A Quality-by-Design Methodology for Rapid HPLC Column and Solvent Selection

Introduction

High Performance Liquid Chromatography (HPLC) method development work typically involves column screening followed by formal method development. Column screening, or scouting, is the activity of selecting the correct analytical column. Formal method development identifies the important instrument parameter settings that best separate all compounds of interest.

Traditionally column screening is done at constant conditions of all other instrument parameters which are studied later as part of formal method development. Such as sequential approach results in a minimal design space that does not allow for the expression of interaction effects which can greatly affect the selectivity of individual columns for one or more critical compound pairs. To address this limitation some column screening approaches use a multi-factor statistical experimental design that includes multiple combinations of solvents and pH to provide a more complete design space. However, in such a dynamical system the chromatographic results will often be severely limited due to peak exchange and peak co-elution, which puts final column selection on a purely qualitative footing.

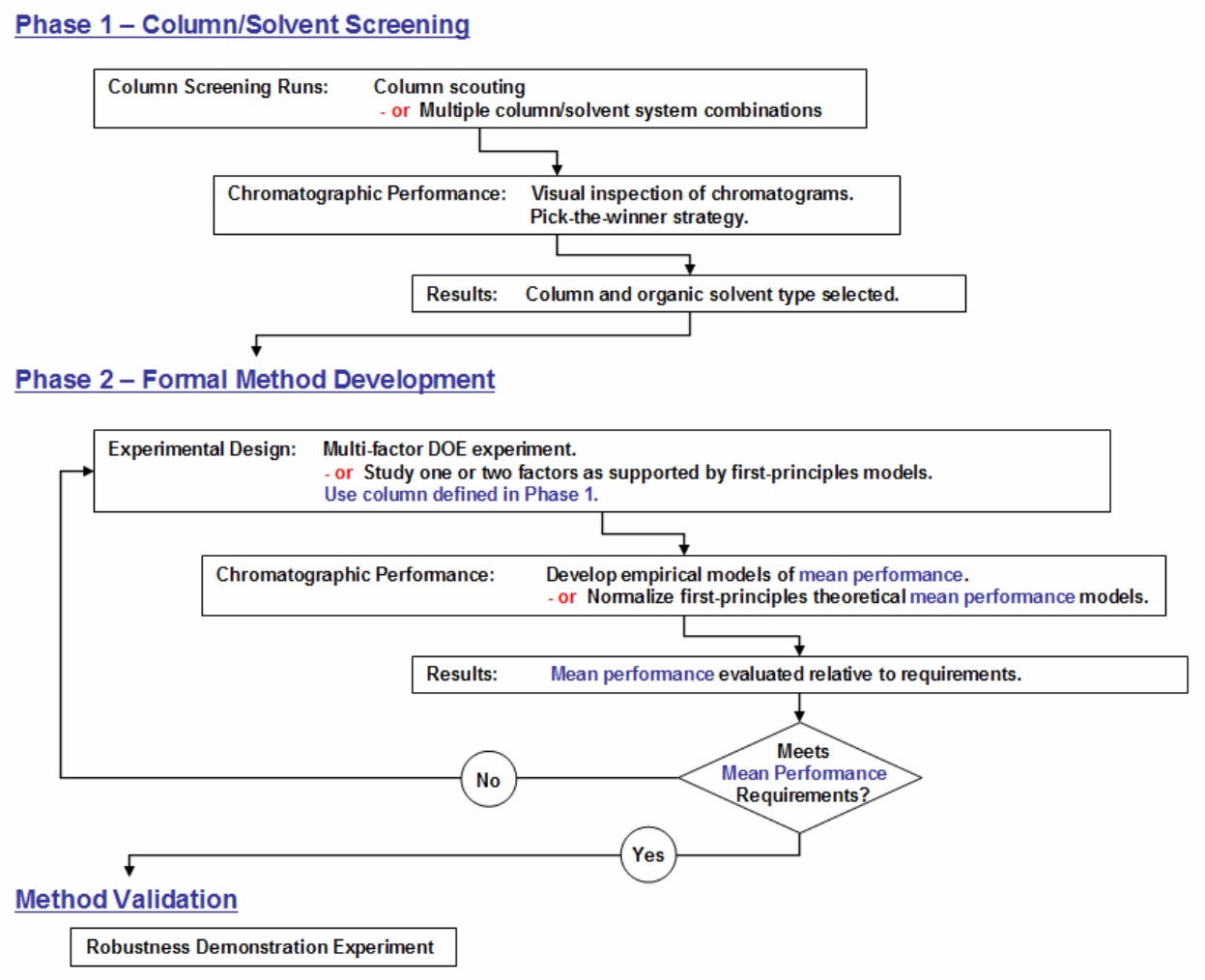
This white paper describes a new methodology for automated HPLC column and solvent system selection using QbD principles. The methodology, adapted to multiple instruments and instrument data systems, overcomes the limitations inherent in both the sequential and classical Design of Experiments (DOE) approaches to place the column screening activity on a rigorous and quantitative footing.

Current HPLC Method Development Practice

Reversed-phase HPLC is by far the most widely used HPLC separation methodology in pharmaceutical and biotechnology analytical applications. Reversed-phase HPLC is therefore the basis of the discussions and examples used in this paper. However, the reader will recognize that the instrumentation, software, and Quality-by-Design (QbD) based methodologies presented here are applicable to other HPLC approaches such as normal-phase and HILIC.

Figure 1 presents a diagram of the HPLC method development workflow as it is commonly practiced today. As the diagram indicates, the first phase is selecting the analytical column. Once the "best" column is identified, a second development phase is carried out that addresses the remaining important instrument parameters. The goal of this second phase is to identify the parameter settings that meet all critical method performance criteria in terms of both compound separation and total assay time. The third phase is to conduct an experiment to demonstrate the robustness of the resulting method. This is normally done as part of the method validation effort.

Figure 1. Current Method Development Workflow

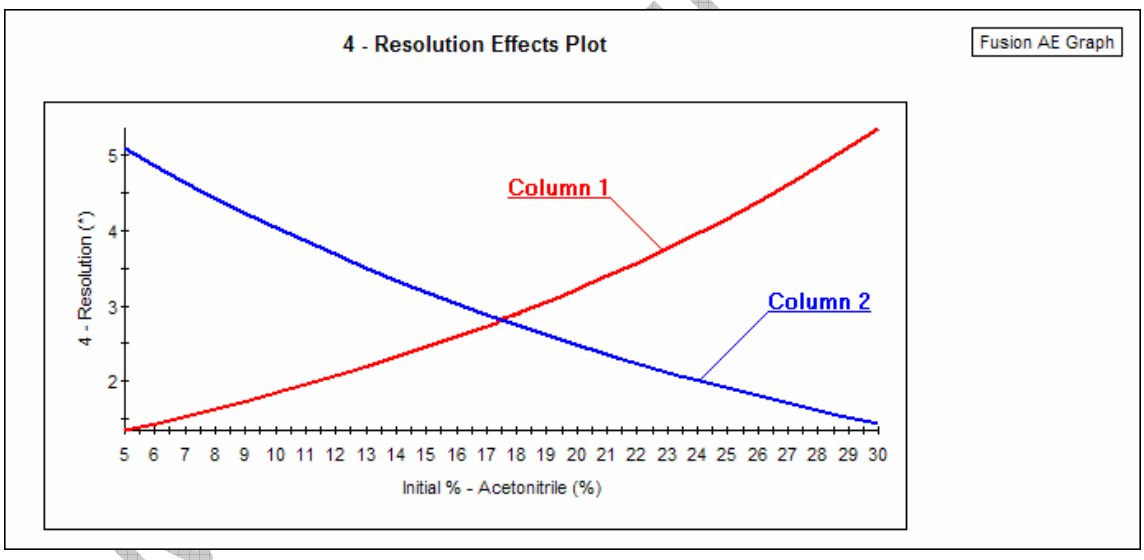


Phase 1 - Column/Solvent Screening

A column screening approach that focus exclusively on column selection will involve one experiment trial for each candidate column carried out at a single fixed level of all other instrument parameters. This approach is therefore very limited in terms of the overall design space that it addresses. A major risk in this approach is that an analytical column's selectivity often depends on the settings of other parameters such as pH, type of organic solvent, and even gradient conditions. These interactions mean that a given column's selectivity (ability to separate one or more critical compound pairs) can range widely depending on the specific pH and gradient conditions used in the column selection protocol.

Figure 2 illustrates the effect of an interaction between column type and gradient slope on the resolution of a critical compound pair (in this case compounds 3 and 4). The interaction was expressed in the data from a 4-column screening experiment in which the gradient slope was varied by changing the initial % of the organic solvent (X axis) at constant gradient time. The interaction graph expresses two important aspects of the relative selectivity of the two columns for this compound pair. First, the selectivity of each column is shown to strongly depend on the gradient slope used. Second, the nature of the dependency is not the same for all columns, or even all qualitatively similar columns.

Figure 2. Column–Gradient Slope Interaction



As the graph in Figure 2 shows, which column is the better choice for resolving the critical pair depends completely on the initial % of organic solvent (steepness of the gradient slope). A one-factor approach does not allow interaction effects such as the one illustrated in Figure 2 to be expressed in the experiment data, and so the effects of interactions on method performance can not be identified or quantified. This demonstrates the need to visualize and characterize these interactions in the column selection process.

Alternatively, QbD principles can be applied to the task of screening analytical columns to include factors such as pH, gradient slope, and organic solvent type, as these factors are recognized effectors of column selectivity. In a QbD approach a multi-factor statistical experiment design plan would be used to systematically vary the multiple study factors in a series of experiment trials that together comprehensively explore the design space. The statistical experiment design would provide a data set from which the interaction effects of the factors could be identified and quantified along with their linear additive effects and curvilinear effects.

However, in practice column screening experiments, even those done using a Design of Experiments (DOE) approach, often have significant inherent data loss in critical results such as compound resolution. The data loss is due to both compound co-elution and also changes in compound elution order between experiment trials (peak exchange). These changes are due to the major effects that pH, organic solvent type, and gradient conditions can have on column selectivity. Switching columns between trials while simultaneously adjusting these factors therefore dramatically affects compound elution and therefore the complete3ness of the resolution data computed from the experiment trial chromatograms.

As is discussed in detail below, the inherent data loss in current practice column screening experiments makes accurate peak identification across the experiment trial chromatograms (peak tracking) extremely difficult. This in turn makes numerical analysis of the results very problematic. Often it reduces data analysis to a manual exercise of viewing the experiment chromatograms and picking the one that looks the best in terms of overall chromatographic quality – a "pick-the-winner" strategy.

Inherent Data Loss in the Current Approach

The inherent data loss due to co-elution and peak exchange in column screening experiments can be seen by comparing the chromatograms in Figures 3 and 4, obtained from a simple 2-column screening experiment which also included Gradient Time as a study factor. The Compound 4a resolution data value obtained from the chromatogram in Figure 3 is the measure of the separation of Compound 4.a from Compound 3 – the immediately prior eluting compound. However, as seen in Figure 4 the instrument parameter settings associated with trial 22 cause Compounds 3 and 5 to co-elute. As a result, the Compound 4a resolution data obtained from the trial 22 chromatogram is the measure of the separation of Compound 4.a from Compound 5. Additionally, no Compound 3 resolution data can be obtained from the trial 22 chromatogram.

Figure 3. Chromatogram from Trial 12

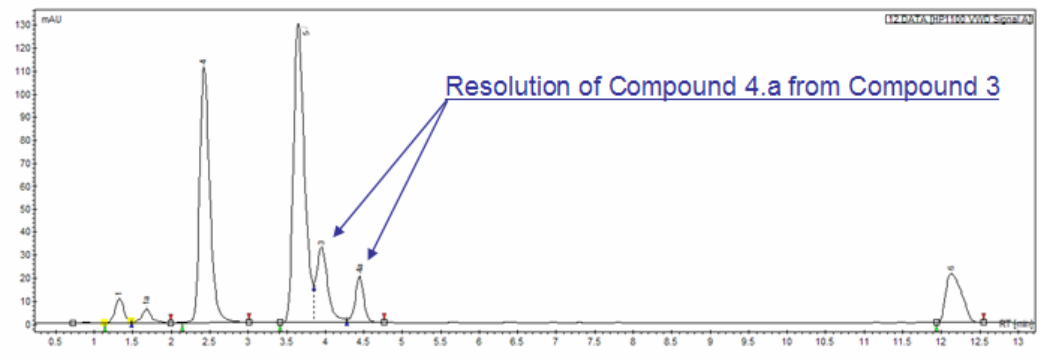


Figure 4. Chromatogram from Trial 22

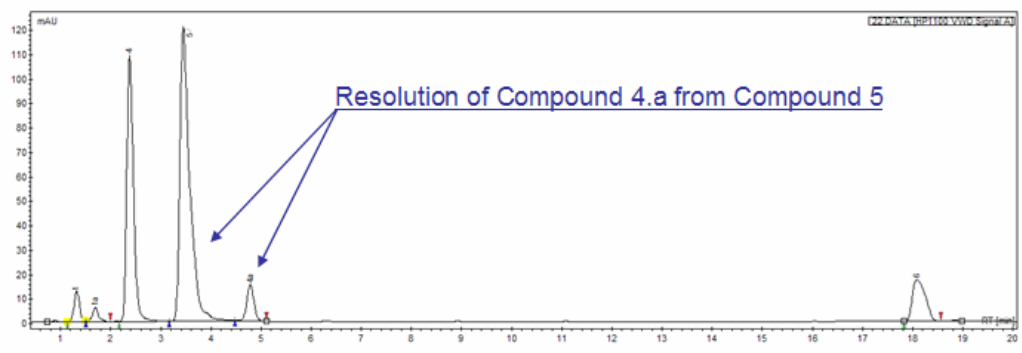


Table 1 presents the experiment design used in the 2-column screening experiment from which the chromatograms in Figures 3 and 4 were obtained, along with resolution results for five of the 12 compounds in the experiment sample. The data were generated by identifying the compound associated with each peak in each chromatogram (peak tracking). One can see the large number of missing resolution values in the Table 1 data set for Compounds 3 and 4a - two impurities that must be able to be separated from Compounds 4 and 5, the two APIs in this drug product sample.

Table 1. Example Data Set - Current Practice Data

		Study Factors				Experiment Results		
Trial No.	Initial Hold Time	Gradient Time	Final % Organic	Column Type	3 - Resolution	4 - Resolution	4a - Resolution	5 - Resolution
1	1	20	50	A	2.94	1.18		16.54
2	1	20	65	B	1.17	2.54	2.53	
3	1	25	50	A	2.94	1.13	1.66	3.42
4	4	15	80	B	1.27	3.24	2.12	
5	1	15	80	A	2.48	1.04		4.09
...	...	...	...	...	...	...	...	...
11	1	15	65	A	2.44	1.13		4.36
12	1	25	80	B	1.21	3.28	2.2	4.86
13	4	15	80	B	1.39	4.9	2.64	5.08
14	2.5	25	65	B	0.69	3.52	1.19	5.1
15	4	25	80	B	1.23	3.27	2.12	4.68
...	...	...	...	...	...	...	...	...
40	1.75	22.5	72	A		1.58		4.32
41	1	15	80	B		3.21	3.98	5.06
42	3.25	22.5	72	A		1.33		3.98
43	2.5	20	80	B	0.8	3.6	1.1	5.25
44	4	15	50	A		1.31		4.71
45	2.5	25	50	A		1.42		3.99
46	1	20	50	B	0	3.62	2.56	5.42
47	1	15	50	B		3.16	3.9	3.99
48	4	15	50	B		3.01	4.76	3.78
49	4	25	65	A		1.44		4.62
50	4	15	80	A	0	2.5		4.49

The result of inherent data loss in HPLC method development experimental work is that the data do not accurately represent a compound's actual chemistry-based behavior, and so provide no basis for legitimate analysis and interpretation of the results. This can be seen from regression analysis (equation-fitting) of the Compound 4a data, the results of which are presented in Table 2. The R^2-Adj. (read as Adjusted R Square) in Table 2 is a key measure of equation predictive accuracy. Hypothesis testing of the value of 0.0639 shows it to be ***not*** statistically different from zero, meaning that the equation has no predictive accuracy whatsoever.

Table 2. Regression Statistics – Compound 4a Resolution

Regression Statistic Name	Statistic Value
R^2 **R^2-Adj.**	0.0986 **0.0639**

However, the study parameters included Column Type (two columns with very different stationary phases) and a wide range of Final % Organic (the gradient endpoint percent of organic solvent) – two instrument parameters known to greatly affect compound separation under almost all conditions. Additionally, the observed changes in the resolution data across trials are substantially greater than can be accounted for by HPLC instrument error. Therefore, it can only be concluded that inherent data loss is the cause of the inability to derive any statistically valid results from numerical analysis of experiment data.

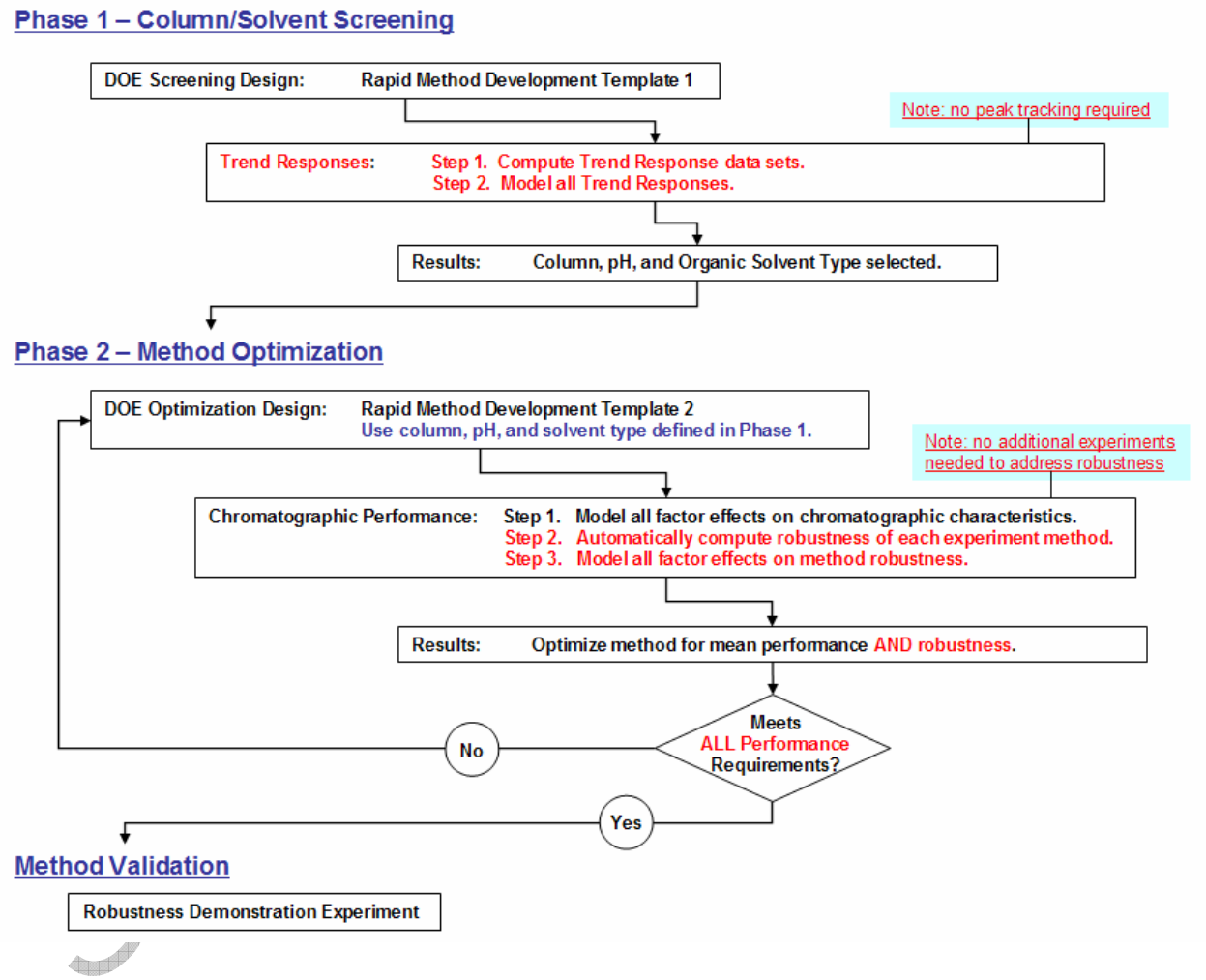
The inherent data loss problems described above are systemic to HPLC method development approaches using DOE methods as currently implemented. The reason is simple: it is standard practice to start the method development process by studying the factors known or expected to have the greatest affect on peak shape and compound retention time, and therefore peak separation. This is especially true of current "phased" method development approaches that start with column screening. But it is exactly these changes that make correct compound assignments between trial chromatograms extremely difficult. As a result, the most important information sought from the experiment – the best column and gradient conditions – is normally not available.

As described in the next section, the lack of effects visualization associated with a one-factor-at-a-time approach and the inherent data loss problems associated with current DOE approaches are solved by using DOE methods in combination with unique surrogate responses that eliminate the need for peak tracking. These responses enable replacing a qualitative pick-the-winner strategy with a quantitative practice consistent with QbD principles.

A New Quality-by-Design Based Methodology

Figure 5 is a flowchart of a new QbD-based method development workflow. The new methodology harmonizes with current practice in many labs in that it is becoming common for method development to be carried out in two phases. However, novel experimentation and data treatment methods have been integrated into each phase of the new practice to transform the qualitative elements of current approaches into statistically rigorous quantitative practice. Most importantly, these novel methods have been automated to accomplish this in minimum time and effort.

Figure 5. New Method Development Practice Workflow



This new practice has been successfully demonstrated in "live" studies carried out independently at several pharmaceutical laboratories. The studies involved either (1) test mixes of active ingredients and impurities designed to challenge the practice, or (2) current method development projects in which obtaining an acceptably performing method proved resistant to current practice attempts.

Although the flowchart presented in Figure 5 identifies the important changes to current practice that have been developed and implemented in both phases of method development, the details of the Phase 2 changes are beyond the scope of this white paper and will be presented in a subsequent white paper. The next two sections of this white paper present the theoretical background of the novel experimentation and data treatment methods in Phase 1 and one of the successful proof-of-technology experiments conducted according to the new practice.

Technical Background

The new HPLC method development practice implemented in the Fusion AE software program executes the Column/Solvent Screening phase using statistical design of experiments (DOE) methods. However, this new practice combines DOE experimentation with a novel response data computational approach that quantitatively identifies the best column, pH, and organic solvent type without any need for peak tracking (S-Matrix patents pending). The Phase 1 Column/Solvent Screening experimental work involves the following four workflow steps:

1. **Define the design space.** The new practice uses a *Phase 1 - Column/Solvent Screening* experiment template that includes the factors identified below. The template is modified in terms of the gradient time range, the need for and levels of pH, the complement of candidate columns, and the type(s) of organic solvent to be used based on the specific compounds that must be resolved.

Experiment Variable	Range or Level Settings
Gradient Time (min)	15.0 — 40.0
pH	2.5, 5.0, 7.0
Column Type	YMC Pro C4 5um Pursuit C18 3um Polaris C18 3um Pursuit DP 3um Pursuit XRS C8 3um
Gradient Slope (% Organic)	5.0 — 95.0
Organic Solvent Type	Acetonitrile, Methanol, Blend

2. **Generate a statistically designed experiment.** The experiment is a series of study factor level setting combinations (experiment trials) to be run on the instrument. The software defines the trials that will thoroughly address the design space in the minimum required number of trials. This in turn assures that all important study factor effects will be expressed in the experiment data.

3. **Run the various design conditions on the instrument.** The software automatically constructs a sequence and builds the instrument methods within the CDS.

   Once the experiment is run, the software automatically retrieves the results data for each experiment chromatogram from the CDS and computes unique *surrogate* responses from the results data. Definitions of these surrogate responses follow Step 4.

4. **Derive predictive models of the Surrogate Responses.** The software applies a pre-scripted series of linked analysis routines to the experimental data that include response data nonlinearity metrics, error analysis, linear regression analysis, and outlier analysis to develop Trend Response models (prediction equations).

Two of the surrogate responses mentioned in Step 3 above are Total Peaks and Resolved Peaks. Below are definitions for these responses.

Total Peaks: the total number of integrated peaks in a chromatogram.

- The number is normally the number of "integrated" peaks obtained from a chromatogram that has been reprocessed based on user set minimum peak height and/or minimum peak area thresholds.

Resolved Peaks*: the number of integrated peaks in a chromatogram with resolution ≥ X.

- Responses can be computed for various values of X such as 1.50, 1.75, 2.00, 2.50.
- Each value of X is settable by the user.
- The default X value for HPLC applications is normally 1.5.

* - chromatographic results such as area or area percent can be used as trend responses to address other method development goals such as API purity.

In the HPLC application these unique surrogate responses are termed "Trend Responses", since these data contain the information on the key trends in chromatographic quality as the experiment variable settings are systematically changed across the experiment trials. Since the key trends in these surrogate responses are expressions of peak shape and compound separation, the two characteristics most consequential to the ability to accurately measure compound amount in a sample, the trend responses directly support standard HPLC method development goals.

Note that, as opposed to a pick-the-winner strategy, the trend responses are statistically analyzed and modeled. These models provide quantitative estimates of the study factor effects – it is this Quantitation that defines the best column, pH, organic solvent type, and method conditions of any other included study factors.

Note also that computing the trend responses defined above does not require any assignments of peaks to sample compounds in the chromatograms providing the source data for the computation. Fusion AE can automatically compute the trend responses from integrated peak data available in several chromatography data systems (CDS). From analyses of the trend response data sets Fusion AE determines the best performing combination of the experiment study factors.

It should be understood that both trend responses are addressed since both goals must normally be met; achieving one goal does not necessarily guarantee that the other goal will also be achieved. For example, the best instrument settings for the Total Peaks response may result in peaks being present for all compounds, but only some compounds being separated to the degree required. Conversely, the best instrument settings for the Resolved Peaks response may resolve almost all compounds but leave some peaks completely unresolved.

Proof-of-Technology Experiment

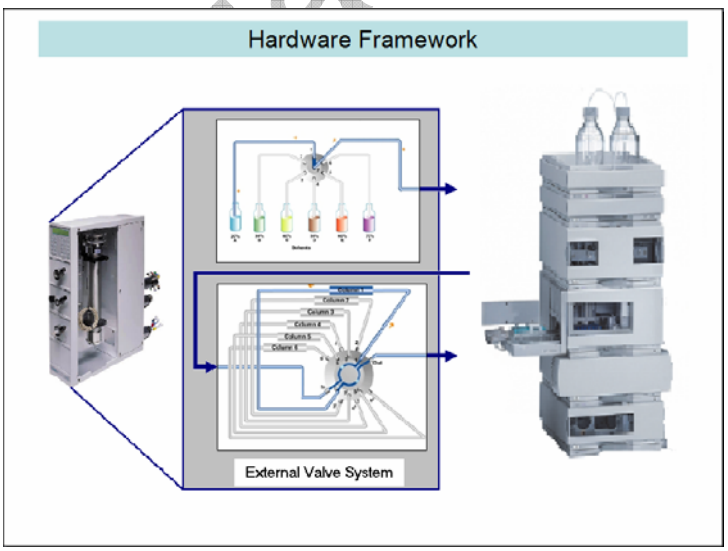
A live experiment was carried out at Pfizer's Ann Arbor laboratories to validate the new QbD practice. To seriously challenge both the approach and the software a special sample was prepared that combined the active ingredients and impurities of two different drug products. The *Phase 1 - Column/Solvent Screening* experiment template was modified in terms of target pH, columns used, and organic solvent type as shown below for the sample compounds.

Experiment Variable	Range or Level Settings
Gradient Time (min)	15.0 — 40.0
pH	2.5, 5.0, 6.5
Column Type	Gemini C18 (Phenomenex, Inc.) Synergi Fusion RP (Phenomenex, Inc.) Luna C18 (Phenomenex, Inc.) Pursuit DiPhenyl (Varian, Inc.) Sunfire C18 (Waters, Inc.)
Gradient Slope (% Organic)	5.0 — 95.0
Organic Solvent Type	Acetonitrile

The Experiment Platform - Hardware

Figure 6 illustrates the HPLC instrument system on which the screening and optimization experiments were run. The instrument system consisted of an Agilent Technologies 1100 HPLC configured with a G1311 quaternary pump, a G1313 autosampler, and a G1315 diode array detector. The HPLC instrument was augmented with a Varian, Inc. Column Valve Module (CVM) that contained a six-position solvent selector valve, a six-column switching valve, and column heating units. The solvent selector valve was connected to the Reservoir A line on the HPLC. This allowed automated multi-solvent screening of solvent linked study factors such as pH and additive. The column switching valve was connected to the main HPLC flow line. This allowed automated screening of up to six columns, or five columns and a bypass line.

Figure 6. HPLC Instrument with CVM



The Experiment Platform - Software

The experiments were generated and analyzed using the Fusion AE™ software program developed by S-Matrix Corporation. Fusion AE implements the phased QbD approach described in this paper using automated statistical experimental design, data analysis, and data modeling protocols. The software automatically converts the experiment designs into the native file and data formats required by the chromatography data system (CDS) software that directly controls the instrumentation. Additionally, the novel approaches to column/solvent screening and the integration of robustness into the method development process are fully automated in this software. The Varian® Galaxie™ CDS software program was used in these experiments. Galaxie provides full level-4 control of all Agilent Technologies HPLC instrument modules and also full control of all CVM components (column ovens and switching valves).

Table 3 presents the *Phase 1 - Column/Solvent Screening* experiment design generated from the template along with the Trend Response results computed directly from the chromatogram data.

Table 3. Column/Solvent Screening Experiment Data Set

Experiment Design				Trend Response Results		
Trial No.	Gradient Time	pH	Column	Compound 1 - Total Peaks	Compound 1 - Resolved Peaks ≥ 1.5	Compound 1 - Resolved Peaks ≥ 2.0
1	40	6.5	Luna C18	12	11	10
2	15	6.5	Sunfire C18	11	9	8
3	15	2.5	Pursuit Diphenyl	10	7	7
4	15	2.5	Gemini C18	12	10	6
5	15	2.5	Sunfire C18	9	7	7
6	27.5	5	Luna C18	11	9	9
7	40	2.5	Sunfire C18	13	12	10
8	40	6.5	Sunfire C18	12	10	9
9	15	6.5	Synergi Fusion RP	12	9	9
10	15	2.5	Luna C18	9	7	7
11	15	6.5	Luna C18	12	9	9
12	40	2.5	Pursuit Diphenyl	11	8	8
13	27.5	5	Sunfire C18	10	9	9
14	27.5	5	Gemini C18	12	10	9
15	40	2.5	Luna C18	12	11	10
16	21.25	2.5	Luna C18	11	8	8
17	27.5	5	Sunfire C18	10	9	9
18	27.5	5	Pursuit Diphenyl	8	7	7
19	40	2.5	Pursuit Diphenyl	10	9	8
20	40	2.5	Synergi Fusion RP	11	9	8
21	40	5	Gemini C18	9	7	7
22	40	6.5	Synergi Fusion RP	11	9	9
23	15	2.5	Synergi Fusion RP	12	10	10
24	27.5	5	Luna C18	11	9	9
25	21.25	6.5	Synergi Fusion RP	12	9	9
26	27.5	5	Gemini C18	11	9	8
27	15	5	Gemini C18	13	11	6
28	27.5	5	Synergi Fusion RP	11	9	8
29	27.5	2.5	Gemini C18	11	9	8
30	33.75	2.5	Synergi Fusion RP	11	10	8
31	27.5	5	Synergi Fusion RP	12	9	8
32	27.5	6.5	Gemini C18	11	8	8
33	40	2.5	Sunfire C18	12	11	9
34	27.5	5	Pursuit Diphenyl	9	8	6
35	40	6.5	Pursuit Diphenyl	11	9	8
36	33.75	6.5	Gemini C18	10	7	7
37	15	6.5	Pursuit Diphenyl	10	8	7
38	40	2.5	Gemini C18	9	7	7

Table 4 presents the regression analysis results for the Total Peaks trend response. The table contains two important results worth describing in detail. First, all equation (study parameter effect) terms are statistically significant. This is seen from the significance test values associated with each term in the table (P-Value less than 0.0500, F-Ratio value > 4.0000, zero outside the 95% confidence interval). Second, all study parameters are represented in the equation in a form related to the nature of their effects (nonlinear, interaction, etc). In fact, as expected a ranking of the effect coefficients identifies the largest effect as due to changing columns (Column 4 in the table represents the effect of switching from Column 1 to Column 4). These results show that a predictive equation has been developed which accurately and quantitatively relates the study parameter effects to one key aspect of compound separation – the visualization of all compounds present in the sample.

Table 4. Equation Statistics – Total Peaks Trend Response

Parameter Name	Coefficient Value	Coefficient Standard Error	Coefficient t Statistic	P-Value	F-Ratio	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Constant	1,415.59	44.80	---	---	---	1,324.10	1,507.08
Gradient Δt	-662.76	121.51	-5.4545	<+/- 0.0001	29.7514	-910.91	-414.61
Column 4	-462.68	105.27	-4.3952	0.0001	19.3178	-677.67	-247.69
(Grad. Δt)*pH	-113.54	55.15	-2.0589	0.0483	4.2389	-226.16	-0.92
(Grad. Δt)*Column 2	442.20	168.63	2.6223	0.0136	6.8762	97.81	786.60
(Grad. Δt)*Column 3	901.09	170.96	5.2707	<+/- 0.0001	27.7807	551.94	1,250.24
(Grad. Δt)*Column 4	781.87	164.97	4.7396	<+/- 0.0001	22.4635	444.96	1,118.77
(Grad. Δt)*Column 5	1,075.53	163.78	6.5667	<+/- 0.0001	43.1221	741.04	1,410.02

Table 5 presents the regression analysis results for the Resolved Peaks (> 1.50) trend response. As for the Total Peaks response all equation (study parameter effect) terms are statistically significant, and all study parameters are represented in the equation in a form related to the nature of their effects (nonlinear, interaction, etc). These results show that a predictive equation has been developed which accurately and quantitatively relates the study parameter effects to a second key aspect of compound separation – the separation of each compound from all other compounds to the extent required.

Table 5. Equation Statistics – Resolved Peaks (> 1.50) Trend Response

Parameter Name	Coefficient Value	Coefficient Standard Error	Coefficient t Statistic	P-Value	F-Ratio	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Constant	9.13	0.12	---	---	---	8.88	9.37
Gradient Δt	-2.01	0.32	-6.2877	<+/- 0.0001	39.5346	-2.66	-1.36
Column 4	-1.26	0.28	-4.5450	0.0001	20.6568	-1.82	-0.69
(Grad. Δt)*pH	-0.38	0.15	-2.6404	0.0130	6.9715	-0.68	-0.09
(Grad. Δt)*Column 2	1.90	0.44	4.2800	0.0002	18.3185	0.99	2.80
(Grad. Δt)*Column 3	3.53	0.45	7.8523	<+/- 0.0001	61.6582	2.61	4.45
(Grad. Δt)*Column 4	2.56	0.43	5.8987	<+/- 0.0001	34.7952	1.67	3.45
(Grad. Δt)*Column 5	3.51	0.43	8.1422	<+/- 0.0001	66.2962	2.63	4.39

Once the software derives the equations from the Trend Response data sets, the equations are linked to a numerical algorithm that identifies the study parameter settings that maximize both responses. In this study the automated optimization analysis immediately identified the column type, pH, and gradient conditions that should be used in the second phase of the method development workflow. These results are presented in Table 6.

Table 6. Optimizer Answer

Parameter Name	Optimizer Result Level Setting
Gradient Time	40.0
pH	2.5
Column	Column 3

It should be noted that the same Trend Response approach described above was used on the 2-column screening study shown previously. The same quantitative results and models were obtained from that experiment data set (results not shown). In this case the automated optimization analysis identified the column, initial hold time, gradient time, and final percent of organic solvent that maximized the Total Peaks and Resolved Peaks responses.

In practice the Trend Response approach will not always yield the optimum HPLC method (instrument parameter settings) in a single experiment, and indeed it is not meant to. The Trend Response approach is part of a phased workflow in which the trend responses enable the experimenter to identify the best settings of parameters such as Column Type and pH; parameters that normally have the greatest effect on separation and therefore cause the most inherent data loss. Once these settings are identified, these parameters are then held constant in a second experiment to optimize the HPLC instrument method. In this experiment current practice experiment results such as resolution can be directly analyzed to identify optimum instrument parameter settings. In fact, the second phase experiment will again include gradient conditions (time and slope) to address total assay time as a key method optimization goal.

Conclusions

Chromatographic analytical method development work normally begins with selection of the analytical column, the pH, and the organic solvent type. A major risk of using a one-factor-at-a-time (OFAT) approach in this phase is that it provides no ability to visualize or understand the interaction effects usually present among these key instrument parameters. Alternatively, a Quality-by-Design (QbD) based methodology would employ a statistical experiment design matrix to study these parameters in combination. However, this approach often results in significant inherent data loss in key chromatographic performance indicators such as compound resolution due to the large amount of peak exchange and compound co-elution common in these experiment data sets. This inherent loss makes it difficult or impossible to quantitatively analyze and model these data sets, reducing the analysis to a pick-the-winner strategy based visual inspection of the chromatograms. The QbD-based practice described here uses a statistical experimental design coupled with automatically computed *Trend Responses*™. This new practice successfully overcomes these problems to provide a rigorous and quantitative methodology for column/solvent screening without the need for difficult and laborious peak tracking implemented in a fully automated HPLC experimentation platform.

Acknowledgements

The authors are grateful to Dr. Graham Shelver, Varian, Inc. for providing hardware, software, and expertise in support of the live experimental work conducted to prove out the Quality-by-Design approach to method development presented in this paper. The authors also want to thank Dr. Raymond Lau, Wyeth Consumer Healthcare, and Dr. Gary Guo and Mr. Robert Munger, Amgen, Inc. for the experimental work done in their labs which supported refinement of the phase 1 and phase 2 rapid development experiment templates.

Acronyms, Tables and Figures

Acronyms:

21 CFR 11 –	Title 21, Part 11, of the Congressional Federal Register
CDS –	chromatography data system
DOE –	design of experiments (also DOX)
FDA –	U.S. Food and Drug Administration
GC –	gas chromatography
HPLC –	high performance liquid chromatography
ICH –	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
PhRMA –	Pharmaceutical Research and Manufacturers of America
QbD –	quality by design
SDK –	Software Development Kit (third-party software development interface)
SOP –	Standard Operating Procedure

Table and Figures:

Bibliography

1. Cornell, John A., (1990), *Experiments With Mixtures*, 2nd Edition, John Wiley and Sons, New York.

2. Dong, Michael W., (2006), *Modern HPLC for Practicing Scientists*, John Wiley and Sons, Hoboken, New Jersey.

3. Montgomery, Douglas C., (2001), *Design and Analysis of Experiments*, 5th Edition, John Wiley and Sons, New York.

4. Myers, Raymond H. and Montgomery, Douglas C., (1995), *Response Surface Methodology*, John Wiley and Sons, New York.

5. Turpin, Joseph (Eli Lilly and Company, Inc.) and Verseput, Richard (S-Matrix Corp.), (2003). "Analytical Method Development & Validation - Increasing Efficiency through Automation". Presented at the PerkinElmer TotalChrom User's Conference, Atlanta, GA.