamental criterion for judging data competence is: "Are we reasonably confident that the data presents a picture that is not significantly different from reality?" The criterion is NOT simply "Are we sure the data is *accurate*?" In order to address competence, the data must be more than accurate, it must also be *valid*, *complete*, and *unaltered*.

***Validity* refers to whether the data actually represent what you think is being measured.** For example, is the data field "annual evaluation score" an appropriate measure of a person's job performance? Does a field named "turnaround time" appropriately measure the cycle that it purports to represent? While validity must be considered, this discussion focuses on *reliability*.

DATA RELIABILITY TESTING

Data reliability **refers to the accuracy and completeness of computer-processed data, given the intended purposes for use.** Reliability does not mean that computer-processed data is error-free. It means that any errors found were within a tolerable range - that you have assessed the associated risk and found the errors are not significant enough to cause a reasonable person, aware of the errors, to doubt a finding, conclusion, or recommendation based on the data. Data can refer to either information that is entered into a system or information generated as a result of computer processing. Data is considered reliable when it is:

- *COMPLETE* - includes all of the data elements and records needed for the engagement. A *data element* is a unit of information with definable parameters (e.g. a Social Security #) and is also called a *data variable* or *data field*
- *ACCURATE*:
  - *CONSISTENT* - data was obtained and used in a manner that is clear and well-defined enough to yield similar results in similar analyses.
  - *CORRECT* - the data set reflects the data entered at the source (or if available source documents) and/or properly represents the intended (i.e. calculated) results.
- *UNALTERED* data reflects source and has not been tampered with.

Making a preliminary assessment of data reliability

Simple steps that provide the basis for making a preliminary assessment of data reliability include collecting known information about the data, performing initial testing of the data, and assessing risk related to the intended use of the data.

REVIEW EXISTING INFORMATION. Determine what is already known about the accuracy and the completeness of the entry and processing of the data, as well as how data integrity is maintained. Sources for related information can be found within the agency under review and externally among the customers and data users. This may be in the form of reports, studies, or interviews with knowledgeable users of the data and the system.

Computers are almost always programmed to edit data that is entered for processing. These edits help determine whether the data is acceptable. If a transaction contains errors or fails to meet established edit criteria, it is rejected. A computer record of rejected transactions should be available from the control group responsible for reviewing output. Exercise care in reaching conclusions about these edit tests because a system with insufficient computer edits may routinely accept bad data and reject few transactions, while a system with extensive edits may reject many transactions but actually produce a far more accurate final product. Auditors should ask how management monitors for problems with the computer system while discussing and obtaining standing reports (e.g. for security violations, exceptions, bypasses, overrides, etc.) These discussions and documents are also useful to help review the extent of any known problems with the system.

PERFORM INITIAL TESTING. Apply logical tests to electronic data files or hard copy reports.

- For electronic data, use computer programs to test all entries of key data elements you plan to use for the engagement. Testing with computer programs (e.g. Excel, Access, ACL, SPSS, etc.) often takes less than a day, depending on the complexity of the file.

- For hard copy or summarized data (whether provided by the audited entity or retrieved from the internet) you can ask for the electronic data file used to create it. If you are unable to obtain electronic data, use the hard copy or summarized data and, to the extent possible, manually apply the tests to all key data elements or (if the report or summary is too voluminous) to a sample of them. *Be sure to keep a record or log of your testing for your workpapers!* Whether you have an electronic file or a hard copy report or summary, you apply the same tests to the data, which can include testing for things like:
  - missing data (either entire records or values of key data elements)
  - the relationship of one data element to another (e.g. male patients with prenatal care)
  - values outside of a designated range (e.g. driver's license age under 14)
  - dates outside valid time frames or in an illogical progression (e.g. died before born!)

ASSESS RISK RELATED TO DATA RELIABILITY. In making the preliminary assessment, consider the data in the context of the final report and how big a role the data will play:

- Will the audit depend on the data alone to answer a research question (objective)?
- Will the data be summarized or will detailed information be required?
- Is it important to have precise data, making the magnitude of errors an issue?

You should consider the extent to which corroborating evidence is likely to exist and will independently support your findings/recommendations. *Corroborating evidence* is independent evidence that supports information found in the database. Such evidence, if available, can be found in the form of alternative databases or expert views. Corroborating evidence is unique to each engagement, and its strength (or persuasiveness) varies - for help in judging the strength or weakness of corroborating evidence, consider the extent to which it:

1. meets Yellow Book standards of evidence (sufficient, competent, and relevant),
2. provides crucial support,
3. is drawn from different types of sources (testimonial, documentary, physical, or analytical),
4. is independent of other sources.

Risk is the likelihood that using data of questionable reliability could have significant negative consequences to the decisions of policymakers and others. A risk assessment should consider the following risk conditions:

- The data could be used to influence legislation or policy with significant impact.
- The data could be used for significant decisions by individuals or organizations.
- The data will be the basis for numbers that are likely to be widely quoted.
- The engagement is concerned with a sensitive or controversial subject.
- The engagement has external stakeholders who have taken positions on the subject.
- The overall engagement risk is medium or high.
- The engagement has unique factors that strongly increase risk.

Bear in mind that any one of the above conditions may have more importance than another, depending upon the engagement. Be sure to document *in the workpapers* your analysis of risk in terms of the role the data is to play in your audit and the use of it by other people versus the strength of corroborating evidence.

SUMMARIZE PRELIMINARY RESULTS. The overall assessment of reliability is a judgment call. The outcome of the assessment is going to vary based upon your combined judgments of the strength of the corroborating evidence and the degree of risk involved. If the corroborating

evidence is strong and the risk is low, the data is more likely to be considered sufficiently reliable for your purposes. If the evidence is weak and the risk is high, the data is more likely to be considered not sufficiently reliable for your purposes. Your preliminary assessment should conclude that the data is either:

- *sufficiently reliable* or
- *not sufficiently reliable* or
- *of undetermined reliability*

Data is *sufficiently reliable* for engagement purposes when you conclude the following: Both the review of related information and the initial testing provide assurance that 1) the likelihood of significant errors or incompleteness is minimal and 2) the use of the data would not lead to an incorrect or unintentional message. You could still have some problems or uncertainties about the data, but they would be minor, given the research question (objective) and intended use of the data. When preliminary assessment indicates that the data is *sufficiently reliable,* use the data.

Data is *not sufficiently reliable* for engagement purposes when you can conclude the following: The review of related information or initial testing indicates that 1) significant errors or incompleteness exist in some or all of the key data elements and 2) using the data would probably lead to an incorrect or unintentional message. When the data is *not sufficiently reliable,* you should seek evidence (to support your finding) from other sources, including 1) alternative computerized data (the reliability of which you should also assess) or 2) original data in the form of surveys, case studies, expert interviews, etc.

If seeking evidence from other sources does not result in a source of *sufficiently reliable* data, you should 1) inform the requestor (City Auditor or requesting Council Member) that the data needed to respond to the request are unavailable and 2) reach an agreement with the requestor to:

- ❑ redefine the research question (objective) to eliminate the need to use the data, or
- ❑ end the engagement, or
- ❑ use the data with appropriate disclaimers.

Remember that YOU, not the requestor, are responsible for deciding what data to use. If you decide you must use data that you have determined is *not sufficiently reliable,* make the limitations of the data clear, so that incorrect or unintentional conclusions will not be drawn. Finally, given that the data you assessed has serious reliability weaknesses, you should include this finding in the report and recommend that the agency take corrective action.

You can call data *of undetermined reliability* when you conclude one of the following:

- The review of some of the related information or initial testing raises questions about the data's reliability
- The related information or initial testing provides too little information to judge reliability
- The time or resource constraints limit the extent of the examination of related information or initial testing

If you conclude the data is *of undetermined reliability,* then you need to go through the same steps outlined above for resolving data that is *not sufficiently reliable* (e.g.. contact the City Auditor, revisit the use, disclaim any presentation, include reliability as a finding, etc.)

Making a final assessment of data reliability

If your initial assessment has concluded that the data is *of undetermined reliability*, more work may be desired in order to help conclude whether to use the data or not. Several additional testing techniques can be employed. If the data cannot be tested adequately, an alternative is to perform a *system control review*. However, it must result in one of the conclusions noted above for the preliminary assessment, and follow the same path for resolution.

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 reflects these documents. In deciding what and how to trace, consider the relative risks to the engagement of overstating or understating the conclusions drawn from the data.

For example, if you have concerns that questionable records may have not have been entered onto the computer system, then you would want to start with a set of source documents and trace them to the database. On the other hand, if you suspect that some records are being inappropriately added to the database, you should consider tracing from the database back to source documents.

Tracing only a sample saves time and money. In order to be useful, the sample should be random and large enough to estimate the error rate within reasonable levels of precision. Tracing a sample file will provide the error rate and the magnitude of errors for the sample, which can then be extrapolated to the entire data file. It is this error rate that helps you to determine the data reliability. Generally, every data file will have some degree of error, and assessment requires judging a combination of:

- *error rate* - the frequency with which an error occurs. For example, a random sample shows 10% of records have the incorrect date. However, the dates may only be off by an average of two days and, depending on what the data is to be used for, two days may not compromise reliability.
- *error magnitude* - the size of the errors found can impact our judgment. For example, the value of a record was listed as $10,000 rather than $100,000. The valid range for this data element in the database is $200 to $1,000,000 and thus the data is within the accepted range of values and the error would not have been caught by a simple preliminary assessment test of value ranges. This type of error would likely only be revealed by tracing the data back to source documents.

Obviously, if source documents were destroyed, were never created, or are not centrally located, the actual tracing cannot be accomplished, and missing data cannot be identified. However, one can still gain some insight into the process by interviewing data sources, owners, and users to obtain any related information or any corroborating evidence obtained earlier, or to review the adequacy of system controls.

USING ADVANCED ELECTRONIC TESTING. Advanced electronic testing goes beyond the basic electronic testing you did in your preliminary assessment and requires the use of specialized software (e.g. ACL, SPSS, Excel) to test for specific conditions within the data. Since the use of software allows the auditor to "mine" the entire dataset rather than rely upon a sample, this type

of testing can be particularly useful in determining the accuracy and completeness of processing by the application that produced the data. Potential examples of advanced tests include:

- following up on troubling aspects of the data (e.g. if your preliminary assessment found extremely high values associated with certain sets of records)
- testing relationships between data elements (e.g. looking for unusual or abnormal correlations between data fields, such as skip patterns)
- verifying that processing is accurate and complete (e.g. thorough testing of a computer formula used to generate specific data elements)

UNCONDITIONAL AND CONDITIONAL TESTS. Another way of articulating electronic data tests is to think of them as either *unconditional tests* (where the data must meet an established requirement) or *conditional tests* (where the data is compared as a logical relationship to other data elements). Logically, both types of tests should be considered for any assignment:

- *Unconditional tests* should be performed first to assure that data elements exist and conform to basic requirements (for example, a field such as "# trained" should not be negative). These tests disclose failures of data elements to meet established requirements. Examples include:
  - range or value limits
  - presence/absence
  - proper dates
  - positive/negative signs
  - alpha/numeric formats
  - formula derivations.

  Unconditional tests can be done as a sample or can be mined for entire universe by using software to test the values. Either way, unconditional tests are a good way to evaluate the necessity of expanding or curtailing *conditional tests*.
- *Conditional tests* involve defining and comparing logical relationships among data elements - these can be formally defined or just common sense (e.g. dates out of logical sequence or male patients with hysterectomies). These tests are not limited to comparing elements within the same computer system, but can include rules that compare against legislative intent, program policies, or other computer systems. Typically, after the unconditional tests of known requirements are performed, critical data elements are subjected to conditional tests to assure that the logical relationships exist and are appropriate. Again, as with unconditional tests, software can be employed to check all the records in question rather than just a sample.

SYSTEM CONTROL REVIEWS. *System Reviews* concentrate on the working of the system, not the administrative program, and they cannot substitute for data testing! Some degree of data testing is always needed because even when system controls are well designed and adhered to, data accuracy is not ensured. However, less data testing is needed when controls are satisfactory than when controls are weak or undetermined.

Why would you undertake a system control review? To provide a level of assurance that either makes YOUR current data tests easier, faster, and less expensive to perform OR makes future auditors' testing of data easier, etc. This brings into play two huge (and doubtful) IFs:

- IF you can assert that control reviews will actually reduce data testing that must accompany it or
- IF anyone else will actually be doing more work (at least within a near enough timeframe that the system and controls will still be substantially the same).

DATA RELIABILITY TESTING

A system review examines and permits expressing an opinion on the entire computer system. However, they are very time consuming to perform, require special expertise, and ultimately only identify potential problems. Thus, system reviews may not reduce data testing on a particular audit, but can facilitate testing over multiple audits using the same systems (i.e. using specific data produced by the system for many different assignments over an extended period.)

LIMITED CONTROL REVIEWS. Auditors frequently skip system reviews and go straight to testing the critical data elements necessary to support their findings. This approach is known as a *Limited Review*, as it is limited to only data reviewed, and provides no basis for reliance on the system overall, or over time. Limited reviews are faster, cheaper, easier, require less expertise, and identify actual problems. If auditors decide to conduct a limited review, then (per GAGAS) they must document for the working papers why it is inefficient to test controls (i.e. do the system review).

In addition, "auditors should also include audit documentation regarding their reasons for concluding that the planned audit procedures, such as direct tests of the data, are effectively designed to achieve specific audit objectives. This documentation should address:

- the rationale for determining the types and extent of planned audit procedures;
- the kinds and competence of available evidence produced outside a computerized information system; and
- the effect on the audit report if the evidence gathered during the audit does not allow the auditors to achieve audit objectives."

Before deciding on what level of review is necessary, the first step is to answer a basic set of questions concerning what is known about the system:

- Has the data already been reviewed for reliability and used to support other work?
- Has an audit system review established the adequacy of system controls?
- Have significant changes taken place since the review?
- Has work by other audit groups assessed either individual applications, data, or general controls in related areas which would provide a basis for reliance?

The overall system of internal control is conceptual in nature. It is the integrated collection of controlled systems used by an organization to achieve its goals and objectives. Internal controls came into general use to distinguish controls within an organization (e.g. policies) from those existing externally to the organization (e.g. laws). Thus, from an organization's viewpoint, internal controls include everything management has put in place to accomplish the organization's goals and objectives.

In terms of computer applications, these controls will usually include:

- *General controls* - the structure, policies, and procedures which apply to all (or a large segment of) an organization's information systems and are intended to help ensure proper operation, data integrity, and security. Typical examples include policies for hiring, training, physical access, data backup and recovery, etc.
- *Application controls* - the structure, policies, and procedures that apply to individual application systems, such as inventory or payroll. These controls address issues such as the accuracy of data input and the verification and distribution of output and are typically intended to help ensure that data is accurate and complete, as well as authorized.

DATA RELIABILITY TESTING

Another way to think of control is by the traditional I-P-O application processing phases:

- *Input controls* provide reasonable assurance that data received for processing :
  - was properly authorized, converted into machine form, and identified
  - has not been lost, suppressed, added, duplicated, or otherwise improperly changed
  - has controls for rejection, correction, and resubmission of initially incorrect data.
  - includes checks programmed into the software for things such as financial totals, self-checking digits, hash totals, sequence checks, limit and range checks, sign checks, reasonableness tests, validity checks, record counts, key verification, redundancy checks, echo checks, completeness checks, etc.
- *Processing controls* provide reasonable assurance that:
  - all transactions are processed as authorized
  - no authorized transactions are omitted
  - no unauthorized transactions are added

  Some input controls such as limit, reasonableness and sign checks are also processing controls. Other processing controls include posting checks, end-of-file procedures, cross-footing, concurrency controls, zero-balance checks, audit trails, run-to-run control totals, key integrity checks, internal header and trailer labels, etc.
- *Output controls* assure:
  - the accuracy and completeness of the processing result (e.g. account listings, reports, magnetic files, invoices, disbursement checks, etc.)
  - that only authorized personnel receive the output.

  Output controls include end-of-job markers, error listings, spooler controls, console logs, daily proof account activity listings, distribution registers, and data control groups (tells you that all reports that s/b generated are generated).

If, during your preliminary assessment of data reliability, you learn that source documents are unavailable or you identify potential system control problems, then reviewing the underlying structures and processes of the computer in which the data is maintained can provide some assurance that the data is sufficiently reliable. Examples of system controls (see above) include passwords and edit checks on data entry. Controls can reduce, to an acceptable level, the risk that a significant mistake could occur and remain undetected and uncorrected. You should limit the review to evaluating the specific controls that can most directly affect the reliability of the data in question and choose areas for review based on what is already known about the system.

SAME MENU OF CONCLUSIONS. If additional testing of data *of undetermined reliability* is conducted, you still must conclude whether the data is reliable, not reliable, or still unknown. If you can't say it is reliable, you must communicate the lack of assurance on reliability of the needed data with the City Auditor and then decide if and how to use it. Again, remember that you are not attesting to the reliability of the data or database. You are only determining the sufficiency of the reliability of the data for your intended use. The following are some considerations to help you decide whether you can use the data:

- The corroborating evidence is strong
- The degree of risk is low
- The results of additional assessment 1) answered issues raised in the preliminary assessment and 2) did not raise any new questions

Sources:
Assessing the Reliability of Computer Processed Data, USDA Graduate School Government Auditor Training Institute, March 1995 Participant Workbook.
Assessing the Reliability of Computer-Processed Data, United States General Accounting Office Applied Research and Methods, Publication # GAO-03-273G, October 2002 External Version 1.
Performance Auditing: A Measurement Approach, The Institute of Internal Auditors, 2001.
Government Auditing Standards, 2003 Revision, United States General Accounting Office, Publication # GAO-03-673G, June 2003.