

Data-Driven Quality Control

CASE STUDY

THE CASE FOR THE TETRA SCIENTIFIC DATA CLOUD

Measure what matters: Chromatography data presents interesting challenges for modern life sciences organizations: Millions of samples are manually processed each year across the pharmaceutical industry. The technique underlies every step of pharmaceutical production, from research leads through post-clinical manufacturing lots. Despite its central place in the discovery landscape, chromatography suffers from a fundamental limitation: You can't easily aggregate data across multiple runs or across your diverse in-house instrumentation. Manual report creation and conversions between data formats are still the norm.

SUCCESS METRICS

<p>Saved Time</p> <ul style="list-style-type: none"> • Save 2+ hours per week, per scientist of manual file transfer work • Provide structured and clean data integration for other systems to consume 	<p>Reduced Complexity</p> <ul style="list-style-type: none"> • Reduce manual errors; Results are directly calculated and derived from raw data • Automation allows streamlined downstream analysis via data science applications (e.g. Streamlit apps)
<p>Accelerated Insights</p> <ul style="list-style-type: none"> • Improved data-driven decision making (Result is calculated and generated much faster) • Harmonized, vendor-agnostic data accelerates ability to run AI/ML and data science projects on top of it 	<p>Improved Transparency and Future-Proofing</p> <ul style="list-style-type: none"> • Versatile pipelines are configurable and extensible to handle rich set of technologies and formats

Who Should Read this Study?

R&D IT Directors and Business Leaders; Quality Control teams in Pharmacology, Development, or Manufacturing; Scientific Project Managers, Operations Data Scientists; Lab Automation Engineers; Metabolism / Fate Study Heads



Product Focus

- Small molecule therapeutics, biomarker development, computational screening

Client/ Customer

- Top 15 biopharma

KPIs / Results

- Saved 8 hours weekly in reporting;
- \$30,000 in consumables expense per HPLC; Alleviated days to weeks to rework production batches;
- Resolved inefficiencies in interdepartmental communications

How can you get more value out of your chromatography data using the Tetra Scientific Data Cloud? Following within are specific use cases equally applicable to enterprise pharmaceutical companies with fleets of 1000+ chromatography instruments as they are to small biotech start-ups running a single analytics lab.

CASE 1 - SYSTEM SUITABILITY TESTS (SST) AND COLUMN DEGRADATION CHECKS

System suitability testing gates usage of quality-controlled analytical methods. Simply put, if the test deems your instrumentation unsuitable on a given operations day, you must discard or ignore all data taken through that process [ACS, 2001]. How can you automate SST using existing HPLC data, and look for trending issues in peak shape, retention, and performance?

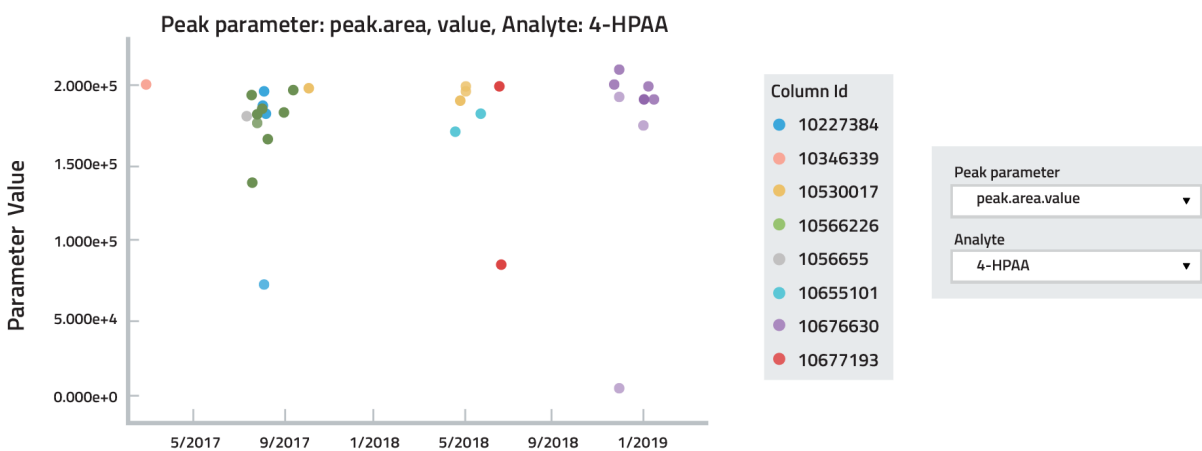
A top 15 global biopharma asked TetraScience to upgrade their SST process. Their previous process:

1. Run test methods on six columns across 15 instruments - 90 total chromatograms
2. Manually open each sample set from Empower
3. Compare by eye to catch adverse behavior or output errors
4. Provide weekly report to ensure operation within specs

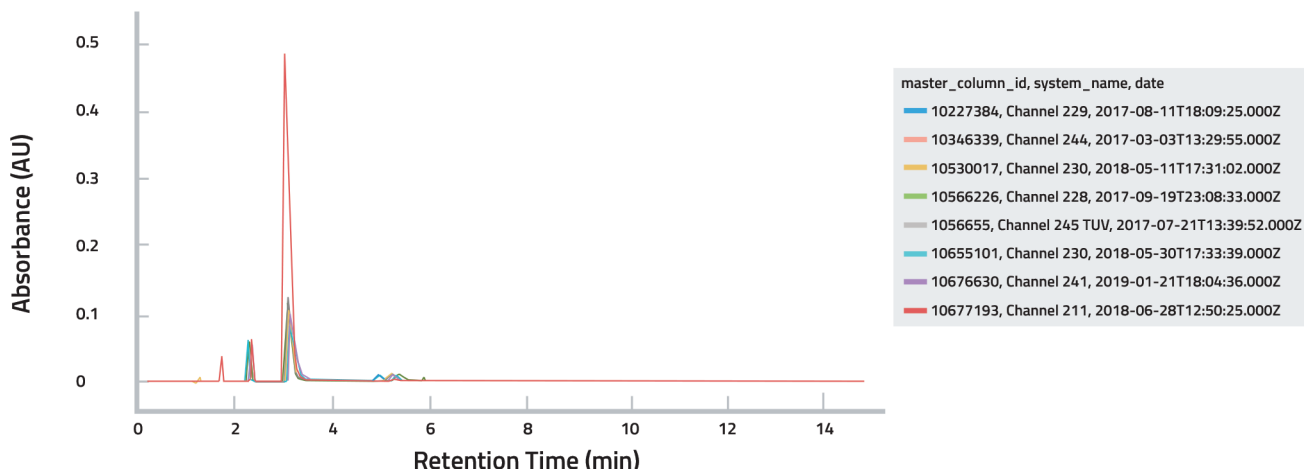
Using our open Empower data connector, built on the Tetra Data Platform, we first ingested and processed all of the company's HPLC data into a FAIR format. We then built a Jupyter notebook downstream, which quickly assembled all chromatograms, grouped them by system, column (reverse-phase or polar), and time period, and produced facile graphical overlays to catch peak shape and retention errors in real-time. Researchers could then search across their SST history to calculate column end-of-life and catch issues before another week's runs were rendered invalid.

Researchers **saved an estimated 8 hours** weekly spent in manual data gathering, visual comparison, and reporting.

Trending analysis (peak parameters by standard analyte over time)



Chromatogram overlay



Speaking of column degradation - placing a final production batch onto a potentially faulty column might result in material loss, or perhaps obscure a tiny impurity in a poorly-resolved shoulder peak - leading to costly rework and missed project deadlines. What if you could monitor the health of each column in your arsenal, in near-real time, knowing when it had reached its end-of-life? What would the value be of this information - aside from ~\$30,000 in solvent and consumables per machine, per year?

Using factors such as peak area, retention time, and tailing factor from the SST, this Global Biopharma used existing runs to "predict" failures before they could occur. The graphic below shows the power of learning from existing data - simple plotting of peak mean (or any other relevant spectral parameter) over time reveals the slow drift into obsolescence.

Peak area mean value vs. time for individual columns, showing considerable drift over a 4-month period

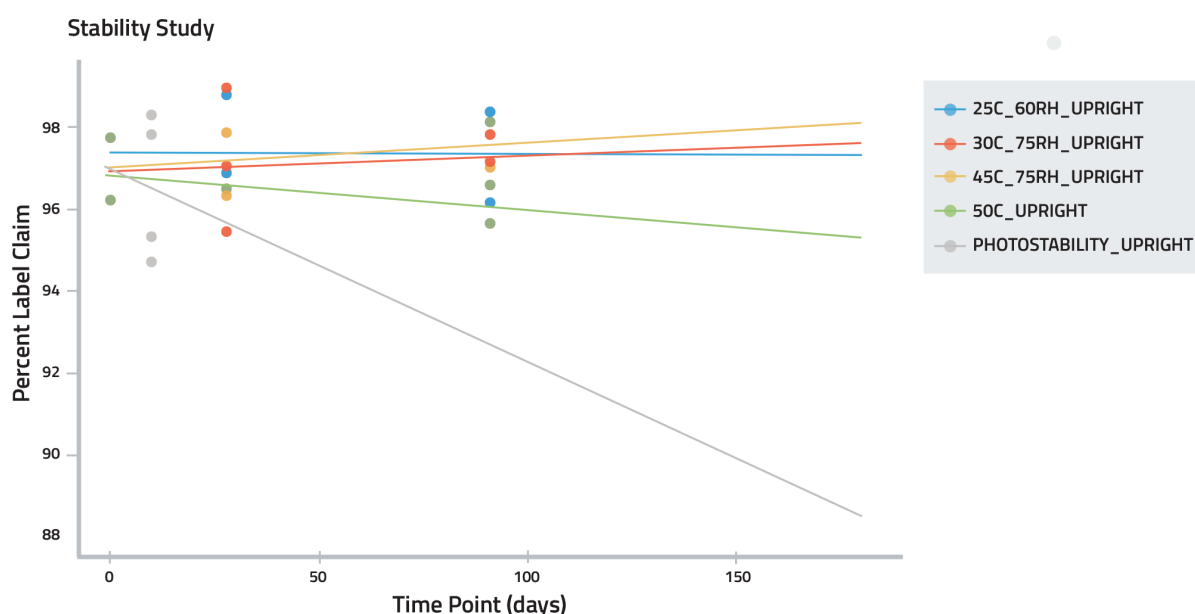


Finally, our Global Biopharma client uses the Tetra Data Platform to calculate system usage for its entire process development division, monitoring parameters like total solvent, system uptime, and run length across their chromatographic fleet. Paired with the ability to inspect specific columns by data detective work, TetraScience automates QC reporting, reduces process risk, and eliminates manual data work.

CASE 2 - SHELF-LIFE FOR ACTIVE PHARMACEUTICAL INGREDIENT (APIs)

Our Global Biopharma partner had also experienced issues with reporting API shelf-life: the tendency of molecules to degrade due to heat, light, water, pressure, and multiple other environmental or handling variables. Having stored and reported their data in a tabular format for years didn't permit quick, intuitive communication to other company divisions. By utilizing the Tetra Data Platform connector to ingest and harmonize all Empower data (see above), a graphical method to quickly chart observational measurements and offer predictions (fitted line) could be realized.

Desired state for shelf-life predictions: A single graphical overlay with immediate information about compound stability.



Better still: multiple such analyses could be grouped, depending on environmental conditions, batch composition, storage containers, country specs, etc.

CONCLUSION

The automation and analyses were made possible through centralized, harmonized data from Empower being available for further inquiry across the organization. This company realized manual labor savings, consumables cost savings, streamlined reporting between functions, and made data available and reusable in adherence to FAIR ideals.



TetraScience is the Scientific Data Cloud™ company with a mission to accelerate scientific discovery and improve and extend human life.

To learn how you can unlock the power of harmonized, FAIR data in the cloud with the only open, cloud-native platform purpose-built for science, visit tetrascience.com

Corporate Headquarters | 177 Huntington Avenue, Suite 1703, Boston, MA 02115

Atlanta | 3280 Peachtree Road NE, 7th Floor, Atlanta, GA 30305

© 2022 TetraScience, Inc.