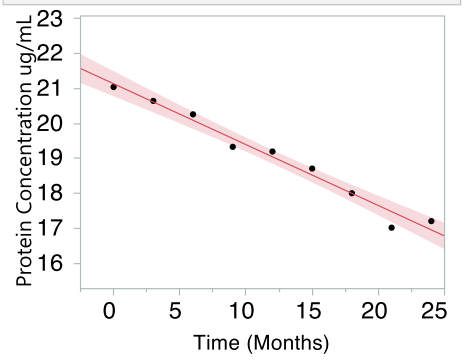
Out of Trend Identification and Removal in Stability Modelling and Regression Analysis

This paper defines the concept, justification and method of removal of out-of-trend points in stability modelling and shelf life prediction. It provides the overview from literature, guidance and best analytical practice and the method to be used during stability modelling and analysis. A clearly defined OOT procedure is needed to correctly and consistently remove outliers from expiry and stability modelling and prediction where technically warranted.

Out of Trend Definition and Application

Out of trend (OOT) is a point (measurement) in a regression analysis that has statistically greater error at a defined risk factor from a regression line or multiple factor regression model than other determinations. A time dependent result which falls outside a predicted statistical interval. Simply put an OOT event is an outlier in a regression analysis. OOT points are considered to be non-representative of the test sample and due to analytical, transcription or other sources of error. Failure to remove OOT point(s) if they exist will produce calculated rates of change that will not be representative of the drug product nor drug substance. The following are indications that OOTs are present in the stability analysis; 1) points do not line up on the regression line, 2) confidence intervals of the fit are excessively wide, 3) root mean squared error (RMSE) of the residuals has excessively expanded well beyond the characterized analytical error, 4) expiry from one time point to another has a large amount of difference and 5) R^2 has a large amount of change with and without the OOT time point.

No OOT



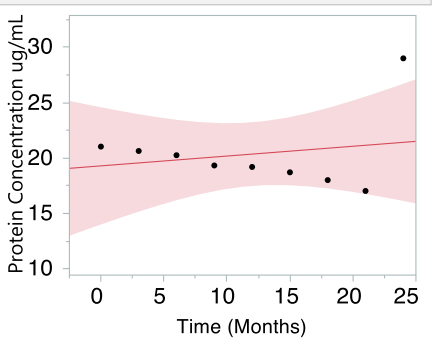
Linear Fit

Linear Fit

Protein Concentration ug/mL = 21.129718 - 0.1733043*Time (Months)

Summary of Fit	
RSquare	0.974613
RSquare Adj	0.970987
Root Mean Square Error	0.245665
Mean of Response	19.05007
Observations (or Sum Wgts)	9

OOT



Linear Fit

Linear Fit

Protein Concentration ug/mL = 19.297252 + 0.0884766*Time (Months)

Summary of Fit	
RSquare	0.043636
RSquare Adj	-0.09299
Root Mean Square Error	3.63802
Mean of Response	20.35897
Observations (or Sum Wgts)	9

Figure 1.0 OOT Illustration and Influence of OOT

"OOT stability data can be described as a result or sequence of results that are within specification limits but are unexpected, given the typical analytical and sampling variation and a measured characteristic's normal change over time (e.g., an increase in degradation product on stability)."1

Regression analysis is normally used to determine change over time and associated 95% confidence limits relative to rates of change and expiry. ICH Q1A(R2) Stability Testing Of New Drug Substances and Product states:

> *"The nature of any degradation relationship will determine whether the data should be transformed for linear regression analysis. Usually the relationship can be represented by a linear, quadratic, or cubic function on an arithmetic or logarithmic scale. Statistical methods should be employed to test the goodness of fit of the data on all batches and combined batches (where appropriate) to the assumed degradation line or curve."*

OOT evaluation and elimination should be used for the following applications and prediction; 1) shelf-life estimation, 2) storage evaluation, 3) impurity formation and trending, 4) in-process monitoring and prediction and 5) tracking and trending of lot performance.

Likely Root Causes for OOT

The following are typical possible sources and mechanisms for OOT events that may occur during stability evaluation:

- Sample selection and sample handling errors
- Dilution and sample preparation errors
- Sample materials and plate errors
- Temperature, reaction time and pH effects errors
- Vendor and lot variation on critical reagents
- Flow rates and process time errors
- Analyst errors
- Instrument variation and calibration error

- Nonstandard test procedures and not following the method SOP
- Drift in standards or reference materials
- Stability of test samples or critical reagents
- Calibration or compensation errors
- Interaction and composite errors
- Expiry of bulk materials

Historical Approaches to OOT

The following are typical historical approaches to OOT identification and removal. They are not considered to be statistically sound methods for OOT identification and removal.

- The difference between consecutive results is outside of half the difference between the prior result and the specification
- The result is outside ± 5% of initial result
- The result is outside ± 3% of previous result
- The result is outside ± 5% of the mean of all previous results

There is no statistical basis for the above definitions of OOT and does not take into account process nor method variation. These are not recommended approaches.

Closed Loop Approach to OOT Identification and Removal

A best practice approach to OOT determination and removal is to see it as a part of a closed loop control system during stability monitoring and expiry prediction. The five steps to a closed loop system for OOT are 1) addition of new time points and data, 2) OOT identification, 3) OOT determination and point removal where warranted, 4) OOT verification and evaluation of OOT influence and 5) stability and performance prediction.

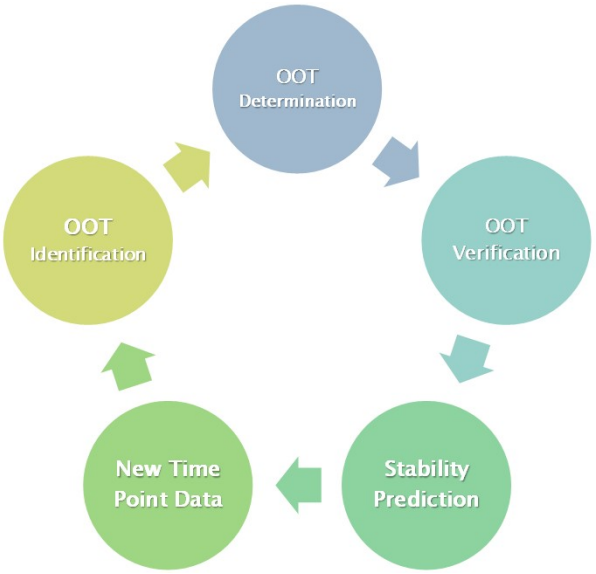


Figure 2.0 Closed Loop OOT Identification and Resolution

Adding New Time Points

As each new time point is added to the stability analysis they should be checked for OOT potential. If they are within the criteria for OOT identification then rates of change, expiry etc. are determined. OOT identification, determination and verification are used if new time points appear to be suspect.

OOT Identification

There are four methods to identify a point as OOT, 1) visually, 2) outlier boxplot of the residuals, 3) multivariate Jackknife Distances and 4) control chart of the residuals. Jackknife distances are the most sensitive in identification of OOT points in a regression analysis as they include and remove each time point in the analysis to evaluate their influence in the model. Once you believe you have identified an OOT the next step is to test it to determine if the OOT will be removed. Root cause as to why the point is OOT is a secondary investigation once the point has been determined to be OOT. Once the point is determined to be a possible OOT the point is tested statistically to determine if it is or is not OOT.

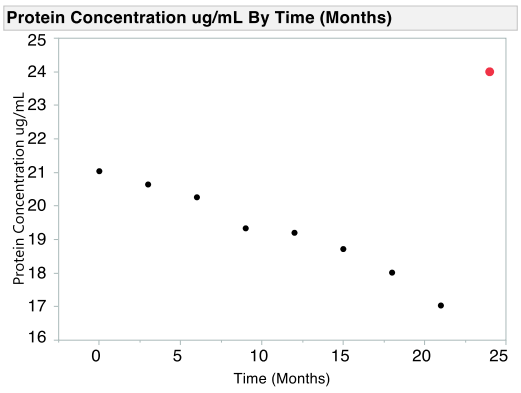


Figure 3.0 Visual Analysis

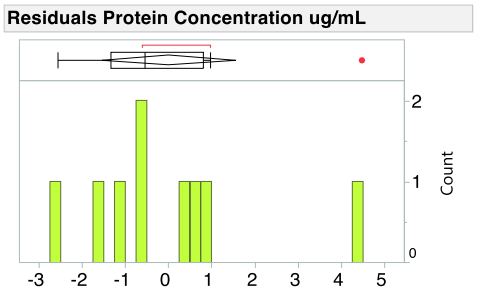


Figure 4.0 Outlier Box Plot of Residuals

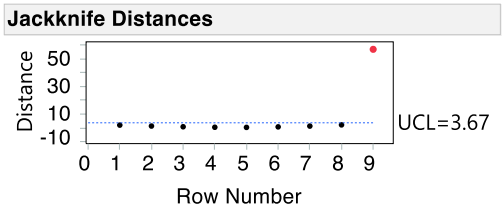


Figure 5.0 Multivariate Jackknife Distances

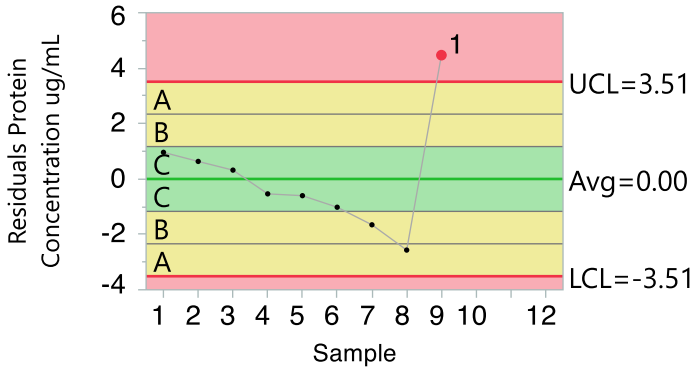


Figure 6.0 Control Chart of Residuals

OOT Determination

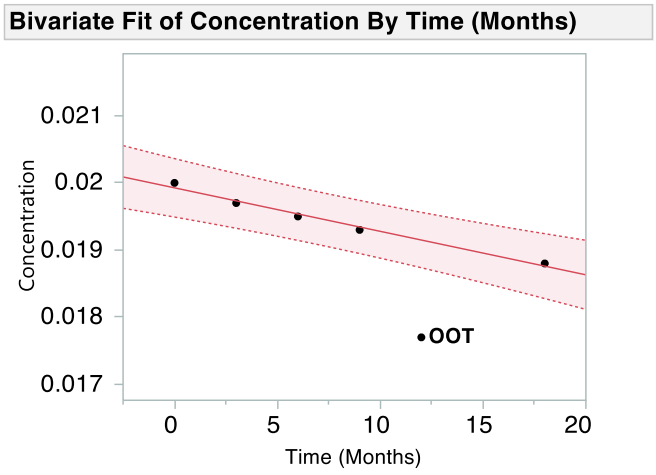
The following is the recommended procedure for OOT outlier determination. Exclude and hide the suspected OOT point in the data analysis. Fit a linear regression line with the potential OOT time point excluded. Save the predicted response (concentration) from the linear fit. Calculate the difference (Delta) at each time point. Calculate a z score for each time point.

z score= (Measurement-Predicted)/stdev(residuals with point removed)

Once the *z* score for the OOT has been calculated it can be compared to a risk threshold. A k-sigma of 2.576 or 99% risk is used to set the limit for OOT detection. Abs(*z*)> 2.576 (99%) is OOT so in this example z is -27.238 so it is OOT. A z score with a limit is the best method of OOT detection. The key difference with this procedure and other z score procedures written in the literature is the z score is evaluated with the point removed. This correctly scales the residual error so that the influence of the OOT point is not included in the residual error and the OOT time point can be correctly evaluated based on the other measurements error.

	Time (Months)	Concentration	Predicted Concentrati...	Delta	z Score	OOT
1	0	0.02	0.0199264151	0.000074	1.3832	OK
2	3	0.0197	0.0197320755	-0.000032	-0.6029	OK
3	6	0.0195	0.0195377358	-0.000038	-0.7093	OK
4	9	0.0193	0.0193433962	-0.000043	-0.8157	OK
5	12	0.0177	0.0191490566	•	-27.2382	OOT
6	18	0.0188	0.0187603774	0.000040	0.7448	OK

Figure 7.0 OOT Determination



Linear Fit

Concentration = 0.0199264 - 6.478e-5*Time (Months)

Summary of Fit

RSquare	0.986058
RSquare Adj	0.981411
Root Mean Square Error	6.143e-5
Mean of Response	0.01946
Observations (or Sum Wgts)	5

Figure 8.0 Stability Analysis with OOT Removed

OOT Verification

To verify the influence of the OOT the following measurements are recommended; 1) change in R^2, 2) change in RMSE and 3) change in expiry calculation. Comparing with or without the OOT time point will verify the influence of the time point and confirm the need for removal. If the change in the three identified measures are trivial then the OOT has not been verified and its removal is not warranted. Differences in R^2, RMSE or Expiry of 3% or less are generally not practically important to drug substance or drug product expiry or stability evaluation. Verification is performed by including then removing the OOT point in the stability evaluation and then measuring the change in the key performance metrics of the fit and the prediction. Also RMSE error can be compared to the repeatability of the analytical method to determine if the residual error is primarily analytical measurement error.

A control chart of the residuals with the OOT time point excluded will be a secondary confirmation of OOT identification and removal as an outlier.

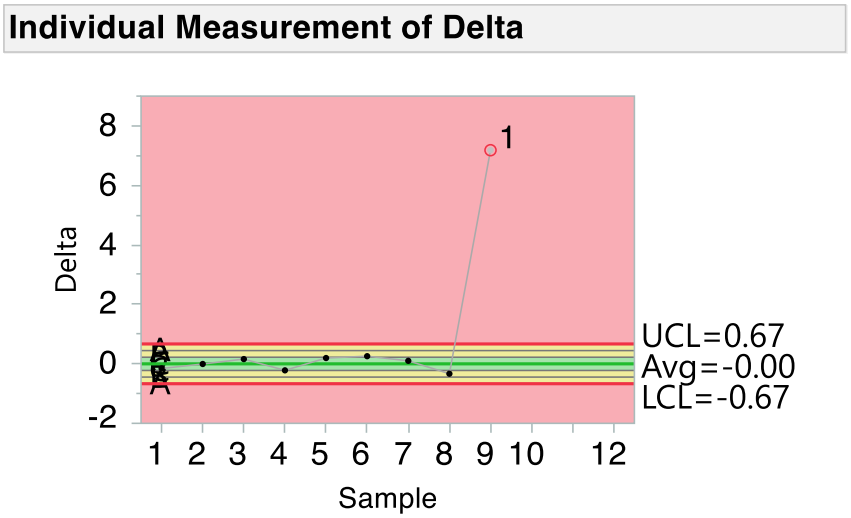


Figure 9.0 Control Chart of Residuals

Stability Prediction

Once the OOT point has been removed and verified, stability prediction can then be performed per ICH Q1E. OOT points should always be included in the plot to indicate the measurement but tagged as OOT to indicate the point was included in graph but not included in the analysis. Including it in the plots will provide full disclosure as to all observations at all time points. Once a point has been determined to be OOT it does not renter the analysis as additional data is added to the stability prediction.

Single Factor, Single Batch and Multiple Factor/Multiple Batch OOT

The technique for outlier works equally well for single batch single factor versus multiple batch and multiple factor stability modelling and expiry prediction. The example provided

was for a single batch, single factor analysis. For the multiple factor and or multiple batch condition the same approach would be used. The model would be fit with the OOT removed and then the model would be saved, z scores would be calculated for all time points and then OOT would be determined based on a 2.576 (99%) threshold.

Conclusions

OOT Determination is a powerful addition to any stability program and needs a clearly defined SOP to consistently apply the logic to day-to-day stability evaluation. Including an OOT protocol to stability testing and data analysis will produce more statistically reliable stability determination and expiry prediction. OOT determination based on the protocol described in this paper opens the door for the automation of OOT determination and removal from any stability analysis and may be the best approach to systematically evaluate all time points in stability analysis.

References

ICH Q1A(R2) Stability Testing Of New Drug Substances, Feb 2003

ICH Q1E Evaluation for Stability Data Feb, 2003

Pharmaceutical Technology (2003) Trajkovic-Joleska, Torbovska. Method for Identification of Out-of-Trend Stability Results.

Pharmaceutical Technology (2003). A Review of the Potential Regulatory Issue and Various Approaches. PhRMA CMC Statistics and Stability Expert Teams.

Pharmaceutical Technology (2005). Identification of Out-of-Trend Stability Results, Part II PhRMA CMC Statistics, Stability Expert Teams.

THOMAS A. LITTLE CONSULTING TLC

Thomas A. Little, PhD

12401 North Wildflower Lane

Highland, UT 84003 USA

drlittle@dr-tom.com

925-285-1847