**Benchmarking Protocol Deviations and Their Variation by Major Disease Categories**

Kenneth Getz¹ · Zachary Smith¹ · Ananya Jain² · Randy Krauss³

Received: 21 January 2022 / Accepted: 18 March 2022 / Published online: 4 April 2022
© The Drug Information Association, Inc 2022

**Abstract**

**Background** Little to no data exist quantifying and benchmarking the magnitude of protocol deviation experience.

**Methods** Nearly two-dozen companies provided the Tufts Center for the Study of Drug Development (Tufts CSDD) with data on the design and the performance of 187 protocols.

**Results** The results of this working group study show that phase II and III protocols have a mean total of 75 and 119 protocol deviations, respectively, involving nearly one-third of all patients enrolled in each clinical trial. Oncology clinical trials have the highest relative mean number of protocol deviations affecting more than 40% of patients enrolled in each trial. The number of endpoints, the number of procedures per visit, and the number of countries were modestly positively associated with and predictive of, the incidence of deviations per protocol. A strong positive relationship was shown between the number of investigative sites and the number of protocol deviations.

**Conclusion** The results of this initial study provide useful measures that sponsor companies can use to benchmark their own protocol deviation experience, identify factors most associated with protocol deviations, and determine whether remediation is warranted.

**Keywords** Protocol deviations · Protocol changes · Protocol design

**Introduction**

Despite their regimented structure, protocols are rarely followed exactly as planned. Widely referred to as protocol deviations—these changes are generally executed by the study staff, while the clinical trial is underway to accommodate study volunteers encountering difficulties complying with the schedule of visits or adhering to study medication administration requirements. Investigative site staff may also deviate from the protocol when facing difficulties finding eligible patients and challenges following and performing clinical and administrative procedures dictated by the protocol.

Protocol deviations are the top reason for clinical trial enforcement actions. Based on Food and Drug Administration (FDA) inspections of investigative sites, approximately one-third (30%) of all warning letters are due to the failure to follow the investigational plan [1]. This proportion is nearly twice the second most common reason: 17% of all warning letters are due to the failure of the investigative site to adequately maintain source documents [2].

Deviations from protocol procedures may be minor. A subset of all deviations, however, are more significant and may harm the integrity of the study, may put the rights and safety of study volunteers at risk, and may detract from the completeness, accuracy, and reliability of the study data [3]. Based on a cross-functional working group of pharmaceutical companies, Galuchie et al. have proposed definitions to differentiate deviations from protocol violations [4].

During the past decade, protocol deviations have received increased attention as part of a broader effort to improve protocol design and executional quality [5, 6]. Consortia and collaborative groups have developed consensus definitions of protocol deviation types as well as frameworks to prevent and manage them. The Drug Information Association’s
Good Clinical Practice and Quality Assurance community, for example, conducted an industry-wide survey resulting in common definitions and best practices for identifying, classifying, and managing protocol deviations [3]. More recently, industry consortium TransCelerate BioPharma developed a conceptual protocol deviation identification, management framework, and toolkit—incorporating risk- and issue-management principles [4].

Despite this growing attention over the past ten years, to our knowledge, little to no data exist quantifying and benchmarking the magnitude of protocol deviation experience. The purpose of this paper is to share the results of a recent study conducted by the Tufts Center for the Study of Drug Development (Tufts CSDD). It is our hope that these results will serve as a useful baseline measure that sponsor companies can compare against their protocol deviation experience and determine if remediation may be required.

**Methods**

Data on protocol deviations and other protocol design characteristics were collected and reported by twenty major and mid-sized pharmaceutical companies and contract research organizations (CROs): Amgen, AstraZeneca, Biogen, Boehringer Ingelheim, Bristol-Myers Squibb, CSL Behring, Eli Lilly, EMD Serono, GlaxoSmithKline, Janssen, Merck & Co. (Kenilworth NJ), Novartis, Otsuka, Parexel, Pfizer, Roche, Sanofi, Takeda, UCB, and Veristat.

Clinical and clinical operation professionals from each participating company gathered data from a convenience sample of clinical trial protocols that had received final protocol approval between January 2013 and December 2018 and had a primary completion date or database lock date before December 31st, 2019. CROs participating in the study gathered protocol data specifically from client companies that were not represented by the other participating sponsor companies in the working group.

Each participating company was asked to select protocols representative of their current portfolio of clinical trial activity and to include protocols from each of three phases (i.e., Phase I, Phase II, and Phase III). Each participating company submitted between 6 and 21 protocols. The analysis dataset focused on traditional protocol designs and excluded master protocols and adaptive designs.

The data collection process deployed in this study followed the approach that Tufts CSDD has used to evaluate trends in, and the impact of, protocol design practice since 2008. The results of these studies have been published extensively. A number of design variables were gathered, including number and type of endpoints, number of eligibility criteria, number of distinct and total procedures performed, number of countries and investigative sites where the protocol was conducted, and number of planned study volunteer visits per month.

Clinical trial performance and quality variables were also gathered, including clinical trial milestone durations, retention rates, the number of substantial protocol amendments, and the total number of protocol deviations. Planned and actual milestone durations were assessed. The following performance and quality variable definitions informed the data collection process for this study:

- **Study start-up duration**—days from Protocol Approval to First Patient First Visit (FPFV);
- **Duration to complete all first patient visits**—days from First Patient First Visit to Last Patient First Visit (LPFV);
- **Treatment duration**—days from Last Patient First Visit to Last Patient Last Visit (LPLV);
- **Study close-out duration**—days from Last Patient Last Visit to Database Lock (DBL);
- **Total trial duration**—days from Protocol Approval to Database Lock;
- **Drop-out rate**—number of patients completing the clinical trial divided by the total number enrolled;
- **Protocol deviation**—total number of changes, divergences, or departures from the study design or procedures as defined by the protocol;
- **Substantial protocol amendments**—total number of changes made to the protocol, in all countries where it is executed, requiring suspending enrollment, obtaining internal approval followed by approval from an ethical review board or regulatory authority, and re-consenting study volunteers.

Several other measures were created for this analysis. Proportion of patients with protocol deviations were derived by dividing the total number of patients enrolled by the total number of deviations per protocol. Differences between planned and actual cycle times were also calculated using the planned and actual dates provided by companies. Protocols in the ‘LOW’ deviation cohort are those with the relative number of deviations below the median for the entire sample. The ‘HIGH’ deviation cohort had protocols with a relative number of deviations above the median for the total sample.

Data for Phase II and III protocols were combined for comparisons by major disease category given the small sample size by individual phase. Descriptive statistics including means, coefficients of variation, and medians were calculated for the number of protocol deviations per protocol and the proportion of patients with protocol deviations, by phase, and by disease category. Logistic regressions which included individual design variables (number of endpoints, number of eligibility criteria, number of unique procedures, total number of procedures, number of procedures per visit,
number of countries, number of clinical sites, total number of patient visits, and number of patient visits per month) and protocol deviation cohort were conducted. Odds ratios that a protocol would be in the ‘HIGH’ deviation cohort were calculated. Correlations between the number of protocol deviations and several clinical trial outcomes were calculated: study start-up duration, duration to complete all first patient visits, treatment duration, study close-out duration, total clinical trial duration, drop-out rate, and the number of substantial protocol amendments. Finally, mean differences between planned and actual cycle times were calculated for the ‘LOW’ and ‘HIGH’ deviation cohorts.

Protocol data were stored as an excel file and saved on a secure, shared, online drive. Analysis was conducted in SAS 9.4.

**Results**

A convenience sample of 187 protocols met the criteria for this analysis. Table 1 shows the distribution of protocols by phase, disease category, and Low/High number of deviations.

Means for the total number of protocol deviations, per protocol, and by phase are presented in Table 2. Each phase III protocol has a mean number of 118.5 total deviations, involving approximately one-third of all patients participating in that protocol. Phase I protocols have the lowest mean number of deviations and involve half the proportion of patients than those observed in phase II and phase III protocols. The coefficients of variation around the mean number are very high—between 1.7 and 2.0—indicating that total deviations per protocol are widely dispersed and inconsistent.

Combined, phase II and III oncology protocols average almost 20% more total deviations than the average for non-oncology protocols at 108.8 and 91.9, respectively. Protocol deviations in combined phase II and III oncology protocols involve 47% of the total study volunteers, nearly double the proportion observed in non-oncology protocols. Rare disease indications average a lower relative number of total deviations (78.1) among a smaller proportion of study volunteers (27.7%) compared to that of non-rare disease indications (refer to Table 3).

Table 4 contains the results of the logistic regressions that were conducted. It contains the odds ratio that a protocol will be in the ‘HIGH’ deviation cohort, and the p-value of the individual design variable. 5 of the 9 variables tested were significant (endpoints, eligibility criteria, distinct procedures, countries, and sites) indicating that as each of these design variables increased, the odds ratio that a protocol would be in the ‘HIGH’ deviation cohort also increased. Total procedures, procedures per visit, total visits, and visits per month were not found to significantly affect this odds ratio.

Table 5 summarizes the correlations between protocol deviations and clinical trial outcomes. The incidence of protocol deviations was significantly, positively correlated with all clinical trial milestone durations except for study start-up and study close-out. The duration to complete all first patient visits (First Patient First Visit—Last Patient First Visit), the overall treatment duration (First Patient First Visit—Last Patient Last Visit), and the total clinical trial duration from protocol approval to database lock showed weak but significant positive correlations (0.283, 0.342, and 0.323, respectively) with the number of protocol deviations per protocol.

**Discussion**

The results of this study present baseline measures of the mean total number of protocol deviations per protocol by phase and by disease category. Each phase II and III protocol has a total of 75 and 119 protocol deviations, on average, involving nearly one-third of all patients enrolled in each clinical trial. Oncology clinical trials have the highest relative mean number of protocol deviations affecting more than 40% of patients enrolled in each trial. This finding is
consistent with other Tufts CSDD studies that have demonstrated the high relative burden for study volunteers participating in clinical trials targeting cancer-related illnesses [7].

A relationship between select protocol design complexity variables and the incidence of protocol deviations was observed. The number of endpoints and the number of procedures per visit were modestly positively associated with and predictive of, the incidence of deviations per protocol. Adjustments to these select design variables may play a role in preventing protocol deviations and in minimizing their risk to the study and data integrity.

Two operational variables were also positively associated with protocol deviations. The number of countries showed a modest relationship. A positive relationship was also shown between the number of investigative sites and the number of protocol deviations suggesting that decisions and behaviors that diverge from the investigational plan may be a function of variation in investigative site personnel training, site infrastructure, operating environment, and culture. Further examination is needed to understand this insight and its implications (Table 4).

It is not surprising that the incidence of protocol deviations was significantly and positively correlated with all clinical trial milestone durations except for study start-up and close-out. Protocol implementation burden is generally low during the study initiation process. This burden, however, for both study staff and study volunteers, becomes more pronounced as these parties comply with protocol implementation requirements over time.

Protocols with a ‘HIGH’ number of deviations above the mean generally had actual timelines closer to, or better than planned durations compared to protocols with a ‘LOW’ number of deviations (Table 6). Timelines for the ‘HIGH’ protocol deviations subgroup may have had their original planned timelines updated or revised leading up to the ‘actual’ start date [8].

Protocols with a ‘HIGH’ number of deviations were also more complex (e.g., a higher relative number of endpoints, procedures, countries, and investigative sites) and were expected to encounter delays. Because of this, the timelines may have been set with additional time built into the original plan. This would explain why the protocols in the ‘HIGH’ cohort often appeared to perform closer to, or better than, plan.

This study has several limitations. The data are based on a convenience sample of protocols arbitrarily selected by 20 participating companies. In addition, recognizing the challenge of gathering this data, the results derive from a

---

### Table 3

| Combined phase II/III protocols (n = 139) | Oncology | Non-oncology | Rare disease indications | Non-rare disease indications |
|-----------------------------------------|----------|--------------|------------------------|-----------------------------|
| Mean number of protocol deviations      | 108.8    | 91.9         | 78.1                   | 98.7                        |
| Proportion of patients with protocol deviations (%) | 46.6% | 27.4%        | 27.7%                  | 32.1%                       |

### Table 4

| Design variable                              | n     | Odds Ratio (95% CI) | p-value |
|----------------------------------------------|-------|---------------------|--------|
| Number of endpoints                          | 110   | 1.054 (1.015–1.094) | <.01   |
| Number of eligibility criteria               | 135   | 1.046 (1.011–1.083) | <.01   |
| Number of distinct procedures performed      | 134   | 1.096 (1.047–1.147) | <.001  |
| Total number of procedures performed         | 134   | 1.001 (1.000–1.003) | .147   |
| Number of procedures performed per visit     | 132   | 1.033 (0.989–1.079) | .142   |
| Number of countries                          | 135   | 1.191 (1.103–1.285) | <.001  |
| Number of investigative sites                | 130   | 1.036 (1.021–1.052) | <.001  |
| Number of planned volunteer visits           | 135   | 1.001 (0.975–1.028) | .928   |
| Patient visits per month                     | 119   | 0.974 (0.943–1.005) | .099   |

### Table 5

| Protocols (n = 187)                        | Coefficient | p-value |
|--------------------------------------------|-------------|---------|
| Study start-up duration                    | .153        | .090    |
| Duration to complete all first patient visits | .283        | <.01    |
| Treatment duration                         | .342        | <.001   |
| Study close-out duration                   | .056        | .554    |
| Total clinical trial duration              | .323        | <.001   |
| Drop-out rate                              | .125        | .159    |
| Substantial protocol amendments            | .141        | .117    |
relatively small sample of only 187 protocols and should be interpreted with some caution.

Future research will look at the root causes of protocol deviations and how to address them. Future research will also look at deviations with more granularity to understand the incidence and impact of minor and major deviations. Whereas some deviations may improve participation convenience and clinical trial executional feasibility, other deviations may pose substantially greater risk to study volunteer rights and safety and to the quality and integrity of the clinical trial data.

Conclusions

The results of this initial study provide useful measures that sponsor companies can begin to use to benchmark their own protocol deviation experience and determine whether remediation is warranted. Given anecdotal reports on the rapid growth of protocol deviations during the COVID-19 pandemic, when the majority of protocols transitioned to virtual and remote patient participation and data collection, this study also provides a valuable pre-pandemic baseline measure to evaluate ongoing and post-pandemic protocol deviation experience and its impact.

Acknowledgements

The authors wish to thank the companies that participated in this working group study: Amgen, AstraZeneca, Biogen, Boehringer Ingelheim, Bristol-Myers Squibb, CSL Behring, Eli Lilly, EMD Serono, GlaxoSmithKline, Janssen, Merck, Novartis, Otsuka, Parexel, Pfizer, Roche, Sanofi, Takeda, UCB, and Veristat. In addition, the authors thank Michael Wilkinson, formerly at Tufts CSDD, for his assistance on this project.

Author contributions

KG, Tufts CSDD, contributed to all four aspects (substantial contribution to conception, design, analysis, and interpretation; drafting and revising the work; final approval of the version to be published; and agreement to be accountable for all aspects in ensuring accuracy and integrity of the work). ZS, Tufts CSDD, contributed to all four aspects; AJ, Biogen, made substantial contribution to conception, design, analysis, and interpretation; RK, Merck & Co., made substantial contribution to conception, design, analysis, and interpretation and assisted in drafting and revising the work.

Funding

Tufts CSDD received grant funding from the participating working group companies to cover staff time on this study.

Declarations

Conflict of interest

Kenneth Getz, Tufts CSDD, has nothing to disclose. Zachary Smith, Tufts CSDD, has nothing to disclose. Ananya Jain, Biogen, declares that he is an employee and has financial holdings in the company. Randy Krauss, Merck & Co., declares that he is an employee and has financial holdings in the company.

References

1. Weed K. FDA Inspections: Findings for Clinical Investigators. IMARC Research, July 11, 2019. https://www.imarcresearch.com/blog/fda-inspections-findings-for-clinical-investigators
2. Ibid.
3. Mehran, Kurpanek K, Pewtrizzo M, et al. The life cycle and management of protocol deviations. TIRS. 2014;48(6):762–77.
4. Galuchie L, Stewart C, Meloni F. Protocol deviations: a holistic approach from defining to reporting. TIRS. 2021;55(4):733–42.
5. Sprunger K, Nickerson D, Meeker-O’Connell A, et al. Quality by design in clinical trials: a collaborative pilot with FDA. TIRS. 2013;47(2):161–6.
6. Meeker-O’Connell A, Sam I, Bergamot N, et al. TransCelerate’s clinical quality management system: issue management. TIRS. 2016;50(5):530–5.
7. Smith Z, Wilkinson M, Carney C, et al. Enhancing the measure of participation burden in protocol design to incorporate logistics, lifestyle and demographic characteristics. TIRS. 2021;55:1239–49.
8. Getz K, Stergiopoulos S, Short M, et al. The impact of protocol amendments on clinical trial performance and cost. TIRS. 2016;50:436–41.