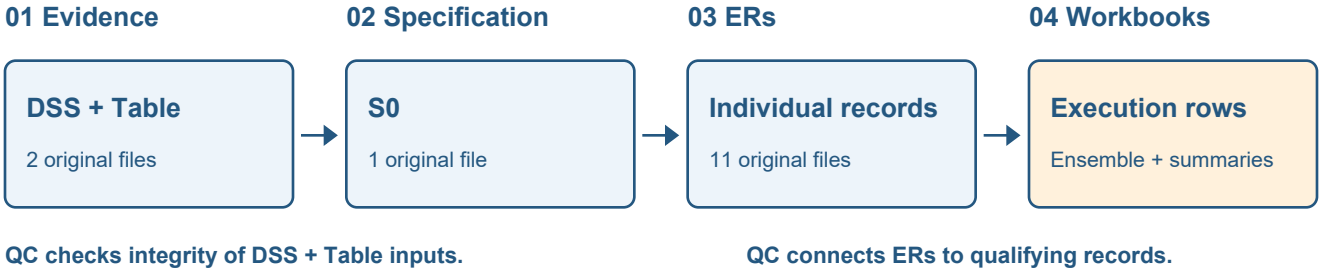
P4A: End-to-End Experimental Record and QC

DJE | Working reader edition | October 29, 2026

Experimental cell: OPVAL-P4A-S0-T

Subject: P4A, 144 Crosscove Circle, Ponte Vedra Beach, Florida

This guide follows one retrospective sale price prediction from the evidence and specification supplied to the Decision Function (DF), through its individual Experimental Records (ERs), to QC dispositions and the workbook records used to construct and verify the qualifying ensemble. The artifacts themselves are the central record; this guide explains how to follow their connections.



OPVAL is the retrospective prediction setting: the subject's realized sale price is withheld from the DF, and eligible comparable sales do not close later than the subject. S0 is the governing specification; T identifies the structured-table condition. The diagram is a reading sequence, not a claim that the evidence and specification were delivered as separate execution stages.

Artifact Inventory and Reader Route

The case provides two evidence artifacts, the S0 specification, and eleven ERs, followed by the shared workbook records. The public-report DSS is a reviewer copy; the structured table, specification, and ERs preserve the historical records.

Reading stage	Record to inspect
01 Evidence	Public-report DSS and original structured quantitative table
02 Specification	Retained S0 experimental specification; derived PDF for reading
03 Experimental Records	E001-E011, including excluded E005 and replacement E011
04 Workbook Records	Registry execution history; qualifying execution rows; P4A cell and population summaries

Use the P4A worked-example landing page to select individual artifacts and the Workbooks menu to download the two workbooks. Follow the shared Record ID and Cell ID from each execution into its disposition, qualifying population, and aggregated results.

01 Evidence: DSS and Structured Table

The **Decision Surface Snapshot (DSS)** available for inspection is a public-report reviewer copy containing P4A and the same eleven comparable sales, in the original order and 124-page sequence. P4A's outcome is redacted; the photographs match the original in content and order, and the structured table's page references remain valid.

The **Structured Quantitative Evidence Table** is the unchanged historical table supplied to the DF alongside the original, outcome-redacted private-report DSS. It organizes selected quantities from that evidence, including some fields absent from the public reports. It is another representation of the same evidence, not an additional set of market observations.

Front-End Preparation and QC Review

ChatGPT constructed and reviewed the structured table. The investigator visually compared it with the independently generated MLS source table to check consistency and accuracy. These checks were part of a human-led experiment; a deployed autonomous process would require engineered and assessed controls rather than dependence on the investigator's visual review. The QC companion develops that distinction.

02 Specification: The Authorized Task

The **S0 experimental specification** defines the task, authorized evidence use, required outputs, and execution constraints. Read it alongside the evidence to understand what the DF was instructed to do with the supplied material. The DF Engineering companion explains the distinction between the supplied evidence and the DF configuration, including its governing specification and computational method.

03 Experimental Records: Executions, QC, and Replacement

Eleven individual ERs document the DF's outputs, reported reasoning, and qualifications. Ten executions were admitted to the qualifying population. The original fifth execution was excluded; the eleventh execution replaced it.

Execution	Historical disposition and role
E001	Qualifying execution
E002	Qualifying execution
E003	Qualifying execution
E004	Qualifying execution
E005	Excluded; retained for inspection
E006	Qualifying execution
E007	Qualifying execution
E008	Qualifying execution
E009	Qualifying execution
E010	Qualifying execution
E011	Qualifying replacement for E005

Each full Record ID begins **OPVAL-P4A-S0-T-**, followed by its execution identifier. The E005 filename explicitly identifies its exclusion; the E011 filename identifies the execution it replaces. These identifiers connect the retained ERs to the workbook rows.

Inspect the Exclusion Judgment

E005 accepted the reported 1.5-car garage capacity and used it in comparative reasoning. ChatGPT challenged that interpretation during post-execution review and proposed exclusion. The investigator reviewed and agreed with the proposed exclusion and its rationale, and E005 was excluded. The retained record therefore shows an AI-proposed QC disposition accepted through human review, rather than an autonomous exclusion. The recorded rationale treated it as inconsistent with a two-car interpretation, but the original experimental DSS explicitly reports 1.5 garage spaces. The available record does not establish that the exclusion rationale was factually correct, or that the challenge was based on visual comparison of garage photographs.

The guide preserves that historical disposition for scrutiny rather than resolving the interpretive dispute. Collaborative use permits a human to challenge an interpretation before proceeding. Under autonomous EA, that same human review and verification of each proposed exclusion and its rationale would not be available as a routine step. Any pre-release intervention would need to be provided by engineered and assessed automated controls; later human review would be retrospective. The QC companion explains these control points in more detail.

Retaining both E005 and E011 makes the lineage visible. The excluded value remains part of the historical record, but does not enter the qualifying ensemble.

04 Workbook Records: From Executions to Ensemble and Summary

The two workbooks continue the record beyond the 14 original artifacts. They distinguish **an execution row**, **a property-level experimental cell**, and **a summary across properties**. These are different levels of the evidence, not interchangeable records.

Execution Rows: Trace Each ER and Its Disposition

In the **Registry Records workbook**, select the **S0 Registry Executions worksheet/tab** and find Cell ID **OPVAL-P4A-S0-T**. This records execution identity, extracted quantities, QC disposition, and replacement lineage. It preserves the excluded E005 alongside the qualifying executions and replacement E011.

In the **Qualified Records workbook**, select the **S0 Qualified Executions worksheet/tab** and find the same Cell ID. Its ten execution rows are the admitted population. E005 is absent from that qualifying set; E011 is included in the replacement slot. Each row in this execution worksheet represents selected information from one qualifying ER. It is not a property-level ensemble summary or a substitute for the full ER's reasoning and qualifications.

The P4A Cell: Aggregate the Ten Qualifying Outputs

Switch to the **S0 Qualified Verification worksheet/tab** in that same workbook and find **OPVAL-P4A-S0-T** in the OPVAL Table detail. This single row aggregates the ten qualifying executions for P4A. The workbook column **Mean F** contains the ensemble prediction, denoted $\bar{F}$. **SD(F)** describes dispersion among the individual qualifying F values; **CV(F)** is that standard deviation divided by $\bar{F}$. The comparison with realized sale price Y supplies the ensemble prediction's absolute percentage error.

Measure	Result
Mean intrinsic estimate (I)	$756,500
Ensemble prediction ($\bar{F}$)	$748,750
Standard deviation of F	$6,373.77
Coefficient of variation of F	0.8513%
Realized sale price (Y)	$749,900
Absolute percentage error of ensemble prediction ($\bar{F}$)	0.1534%

These case values are preserved from the existing guide and correspond to the P4A ensemble summary. They do not represent eleven-run statistics: the excluded execution is outside the qualifying calculation.

Population Summaries: Combine Property-Level Results

The OPVAL Table headline at the top of the **S0 Qualified Verification worksheet/tab** summarizes nine property-level cells, P1A through P9A, including P4A. Its mean APE and mean CV summarize the corresponding property-level measures. They are not statistics from treating all individual executions as independent property predictions. PROSP is shown separately, while LOOCV0 and LOOCV1 characterize their respective evidence states.

In the Registry Records workbook, the **S0 Registry Verification worksheet/tab** summarizes qualifying and excluded attempts, including the broader OPVAL Table/No Table exclusion comparison. E005 contributes to the recorded Table exclusions there; its individual disposition and replacement lineage are inspected in the execution tab, not this summary tab. The workbook pair therefore supports both performance verification and inspection of the execution history that produced the qualifying populations.

Review the Translation into Workbook Records

The trace runs from ER identity and extracted I/F values, through QC disposition and ensemble membership, to property-level aggregation and population summaries. A reviewer can inspect each level using the shared identifiers. Checking the formulas alone does not establish faithful extraction from the ERs, and correct extraction does not establish that every underlying execution or QC judgment was valid.

Workbook construction and reconciliation in this project were collaborative engineering activities. The QC companion discusses the additional controls that would need development and assessment for autonomous record handling.

Scope and Limitations

MLS policy does not permit public distribution of the private-report DSS supplied to the DF. The public-report reviewer copy omits private listing-agent remarks and certain listing, transaction, showing, and agent fields; its layout and preparation date also differ. It preserves the property population, photographs, and page references, but is not the original experimental input.

This case makes an actual experimental sequence inspectable. It is not a representative sample of performance or QC failure mechanisms, and favorable P4A accuracy does not establish the adequacy of the DF, every QC disposition, or authority for autonomous execution.

Execution validity, empirical performance, and the sufficiency of engineering and operating controls to justify Standing are separate questions. A candidate-specific disputed judgment need not revoke Standing for other candidates; a broader weakness can require engineering reassessment and reconsideration of continuing Interpretive Sufficiency. This historical case does not establish that an autonomous deployment was operating.

The unretained printed source tables used for the investigator's visual check are not reconstructed as original exhibits.

The guide preserves historical records and reported case values. It does not establish an automated end-to-end workbook-reconciliation control. The QC and DF Engineering companions develop those engineering distinctions and their implications.