
# Machine Learning to Drive Clinical Trial Efficiency via Adverse Event Data and Predictions

## Integrated Clinical Solutions platform provides data analytics and visualizations for rapid interpretation of study performance

Clinical trial data management has become increasingly challenging as clinical studies evolve in complexity. A common pain point for those managing clinical research and clinical development is being able to quickly access and analyze all the data connected to a study to understand how a trial is progressing against patient safety goals —is the drug showing efficacy, what are the adverse events, and if some trial sites are underperforming and need more active management. Having all these data available, via a single solution, allows clinical managers to make timely adjustments to the trial to ensure the drug candidate is safe and efficacious and continues to progress toward regulatory approval.

In this white paper, we will examine how clinical trial data can be accessed in near real-time and then analyzed to empower information-driven, proactive study management by clinicians. We will also explore how adverse event (AE) data can be used to monitor a trial's safety design and performance. We will discuss a clinical analytics platform that provides cross-functional collaboration along with a workflow to track the review of emerging clinical trial data and that facilitates the understanding of that emerging data. Furthermore, a new Machine-Learning (ML) algorithm is demonstrated that uses bootstrapping to analyze site and patient AE data to predict whether a site is overperforming or, more importantly, underperforming in its AE reporting. Using the AE data, sponsors can quickly identify at-risk clinical sites that need closer management to ensure optimal performance. The combined use of these tools can identify potential improvements in trial design through amendment or clinical strategy decision-making.

## Fast easy access to critical data

Decision-making in clinical trials is too often hampered by the inability of clinical research leaders to get a holistic view of all the data across the entire study. Slow delivery of data, missing data from some trial sites, and lack of data harmonization prevents sponsors from drawing insights from related occurrences across multiple sites. This lack of visibility often arises from delays in data access, mapping, standardization, and aggregation; dependency on multiple systems to complete, track and review data; and reliance on limited analytics such as spreadsheets.

That is why a single, powerful, centralized data analytics solution is needed: to help managers monitoring trials to identify and analyze data across all sites. The Revvity Signals Clinical Solutions platform, powered by TIBCO® Spotfire®, is designed to facilitate clinical use cases via a portfolio of pre-built modular analytics solutions. Each solution integrates seamlessly to manage the range of clinical development needs, from clinical data review and medical review to risk-based monitoring and clinical operations, to pharmacovigilance and efficacy analysis.

For example, this approach allows Clinicians and Data Managers to review AE data visually across all sites to quickly see relationships within the data. This allows Clinicians and Study Managers to address issues much earlier and prioritize resources aimed at generating high-quality data, accelerating submissions to regulators, and driving decision-making.

### Monitoring adverse events

Understanding adverse events in a clinical trial and responding to potential adverse event reporting issues is critical to keeping the trial on track. This requires both the ability to quickly visualize the big picture across the entire study and to easily drill down into individual patient data to develop a complete understanding of what factors may be contributing to the reported adverse events.

Below is a bar graph from the Clinical Solutions platform that shows all the adverse event preferred term data collected to date, labeled by the type of event on one axis of the graph and number of cases of that event on the other (Fig. 1). The longer the bar, the more patients have experienced that particular AE, while color coding within the bar shows the relative severity of each instance. This information gives cross-site insight into patient safety concerns that may not be apparent to principal investigators at each site, driving sponsor safety oversight. Using this information, trial managers can further drill into the individual patient data and subgroups of patients for each AE severity, which may reveal risks or other disease or treatment insights related to that specific AE during trial conduct for accelerated clinical decision-making and safety management.

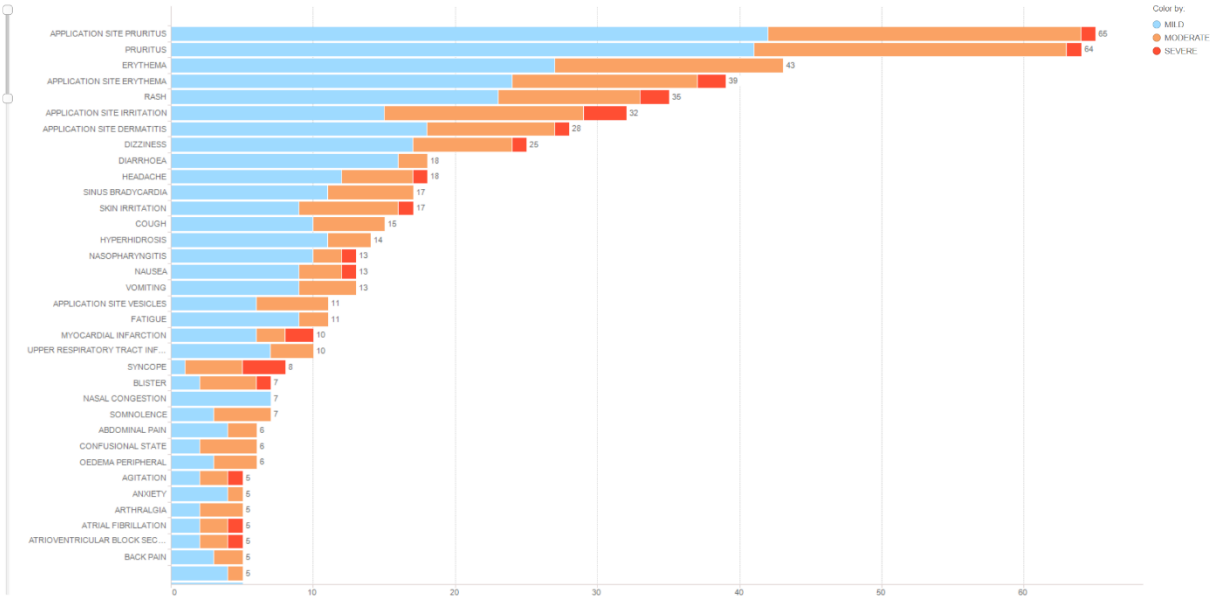


*Figure 1: Adverse event preferred terms by unique patient counts.*

The Signals Line Listing Review (SLLR) is another Clinical Solutions' analytic that provides a workflow guiding AE review in a clinical trial. It equips reviewers with scalable tools to review, query, and track in real time line listing data, and collaborate cross-functionally during the review. As clinicians or study managers review a case, raise questions about the data, or route messages to other clinical team members for action or escalation, an audit trail is created and the status of review is updated. Monitoring progress of the line listing review is an essential regulatory component of the safety review process, and SLLR ensures compliance while saving hours of time for safety review and data management teams.

Visualizations within the SLLR help users understand trends in emerging responses to treatment regimens – whether issues with the initial study design need correcting, or among AEs that have occurred, whether there is a relationship with the underlying disease or with the treatment. The analytics in the Clinical Solutions platform help clinical research teams improve data review, make highly informed decisions about trial conduct and potential trial amendments, and suggest alternative designs for future trials to ensure patient safety.

## Getting a handle on AE underreporting

The FDA's stance on identifying and managing AEs within clinical trials centers around its core tenet of first doing no harm to the patient. To help meet this goal, the agency encourages sponsors to identify not only what is happening with one patient, but across all patients at every site. These requirements drive sponsors to centralize data collection and to take a statistically driven approach to analyzing trial data to improve safety.

Whether the AE is mild or serious, it is imperative that both the patients and the clinical sites report every AE they encounter and provide that information to the sponsor to ensure full transparency with regulatory bodies. These data are critical to patient safety, identifying potential side effects of the medications and informing dosing limits.

For a host of different reasons, some trial sites simply don't always report to sponsors all the AEs that occur at their site in a timely way. Called adverse event underreporting, this gap is not easy to detect while the trial is running, as sponsors can only see the data that sites are reporting, not what isn't being reported.

Standard practice for attempting to identify AE underreporting today might look something like this graph (Fig. 2). In theory, such a graph is intended to indicate which sites may be underreporting by comparing them to other sites in the trial and the average count of AEs reported. When you add a site average line, more than half of the sites fall below that, making it difficult to decide which may need a visit from a monitor to get them on track. Further, if all of those that fall below the average require a visit, this can quickly become untenable as many mid-sized and smaller companies simply don't have the personnel.

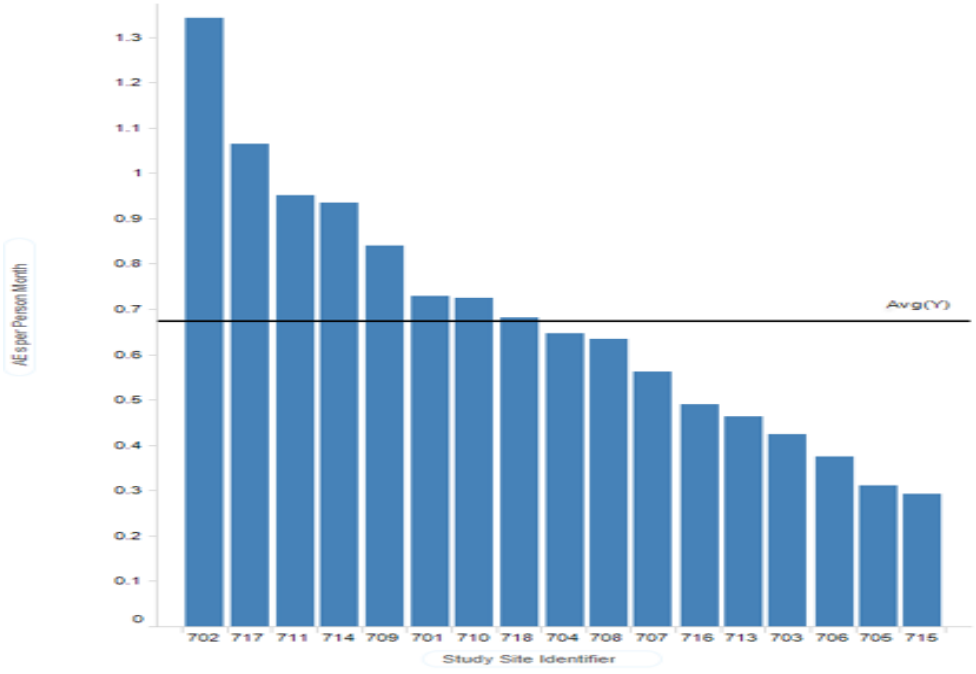


*Figure 2: The standard practice today is represented as a bar chart, but the goal should be to visualize data to prioritize recommendations for study teams.*

In addition to human resource constraints, the very basis of the decision making may be flawed. The assumption is that other sites are a reasonable comparison, but that may not be true. Subjects are not uniformly distributed across sites; for example, some sites may have older or younger populations than others, influencing the number of AEs you would expect to see from that site, for example.

The reality is, standard practice cannot provide a snapshot of what sponsors can reasonably expect for AE reporting from each of the sites and, therefore, can't accurately gauge which sites are lagging.

## *A better way*

What if we could use a statistical method called bootstrapping that would take the AE data received to date in a study and simulate the AE counts we would expect sites to report? This method could help mitigate the lack of uniformity of patients across sites influencing expectations, creating a specific threshold that is based on the subjects in the trial regardless of site affiliation. With a more accurate, tailored expectation, identifying sites that are potentially underreporting becomes possible.

Revvity Signals, in its Clinical Solutions platform, is leveraging a new, proof-of-concept, open-source algorithm developed by Roche that can do just that: run thousands of simulations to help prioritize sites by their chance of being outliers from expectation and thus at risk of potentially underreporting. While the algorithm is written in the R programming language, Clinical Solutions users are able to seamlessly deploy the algorithm within TIBCO Spotfire for visualizations that don't require in-house R programmers.

Visualizations employing the predictions from this tool can produce a graph like the one seen below (Fig. 3). The simulated expectation line is gold, and we can see each site as a blue line in the graph on the left. On the right, we have individual sites that are likely to be mathematical outliers from this expectation. Those sites most likely to be outliers are highlighted in progressively darker shades of blue as the chance of underreporting increases, with the exact percentage highlighted in that site's breakout graph on the right.

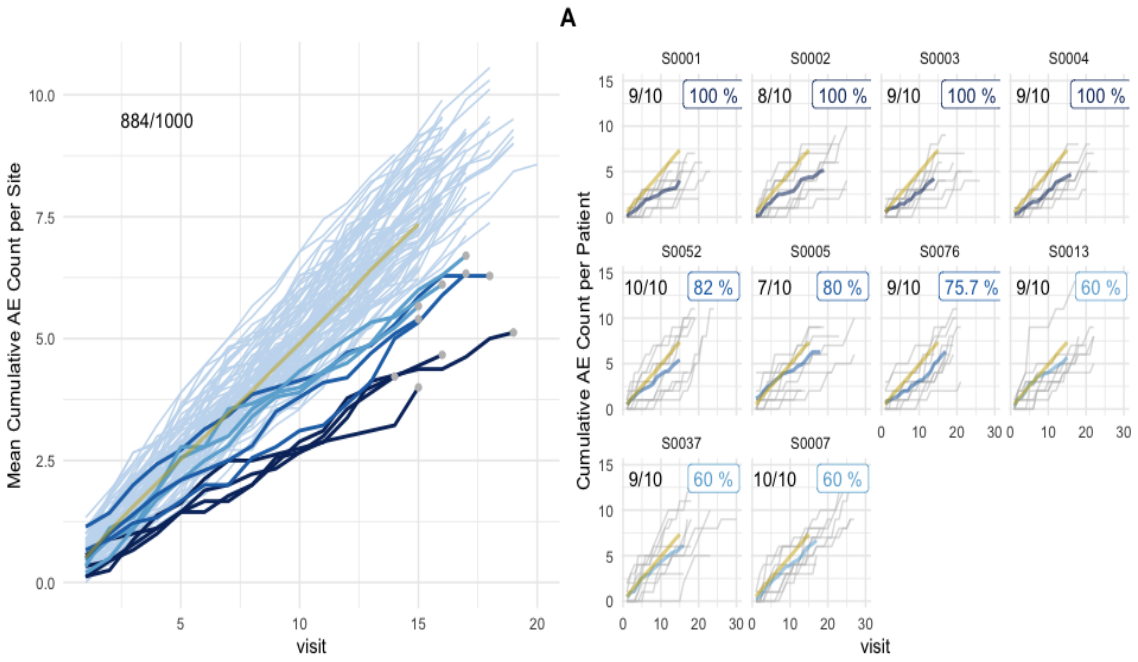


*Figure 3: Output from running Roche's algorithm on its provided example data in R. It compares the simulated expectation to the data reported from sites and, for those sites with more than a 50% chance of being mathematical outliers, it presents more detailed information on the right. This extra information includes individual patient information about their AE reports over time, called out in gray.*

It should be noted that those sites in the top right of the graph with 100% chance of being outliers aren't necessarily underreporting—there are many legitimate reasons why a site may report fewer AEs than expected. Therefore, the odds of being an outlier are not the odds of underreporting, however it still provides a useful metric for study teams to consider when deciding whether they need to deploy a monitor or other oversight. Based on this visual, for example, and the powerful roadmap it provides, the sponsor may choose to have one monitor visit the four different sites that have the highest probability (100%) of outliers, while keeping an eye on the data coming in from the other sites shown as having a 60% or higher chance of being outliers.

Using TIBCO Spotfire analytic capabilities, each bootstrapped site in (Fig. 3) can be linked to the actual AEs collected at that site, allowing drill-down into the actual AE details for each site that bootstrapping predicts is an outlier and could be underreporting. The bootstrapping approach provides more information than was previously available to help study teams focus monitoring activities on the most important sites to achieve more accurate AE data collection and reporting.

*But how does this algorithm work? What is inside the "black box?"*

We start by determining which patients are eligible to be included in the algorithm. In (Fig. 3) above you'll notice that the individual sites have a fraction noted on the line graphs, which represent the number of eligible patients at that site over the number of patients at the site in total. The reason for this distinction is because it is not possible to have a patient newly screened into a trial included in that site's adverse effect count. They simply haven't had enough time to accumulate AEs. So, we calculate a benchmark visit that patients need to reach in order to be eligible. This benchmark can be the default calculation, or if there is some visit that makes sense for the specific structure of the trial, that is also appropriate.

Next, we help mitigate the lack of uniformity of patients in the real world by randomly assigning patients and their associated AEs to trial sites. In the real world, patients are assigned at a particular site, often based on proximity and ease of access from the patient's home. The algorithm makes no such distinction and instead imagines a general pool of patients that could be assigned to any trial site. So, if a trial site has in reality three algorithm-eligible patients, the algorithm will randomly select three patients from the entire pool and assign them to that site—along with all their relevant data. Now, the algorithm will report on the adverse events for this site based on those three randomly assigned patients. This is repeated for every site in the trial and the algorithm then computes the adverse event over time curves. This process is repeated one thousand times in order to understand how the expectation for AE reporting fluctuates, giving us a better final threshold. Another way to look at it is the algorithm pretends to run the trial one thousand different times using a patient population that has the same gender breakdown, health conditions, and ages as those in the actual trial.

Using the simulated trial information allows trial sponsors to understand the distribution of adverse event reporting, and then quantify how likely this is to occur in the actual trial itself. The open-source algorithm, in the TIBCO Spotfire environment, provides a valuable predictive tool to determine which sites may need more management than others over the course of clinical trial, an important consideration for sponsors looking to efficiently allocate resources.

## Improved AE insights

Given the importance of patient safety, the management and evaluation of adverse events plays a critical role to ensure the success and integrity of a clinical trial. Proper AE evaluation is important to ensure patient safety, but also to enable swift response to significant AEs that could have a material effect on a trial's success or failure. Revvity Signals Clinical Solutions provide powerful visual representations of AEs and insights into a trial's performance, while also

allowing clinicians access to patient-level granular data. Combined with new predictive technologies such as Roche's site AE algorithm, clinical trial managers now have a collection of powerful tools for analyzing and collecting AE data for the effective management of a successful clinical trial.

For more information on Revvity Signals Clinical Solutions, visit: Revvity Signals Clinical Solutions

revvitysignals.com
940 Winter Street
Waltham, MA 02451 USA
P: (800) 762-4000 (+1) 203-925-4602

Revvity Signals
RevvitySignalsSoftware
revvitysignals
Revvity_Signals
RevvitySignals

Copyright ©, Revvity, Inc. All rights reserved. Revvity® is a registered trademark of Revvity, Inc. All other trademarks are the property of their respective owners.