Don’t Delay Drug Discovery & Development!

Drive chromatography data value using the Tetra Data Platform

Authors:
Mike Tarselli, Ph.D. - Chief Scientific Officer
Spin Wang - Chief Executive Officer

tetrascience

Recommended Reading for: Executive leaders in R&D and R&D IT, Pharmaceutical Technology & Development, QC/QA, manufacturing, digital transformation, IT professionals, data scientists and engineers

Chromatography, using separations and spectroscopic techniques to assess identity and purity of a sample, is omnipresent in modern pharmaceutical and biotech workflows. At multiple stages of drug discovery, measurement matters:

Research: Purifying small molecules, separating components in a complex biological mixture, polishing runs on peptides & oligonucleotides

Development: Characterizing impurities, determining formulation, metabolism / fate studies

Manufacturing: Ensuring GxP compliance, in-process controls, lot and batch parameters

Post-approval: Determining shelf-life, stability, quality control processes

Despite its central place in the discovery landscape, chromatography still runs into roadblocks: you can't easily aggregate data across multiple runs, systems, or across diverse instrumentation. Searching across organizational data silos brings extra headaches - can Development see the Research group's methods? Can Research access Development's lab execution system?

Here are five cases

Case 1 - Dashboarding Disparate Analytical Results

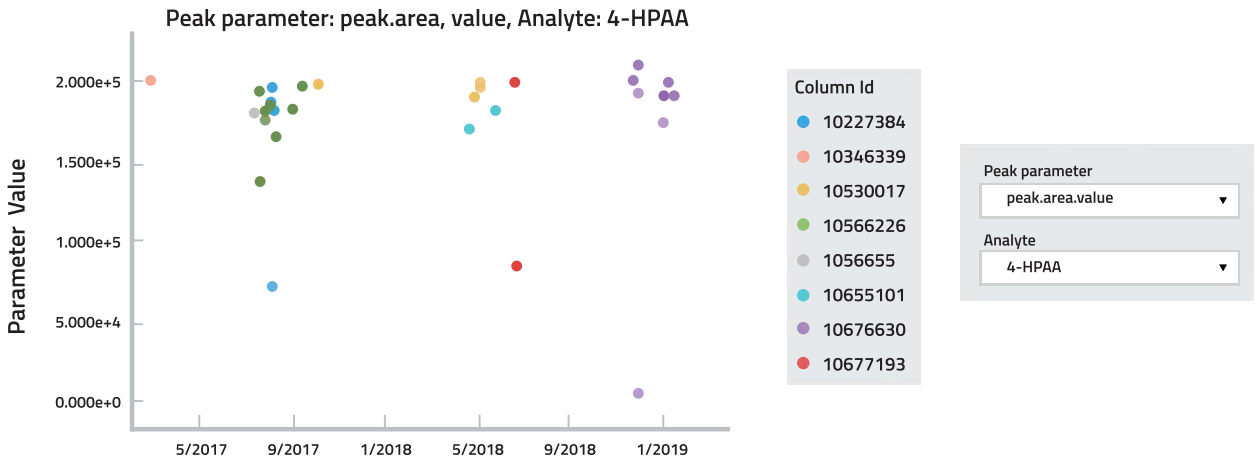
A large European pharma company worked with TetraScience to deploy the Tetra Data Platform into their development organization; expanding eventually into 1.5M injections impacting **1,100 scientists** globally. Annotation of specific molecules under various experimental conditions became trivial; MS, HPLC, solid-state, and NMR chromatograms could be viewed on the same screen, grouped by common ID. In addition to accelerating a scientist's day-to-day operational work, the ability to aggregate characterization data from multiple systems accelerates understanding of the molecule, leading to faster formulation, stability, and eventually taking time off the patent-filing clock. The organization will soon incorporate this infrastructure into their AI/ML data efforts to predict molecular behavior.

Case 2 - System suitability tests (SST) for analytical methods

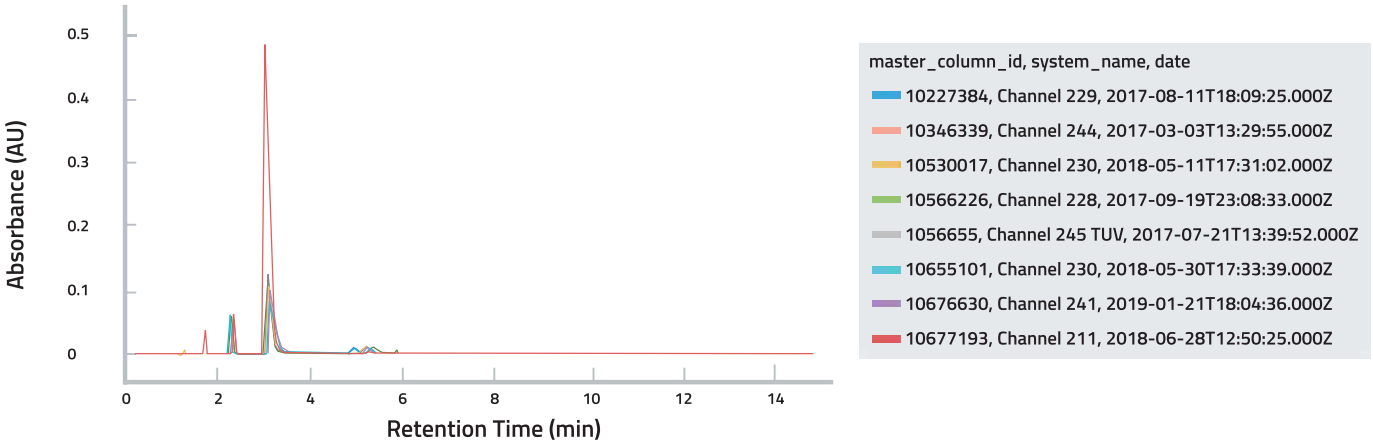
System suitability gates usage of quality-controlled analytical methods. If the test deems your instrumentation unsuitable on a given operations day, you must discard or ignore all data taken through that process.[1]

TetraScience upgraded the SST process for a major global pharma company, including monitoring and reporting. We built **a Jupyter notebook** solution atop Waters Empower data collected by the Tetra Data Platform. This quickly assembled all chromatograms, grouped them by system, column (reverse-phase or polar), and time period, and produced facile graphical overlays to catch peak shape and retention errors in real-time. Researchers could now search across their SST history to catch method issues before another week's runs were rendered invalid. Researchers saved an estimated **8 hours weekly** spent in manual data gathering and comparison. They also gained peace of mind knowing that their methods were **FAIR**, reproducible, and therefore "future-proofed" against future SSTs.

Trending analysis (look at peak parameters by standard analyte over time)



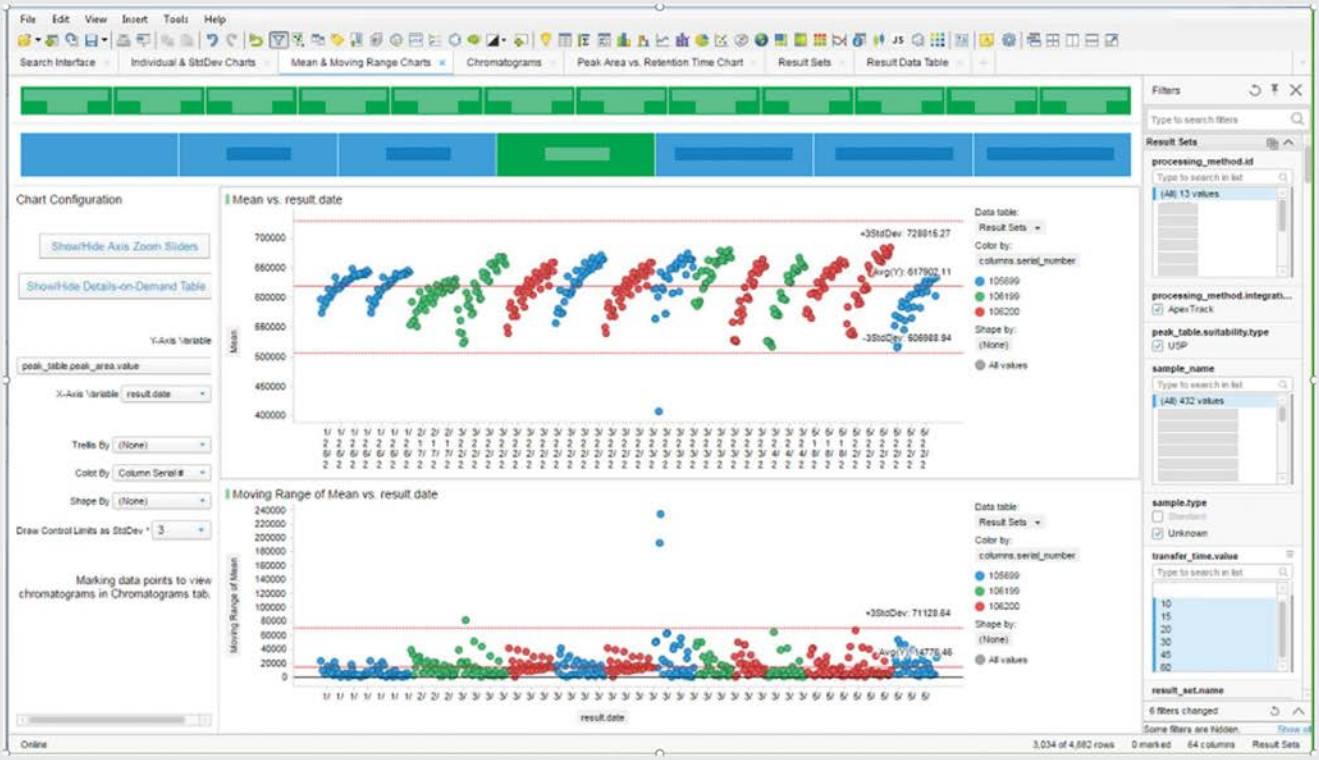
Chromatogram overlay



Case 3 - Column stability / degradation

Columns aren't cheap - a single column for a workhorse LC could run **$400-1000 USD**, and chiral packing adds a 2-3x price multiple on top of that. And yet, entropy wins: eventually columns lose resolution, peaks begin to broaden, shift, or tail. Placing a final material onto a potentially faulty column might result in material loss, or obscure a tiny impurity in a poorly-resolved shoulder peak - leading to costly rework and missed project deadlines. Using the cost estimates from NRDD, even a 1-week delay will expend **$1.5M** in research costs, not including cost of starting materials, reactor time, and staff to re-make the final product; add a factor of two if the material must be produced under GMP. What if you could monitor the health of each column in your arsenal, all the time? What would the value be of this information - aside from **~$30,000 in solvent and consumables per machine**, per year?[2]

Peak area mean value vs. time for individual columns

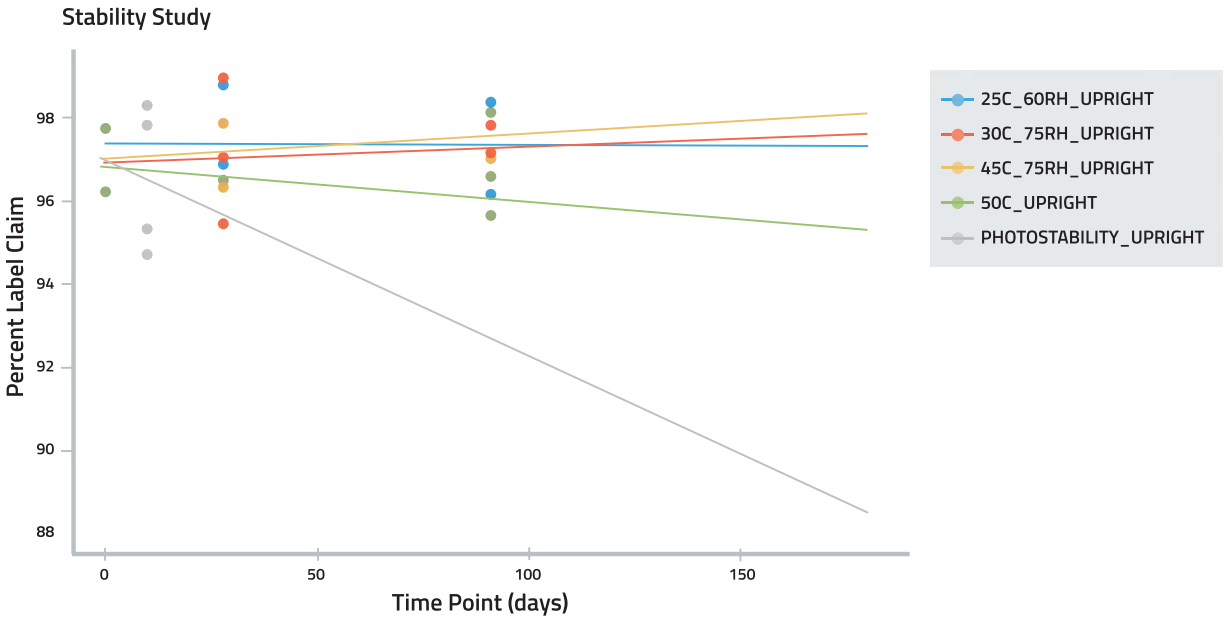


TetraScience, working with a large U.S. biotech, installed a column monitoring system that allowed analysts to search across all chromatography runs, filter out test runs which didn't possess necessary information, and zoom in on polishing runs. They can now perform on-the-fly chromatogram overlays using our **Streamlit** app to examine all scouting runs simultaneously to **determine the highest-fidelity purification method** without one-by-one spectral comparison.

Case 4 - Shelf-life stability study + AI/ML

Could you use the data you have today to deduce sample characteristics like shelf stability, or predict impurity profiles? A second large international pharma experienced issues with reporting active pharmaceutical ingredient **(API)** shelf-life estimates: their tabular format complicated communication of estimated compound parameters - water content, stability, impurities - to other divisions. Once they had access to their chromatography data, they quickly developed a graphical method to quickly compare analyses.

Desired state for shelf-life predictions: get out of tables!



Case 5 - Legal, Compliance, and Financial Risks

Perhaps you're not a researcher...maybe you support a group like Legal, Finance, or R&D IT. Or you're a pharmaceutical executive who just needs to see the window of market opportunity widened as much as possible for your new biologic or drug. Given the "first to file"[3] change to U.S. patent law in 2011, companies face a balancing act - filing too early costs valuable peak year sales, but filing too late allows competitors to catch up. Even for a non-blockbuster drug, which *Nature Reviews Drug* Discovery pegs[4] at $500M USD / year (2020), every lost day of exclusivity results in **$1.4M** of lost potential revenue.

What if you could *shrink* the work needed to complete time-consuming post-patent steps like product development or stability testing? Or, better yet, to expedite analytical data that lead you to *faster filing*? Here are some use cases where the Tetra Data Platform could help:

Before Patent Filing:

- Integrate CRO LC/MS datasets efficiently
- Quickly characterize pharmaceutical leads and streamline biologic bioprocesses
- Measure system uptime and expedite method development
- Calculate column end-of-life to avoid losing products and reduce equipment expense

Post-patent:

- Investigate API stability and shelf-life
- Monitor Standard Operating Protocol (SOP) deviations like manual peak integration or unplanned changes
- Avoid costly delays during regulatory submissions to FDA or EMA
- Quickly identify polymorphs and optimal particle sizing

By having transparent, harmonized access to data across your instruments and the development application landscape, you save not only time on the market exclusivity clock, but also **manual data wrangling, expensive product development**, and **reduce risks** when filing for regulatory approval.

You also save daily burn rates of running a full-scale research enterprise - thus saving time and money on both sides of the patent process. Using data from *JAMA*,[5] we can take the capitalized mean cost of R&D for all approved drugs in the past decade - $374M - and quickly estimate that each research "day" costs **$300-400K USD**!

SOP Deviations

If a pharmaceutical development organization has adopted GXP processes, SOP deviations may prove costly at a future audit. The Tetra Data Platform provides reproducible, audit-friendly data analysis; one pharmaceutical client requested we eliminate manual peak integration as this would trigger their QC process. Seemingly small deviations such as this to protocol, if investigated and caught by a downstream audit, would likely result in corrective actions that will **add weeks to the process** and potentially jeopardize a final submission document.

For more information on how the Tetra Data Platform can drive chromatography data value, visit **tetrascience.com**

Footnotes

[1]Dong, Paul, Gerhanov. "Getting the Peaks Perfect", ACS Publications, 2001: https://pubsapp.acs.org/subscribe/archive/tcaw/10/i09/html/09dong.html

[2]ThermoFisher Scientific, 2020: https://www.thermofisher.com/us/en/home/industrial/chromatography/liquid-chromatography-lc/hplc-uhplc-systems/hplc-system-total-cost-ownership-calculator.html

[3]Justia, May 2019: https://www.justia.com/intellectual-property/patents/first-to-file-rule/

[4]Moss, Wurtzer. "Developing Blockbuster Drugs: both nature and nurture.", Nature Rev. Drug Disc., 2020: https://www.nature.com/articles/d41573-020-00061-9

[5]Wouter, McKee, Luyten, JAMA, 2020: https://jamanetwork.com/journals/jama/fullarticle/2762311