

## QA/QC Automated Data Evaluation Procedure

By Chris Herrington, Analyst  
Environmental Resource Management Division,  
Watershed Protection & Development Review Department (WPDRD)  
City of Austin, Texas

In order to standardize quality assurance and control procedures for the Water Resource Evaluation (WRE) Field Sampling Data Base (FSDB), an automated method of evaluation was programmed. The following report provides a summary of these procedures including logic and potential uses in screening data for analysis and presentation. A two stage data entry method is used to reduce transcription errors and allow field personnel to review laboratory and field results as they are entered into the database. Flags attributing descriptors of data quality are used to characterize QA/QC results for individual samples and parameters. Accuracy, precision, and laboratory blank data are used to generate the final flags for each data point. Additional modifications for the future are presented including ion balances, post calibration checks, and holding time verification in hours.

### Overview of Function and Deployment

- **What is the Data Approval and Flagging Process?**

The Data Approval and Flagging Process is an automated evaluation of the Quality Assurance/Quality Control (QA/QC) data associated with a given sample result which, according to predetermined control limits, results in the generation of a single-character flag enabling both Water Resource Evaluation (WRE) staff as well as members of the general public who request data, to have an instant understanding of the general accuracy and precision of that data point. QA/QC data include, but are not limited to, laboratory control standards, laboratory and field replicate samples, matrix spikes, and blank samples. QA/QC data can be large in quantity (a single data point may have many QA/QC data points) and complex, making assessment of data quality difficult. Thus, with the flag generated by the data approval process the data user may have an instant understanding of the validity of any given data point.

The flagging process consists of two modules, preliminary and final review, that increase in complexity and contain multiple procedures that are activated in a pre-determined sequence in an attempt to minimize database computation time and simply organization of the process code itself. Both modules have been programmed in PL/SQL, and the code for the packages are available upon request from Field Sample Database Staff. These modules are controlled and activated by a central program unit named QC\_FLAG (Figure 1), which is the primary program unit and the unit called by the forms to activate the process.

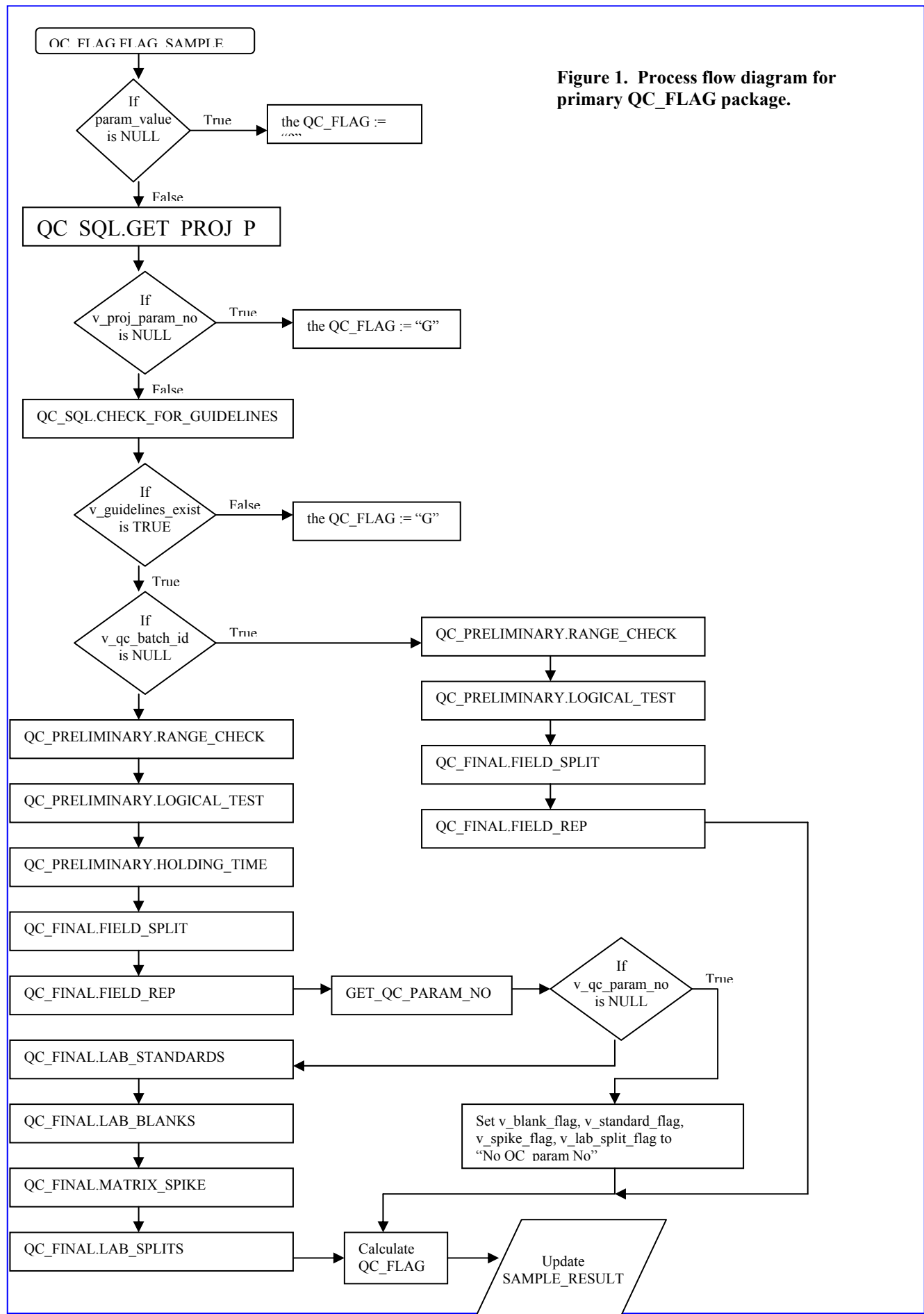
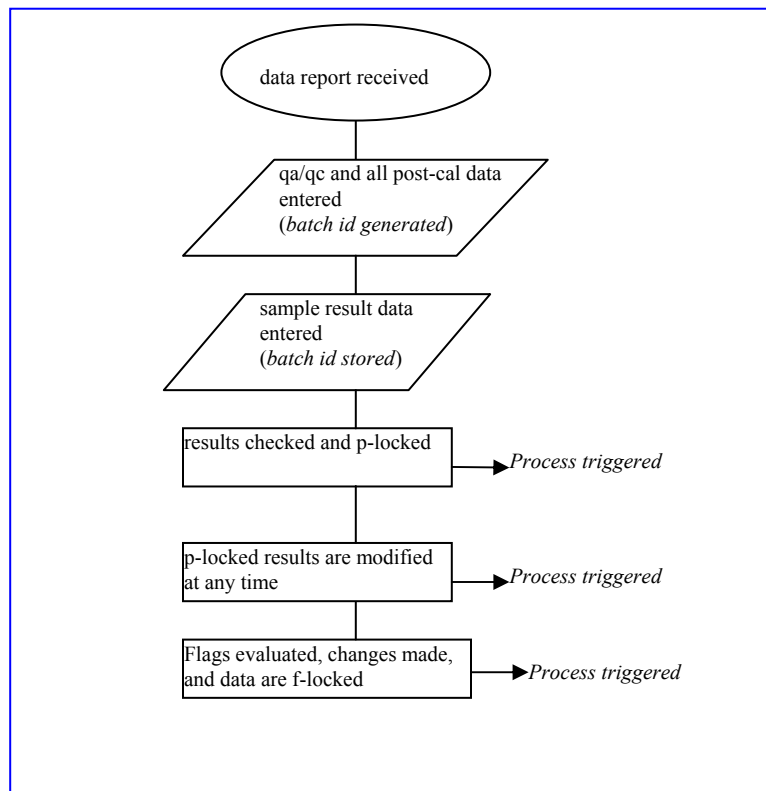


Figure 1. Process flow diagram for primary QC\_FLAG package.

- **When does the process flag the data?**

The process will “fire,” or flag the data, whenever a field sample is locked with either a primary (a value of “P” in the DATA\_LOCKED column of the FIELD\_SAMPLE table) or secondary (a value of “F” in the DATA\_LOCKED column of the FIELD\_SAMPLE table) data lock, or if modifications are made to locked data. *If there are no test guidelines associated with a given sample result, then the process will not be activated.* If guidelines have been specified for a given result, then the portions of the process will fire even if there is no associated QA/QC data. The final review will fire only if guidelines have been specified and if QA/QC data is present.

Database staff recommend that QA/QC data be entered prior to entering sample results (Figure 2). Thus, a QC Batch ID will be generated and may be entered with the sample result data.



**Figure 2. Recommended data entry flow diagram.**

- **What is the two-stage system for locking data?**

Because the process fires when the data is locked and to increase the efficiency of the data entry process, the database staff have established a two-stage locking data locking procedure. In the past, field samples could be locked by anyone simply by placing an “X” in the Lock line of Field Sample Entry Form.

The two-stage procedure now in effect is slightly different in that anyone may still enter data. Once the data has been entered and is correct, then anyone can lock that data by placing a “P” in the DATA\_LOCKED field of the Field Sample Entry Form (Figure 3). Because the sample is locked, the Data Approval and Flagging Process will then fire and generate flags for that data. Once data has been locked with a “P,” or preliminary data lock, then only the database staff and the designated project representative may make any changes to that data.

One of the reasons for a two-stage procedure is that the Data Flagging and Approval Process may catch some data entry errors which will be noted by the final flag applied to the data. With the two-stage procedure, the project representative now has the opportunity to re-examine the data, make any changes, check the data once again, and then apply the final data lock by placing an “F” in the DATA\_LOCKED field of the Field Sample Entry Form.

Please note that the project representatives have the responsibility to apply final locks to all of their respective data once the preliminary lock has been performed and the project representative confirms that the data is completely correct.

The screenshot shows the 'Field Sample Entry Form' in Oracle Developer Forms Runtime. The window title is 'Oracle Developer Forms Runtime - [Field Sample Entry Form]'. The 'Field Visit' tab is selected. The 'Field Visit' section includes a 'Sample Site Name' field and 'Next Visit' and 'Previous Visit' buttons. The 'Field Sample Info' section contains a table with the following columns: Ref No, Time, Lab, Medium, Sample Type, Depth (m), Sample ID, QC Type, Weather, Flow Type, Comments, Taxo, and Lock. The first row of the table is highlighted in yellow. At the bottom of the form are buttons for 'Enter Query', 'Run Query', 'Save Changes', 'Delete Record', 'Flag Sample', 'Help', and 'Exit/Cancel'.

**Figure 3. Screen capture of the Field Sample Entry Form. The data lock field is the right-most column of the form and may be set to “P” or “F” indicating preliminary and final data lock, respectively.**

- **What are the additional data quality considerations not considered by the process?**

The automated flagging process does not remove project staff from all responsibility to ensure the quality of data being collected. Full data review procedures are specified in the current Water Resource Evaluation Section Standard Operating Procedures Manual. Project staff must also be aware that there are additional considerations which should be addressed that are not handled at the present time by the flagging process.

A review of data documentation should be conducted by project staff to verify that the chain of custody has been maintained, sample identification is consistent, and all deviations from planned project guidelines are identified in the database for further consideration. This document review should be performed by the person responsible for officially receiving laboratory reports for a given project. Please note that the preliminary review will not be performed by the database.

Important items to be considered in the document review include:

1. Comparing chain of custody forms, field logbooks, data entry logbooks and laboratory data summary to verify the accuracy of all sample identification names or numbers, sampling dates and project personnel from the time of sample collection to production of final sample report by the primary laboratory.
2. Checking that all samples submitted for analysis have a value for all requested parameters.
3. Note any comments from field data sheets, field logbooks, or laboratory packets, that indicate possible data integrity corruption such as missing samples or incorrect analysis methods.

## Data Flags and Test Guidelines

The automated data evaluation process is a multi-step procedure in which individual sample results are qualified according to the acceptability of associated QA/QC data by the database. Laboratory data will be automatically assessed to determine whether data should be rejected or qualified as usable to some degree by assessing adherence to data quality objectives related to precision, accuracy and distinction from background noise. For the purposes of this document, accuracy will refer to the difference between the reported value of the sample and the true value. Precision will refer to difference between repeated values for the same sample. Accuracy of data is evaluated by comparing laboratory control standards and matrix spikes to test guidelines. Precision of data is evaluated by comparing split and replicate samples generated in both the field and laboratory to test guidelines. Difference from noise is established by comparing blank results to sample results.

Please note that the data qualification process is performed by the database. The procedure consists primarily of two modules, preliminary and final review, that work in conjunction to generate a single-character flag stored in the SAMPLE\_RESULT table.

**Table 1. List of the potential final flags generated by the data evaluation process and stored with the sample result in the SAMPLE\_RESULT table.**

<b>QC Flag</b>	<b>Description of Flag</b>
U	Results are completely usable
R	Results are completely unusable
J	Results are to be considered an estimated value
H	Holding time was violated but data is not rejectable
S	Result was out of standard range but not rejectable
G	No test guidelines have been specified for the result, and the process was not activated
A	QC data are missing where guidelines have been specified
?	No sample result data exists

In the current version, the data evaluation process considers eight types of QA/QC analyses. The guidelines for these tests originate from information distributed by contract laboratories or published in guidance documents such as Standard Methods. Where no information could be found for a given parameter and QA/QC test, guidelines were assumed by database staff using conventional assessment limits (Table 2).

**Table 2. Summary of general guidelines for QA/QC tests considered by the Data Evaluation Process. Note that a “split” identifies two or more samples generated from the same larger container whereas a “replicate” identifies two or more samples collected from the same location at approximately the same time in different sample containers.**

<b>QA/QC Test</b>	<b>General Limits</b>
Standard Range Check	The 15 <sup>th</sup> and 85 <sup>th</sup> percentiles of all data were used as the lower and upper limits of the standard range check for each specific project-medium-lab-parameter combination.
Laboratory Splits	Percent coefficient of variation must be less than or equal to 10% between laboratory split(s) and sample.
Field Splits	Percent coefficient of variation must be less than or equal to 20% between field split(s) and the sample.
Logical Test	Generally, no dissolved, suspended, or filtered result may be greater than a total result for a given project parameter combination. Additional logical tests apply for nitrogen results (NH <sub>3</sub> , Organic N must be less than TKN) and phosphorus results (orthophosphorus must be less than total phosphorus).
Blank Samples	Sample results must be greater than blank results
Laboratory Control Standards	Percent recovery of lab control standards spikes must lie between 80 and 120%.
Laboratory Matrix Spikes	Percent recovery of lab matrix spikes must lie between 80 and 120%.
Holding Times	Current holding times range from 0 (analysis must be performed on the same day) to 180 days, and result from information listed in Standard Methods or lab documents.

## Preliminary Review Module

The first sub-part of the data evaluation process is contained within the package “QC\_PRELIMINARY” and contains the code for the Range Check (Figure 4), Holding Time (Figure 5) and Logical Test (Figure 6) evaluations. Note that the entire process is controlled by the QC\_FLAG package detailed in Figure 1. Though labeled “preliminary,” this designation arises from an organizational need to compartmentalize the code for the application more than an actual separate staging of the process. These units have been separated into the “preliminary” package because the range check and logical tests may be performed on any result with specified guidelines, regardless of the presence of QA/QC data. The holding time evaluation unit was grouped into this package because it also is not a function of specific laboratory chemical analyses but rather is dependent upon the time at which lab analyses were performed.

The first portion of the preliminary review is to check if control limits have been specified for the given result. If no test guidelines exist, then the process stops and the associated sample result is flagged **G** (no guidelines). If guidelines are present, then the preliminary review continues according to the three data elements described below:

1. A **logical test** to compare parameter results to results for related parameters should be performed wherever applicable. An example of a logical test may be that if a parameter X is defined as the sum of the values of parameter Y and parameter Z, then the values for parameters X and Y may not individually be greater than the value for parameter X. Failure of the logical test results in the data point being rejected. If multiple logical tests exist, then a single failure will terminate the procedure and result in a rejected value.
2. A **range check** to compare results to a standard set of expected minimum and maximum values should be performed. The purpose of a range check is to identify any unusual results that should be examined in greater detail by the project personnel.
3. Check **holding times** to ensure that sample analysis occurred within a time period that allows for stable parameter concentrations in order to maintain completeness of the data. For the purpose of this data review, holding time is defined as the time beginning from sample collection and ending at the time of sample analysis measured in days.

If the holding time for a given analyte has been exceeded, if any applicable logical tests fail, or if sample results exceed a standard range, the sample will be qualified accordingly to indicate that the data is implicitly flawed and must be reviewed by project staff members to determine source of the error and the ultimate usability of the data.

1. If the holding time has been exceeded, the sample is flagged **H**.
2. If any applicable logical tests fail (L), the sample is flagged **R** and is considered rejectable.
3. If the results exceed the specified standard range, the sample is flagged **S**.

Once the preliminary review has been completed, the final module is activated if guidelines exist for final tests and qa/qc data is present.

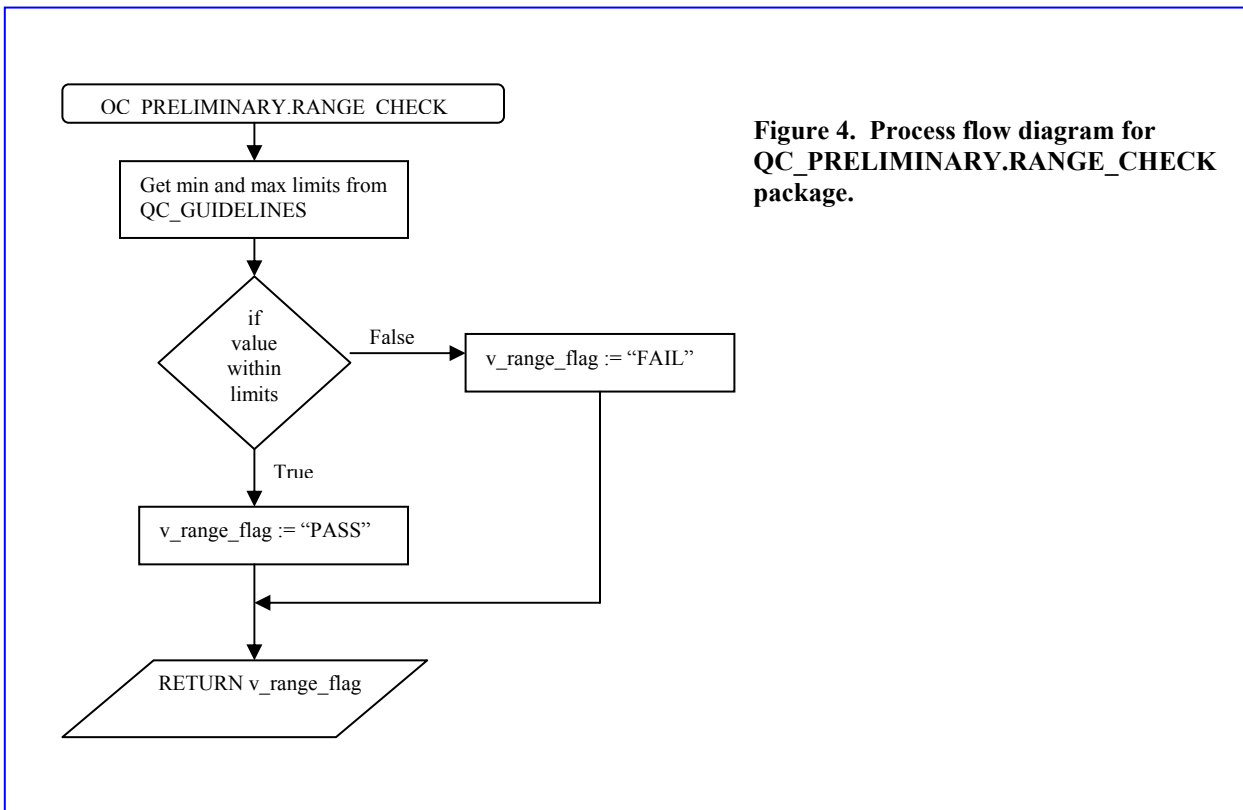
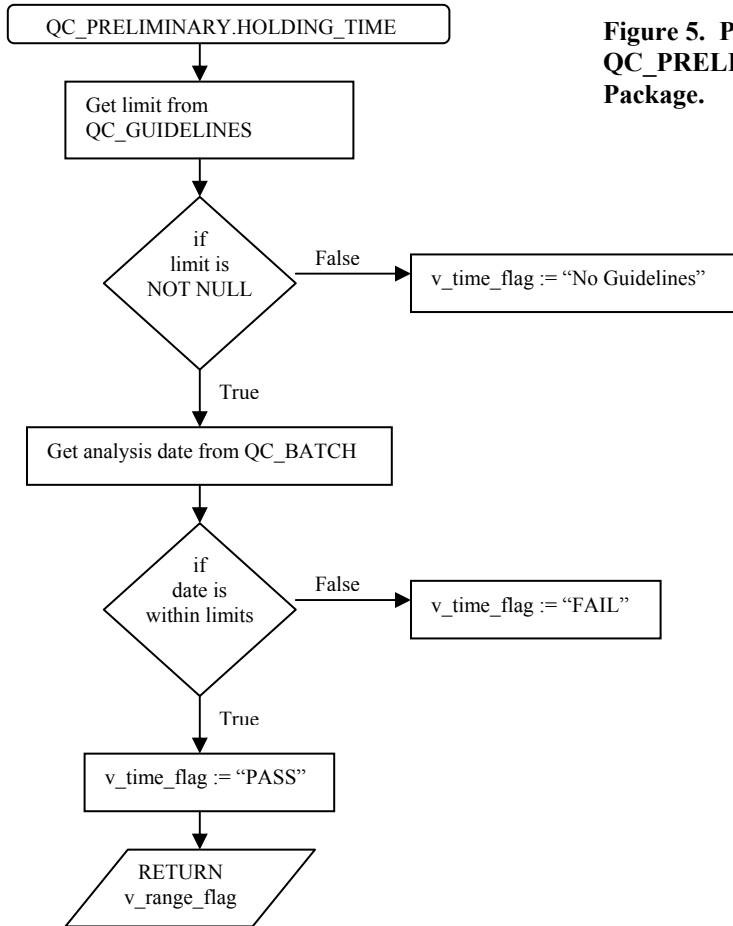
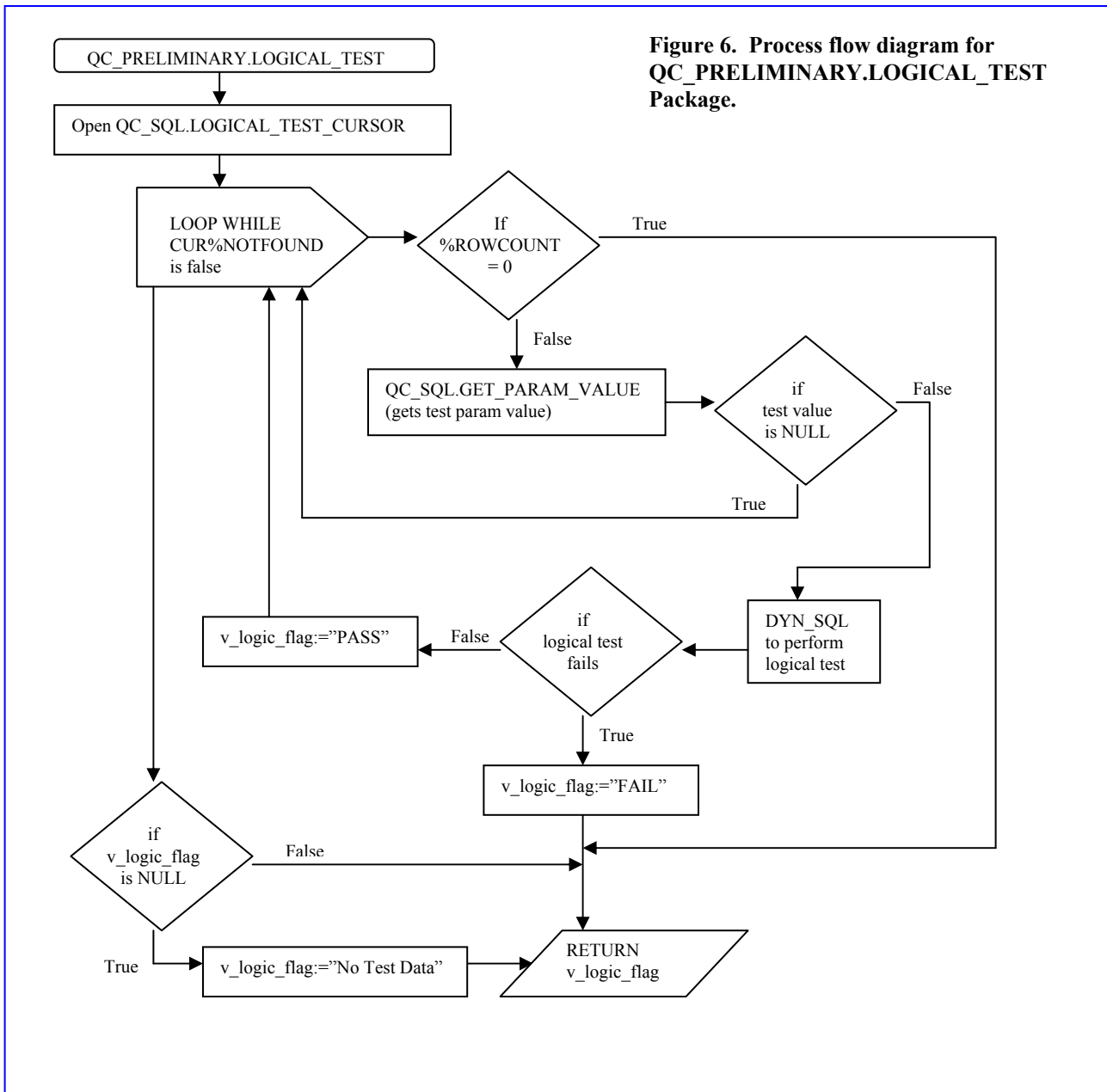


Figure 4. Process flow diagram for QC\_PRELIMINARY.RANGE\_CHECK package.



**Figure 5. Process flow diagram for QC\_PRELIMINARY.HOLDING\_TIME Package.**

**Figure 6. Process flow diagram for QC\_PRELIMINARY.LOGICAL\_TEST Package.**



## Final Review Module

The final module of the QA/QC Data Evaluation Process assesses the accuracy, precision and difference from noise of a given data point. Although the word “Final” is used to describe this module, that moniker is more for organizational purposes. The final module contains those qa/qc tests performed generally performed by the contract laboratory and include standards, spikes, blanks duplicates. Included in the final review module are the Lab Blank procedure (Figure 7), the Lab Split procedure (Figure 8), the Field Split and Replicate procedure (Figure 9), the Matrix Spike procedure (Figure 10) and the Lab Standard procedure (Figure 11).

In general, the final review process uses the “worst-case scenario” when multiple results for a given test are present. Thus, a single failure will result in the termination of that process and a failure flag as the outcome of that process. Conversely, all tests present must be within control limits to obtain a positive test outcome.

The final review module contains the following procedures described below:

1. Results from **laboratory blank** analyses are evaluated to demonstrate that target analytes have not been introduced to the sample as a result of the sample handling procedures used in the laboratory, sample transportation process, or sample collection process, and that detected results are distinct from background interference. Lab blank analyses will be compared to sample result values to determine if sample results are greater than the blank value or if sample result values are less than or equal to the blank value. If more than one blank has been analyzed, then a single failure will terminate the procedure and result in a failed blank flag. *Future modifications to the flagging process may add a multiplier to the blank value based on a relative standard deviation for multiple blank values, when present, for a given analyte.*

<b>If the sample result is:</b>	<b>The blank flag is:</b>
≥ blank value	OK
≤ blank value	Bad

2. Quality assurance of the **precision**, or reproducibility, of the data will be evaluated with lab and field replicate and split samples. Field replicate samples are samples collected in individual bottles at approximately the same place and time, thereby yielding some estimate of the potential heterogeneity present at the time of sample collection. Field split samples are similar to field replicate samples, although field splits are generated from the same sample container separated into multiple smaller containers and thus should exhibit a minimum of natural variation. Laboratory replicates, such as those generated by production laboratories, are evaluated by the same standards as split samples. Because field split and lab replicates should vary less than field replicate samples, precision tests are differentially weighted. For *any* one precision test, the results of multiple results for a given analyte are evaluated based on comparison of project guideline control limits to the percent coefficient of variation

between the sample and split or replicate value. The process makes a distinction between “inside” and “outside” split or replicate samples. It is the customary practice to collect one split or replicate samples for every ten samples collected in the field. The result from the comparison of the sample to the split or replicate will then be applied to all sites for that sampling event. The sample which directly corresponds to the split (i.e., the site at which the split/replicate sample was generated) is referred to as an “inside” split or replicate. All other samples will use the result of the comparison of the inside split or replicate, and those samples are referred to as “outside” split or replicates. If more than one split or replicate sample has been collected, a single failure will terminate the process and result in a “failure” flag. Inside split or replicate samples are evaluated by the process first, then outside split samples. *Future modifications to the process may be made to dynamically set the precision criteria based on historically observed natural variation for a given analyte.*

<b>Lab/Field Splits Result:</b>	<b>Field Replicate Result:</b>	<b>Precision Flag is:</b>
Outside control limits	Outside control limits	Bad
Outside control limits	Within control limits	Bad
Within control limits	Outside control limits	Estimated
Within control limits	Within control limits	OK
<i>Conflicting lab &amp; field results</i>	n/a	Estimated

If only one member of the precision checks are present, then the precision flag will be based entirely on that check. If both members of the precision check are absent, then the precision flag is absent. If lab and field splits are performed for the same batch but have conflicting results, then the precision flag will be estimated.

- Quality Control of the **accuracy**, or correctness, of the data will be evaluated with laboratory control standards and matrix spikes. Laboratory control standards, obtained from an independent source of the calibration standards, are used to determine if the laboratory is operating within control limits in order to demonstrate accuracy in the sample results. Laboratory matrix spikes are known additions of analyte to known sample in order to indicate what effects the sample matrix might be exerting on the accuracy of the sample analysis methods. Relative percent recovery values for matrix spike and control standard results are compared to established guidelines. If multiple QC results for any one accuracy test are present, the results of QC test closer to the range of the sample result are used. If multiple QC results for any one accuracy test are present and both are in the range of the sample result, then a single failure will terminate the process and result in a “failure” flag.

<b>Control standard result is:</b>	<b>Matrix spike result is:</b>	<b>Accuracy flag is:</b>
Within control limits	Within control limits	OK
Within control limits	Outside control limits	Estimated
Outside control limits	Outside control limits	Bad
Outside control limits	Within control limits	Bad

If only one member of the accuracy checks are present, then the precision flag will be based entirely on that check. If both members of the accuracy check are absent, then the accuracy flag is absent.

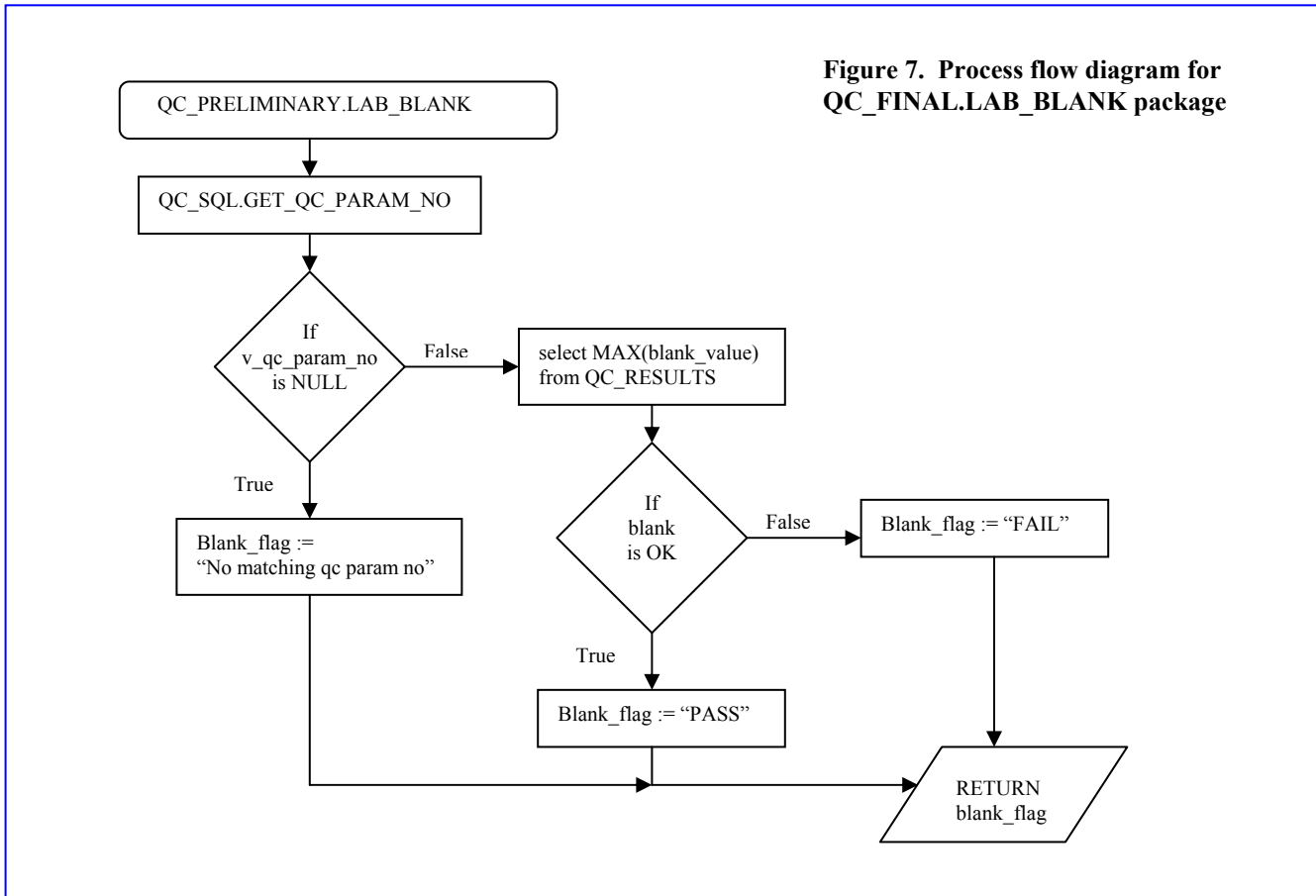
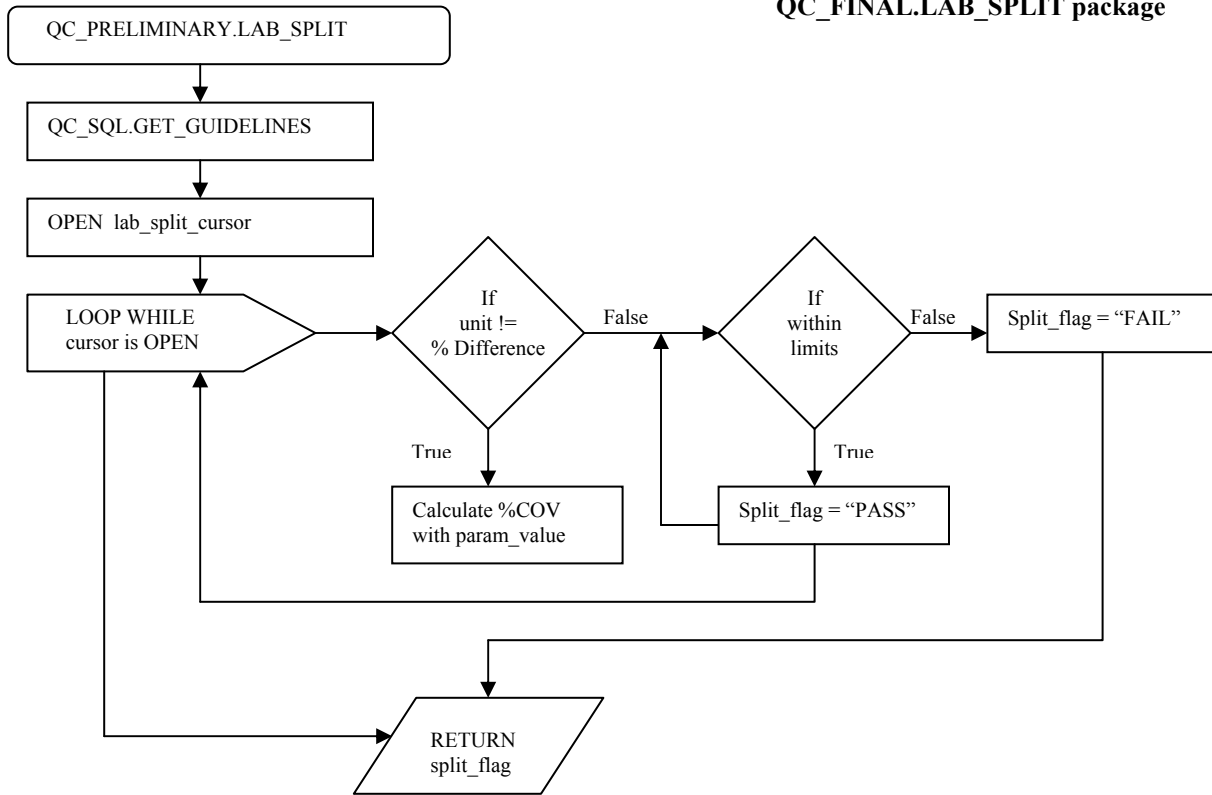
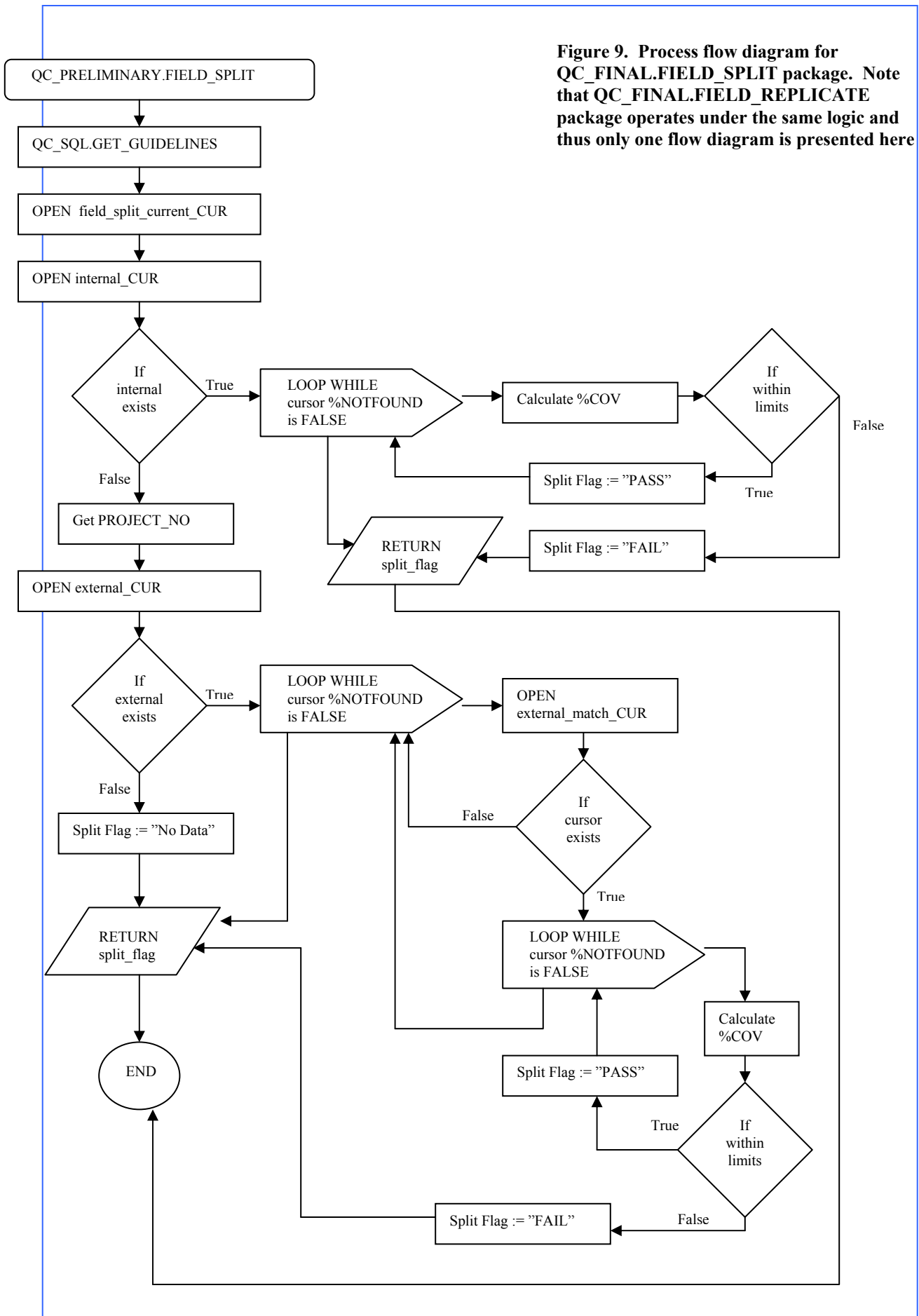


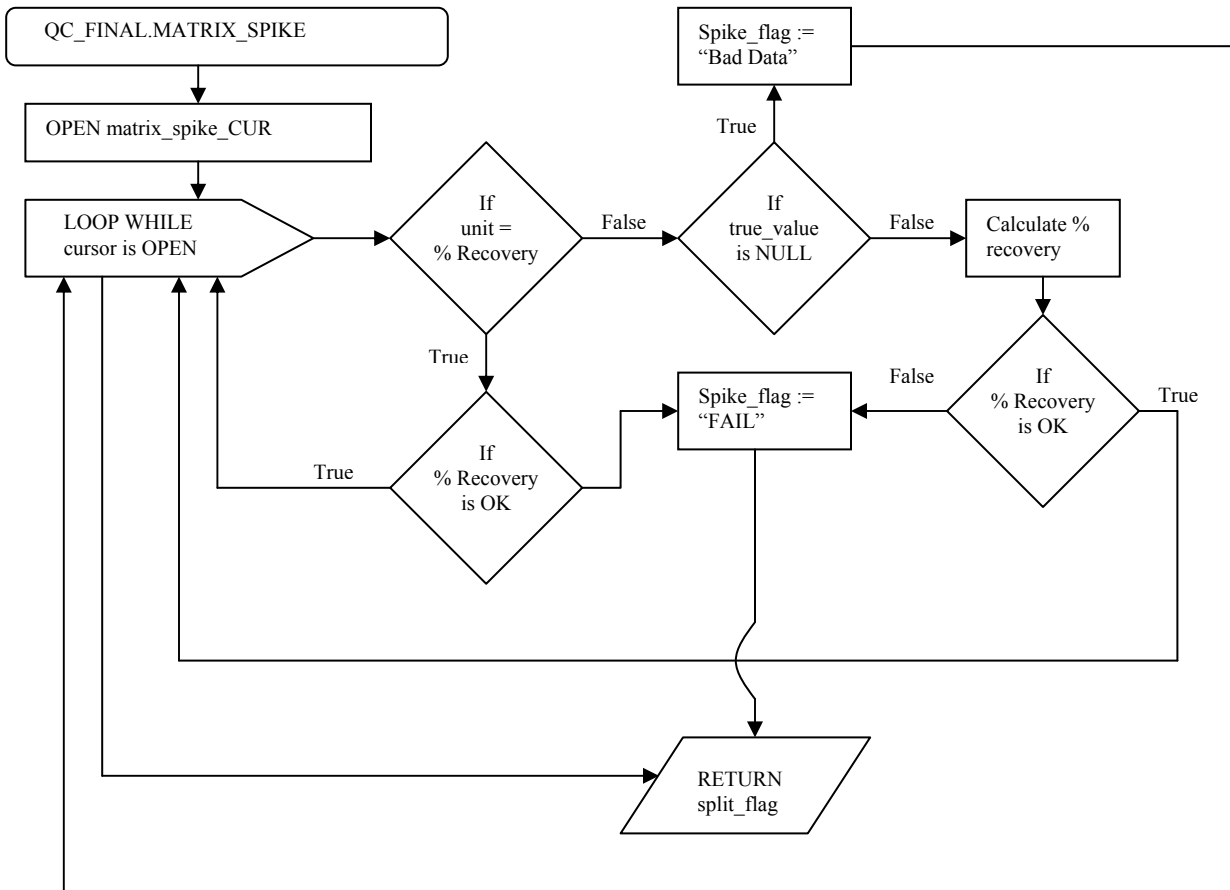
Figure 8. Process flow diagram for QC\_FINAL.LAB\_SPLIT package



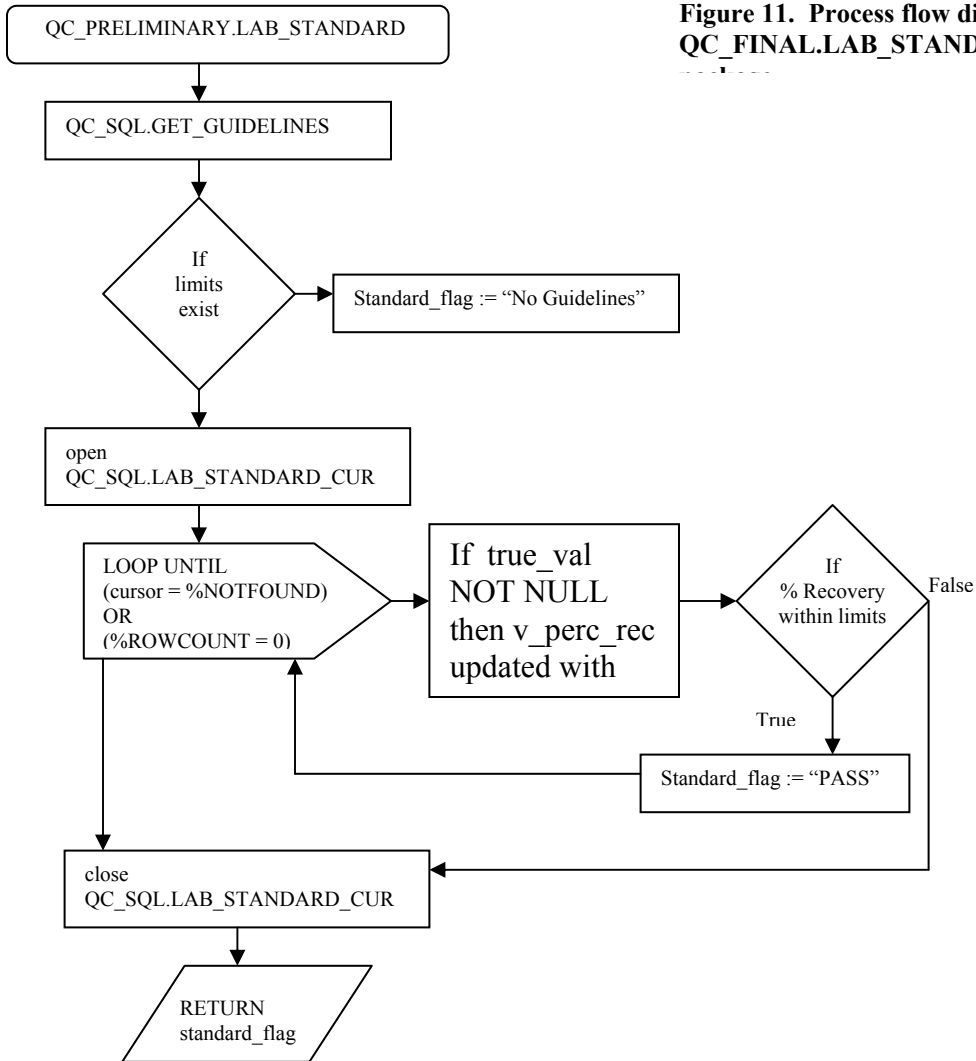
**Figure 9. Process flow diagram for QC\_FINAL.FIELD\_SPLIT package. Note that QC\_FINAL.FIELD\_REPLICATE package operates under the same logic and thus only one flow diagram is presented here**



**Figure 10. Process flow diagram for QC\_FINAL.MATRIX\_SPIKE package**



**Figure 11. Process flow diagram for QC\_FINAL.LAB\_STANDARD**



## Calculation of Flag Calculation

The full and final qualification of the data depends on the results of each of the preliminary and final tests performed, and is handled by the QC\_FLAG package which generates the single-character final flag that may be applied to the data to provide data users with a rapid assessment of the potential usability of the data. The final flag is stored with the sample result in the QC\_FLAG column of the SAMPLE\_RESULT table.

The final flag is determined by a function named FLAG\_CONDITION located within the QC\_FLAG package. The logic contained within that function is briefly described below (Table 3).

**Table 3. Description of final flag calculation logic.**

<b>Condition</b>	<b>Final Flag</b>
Failure of one or more major test (Logical Test, Lab Standard, Lab Split)	R (Rejection)
Failure of holding time (but no major test failure)	H (Holding Time Violated)
Outside of standard range (but no major test or holding time failure)	S (Outside Standard Range)
Qualification of result as an estimated value for any major test OR qualification of result as an estimated value for any minor test (but no failure of major tests) OR failure of any minor test (but no failure of major tests) OR failure of the blank test	J (Results Considered Estimate)
No guidelines specified for any QC test	G (No guidelines)
All tests with guidelines specified were within acceptable limits	U (Usable)

## Suggestions for Additional Modifications

Described below are suggestions for additional improvements that could be made to the existing Data Evaluation Process pending future staff availability and need.

- Holding times are currently calculated in days as most laboratories do not report the time of analysis, only the analysis day. However, new methods for samples such as bacteria require holding times that must be measured in hours. If contract labs could report to Field Sampling Database staff the actual time at which the sample was analyzed, the process could be modified to calculate holding times in hours.
- Groundwater staff have expressed an interest in including ion balance calculations in the Data Evaluation Process. Including the balance is not programmatically difficult assuming all specified member ions were analyzed for every groundwater sample,

although inclusion of ion balances into the determination of the final flag may be difficult. How should a bad ion balance affect the usability of the data?

- Post-calibration must be added to the process as soon as a table structure is created and post-calibration data is entered. Failure of post-calibration would result in total rejection of data.