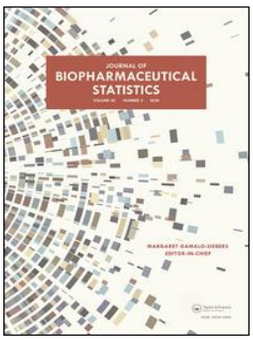


Journal of Biopharmaceutical Statistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lbps20

Enhancing clarity of clinical trial safety reports for data monitoring committees

Sonia M. Thomas , Kwanhye Jung , Hengrui Sun , Matthew A. Psioda , Pedro Miguel Quibrera & Stephen M. Strakowski

To cite this article: Sonia M. Thomas , Kwanhye Jung , Hengrui Sun , Matthew A. Psioda , Pedro Miguel Quibrera & Stephen M. Strakowski (2020): Enhancing clarity of clinical trial safety reports for data monitoring committees, Journal of Biopharmaceutical Statistics, DOI: 10.1080/10543406.2020.1815034

To link to this article: https://doi.org/10.1080/10543406.2020.1815034

Published online: 08 Sep 2020.

JOURNAL OF BIOPHARMACEUTICAL STATISTICS
https://doi.org/10.1080/10543406.2020.1815034

Taylor & Francis
Taylor & Francis Group

Enhancing clarity of clinical trial safety reports for data monitoring committees

Sonia M. Thomas[a], Kwanhye Jung[b], Hengrui Sun[c], Matthew A. Psioda[b], Pedro Miguel Quibrera[b], and Stephen M. Strakowski[d]

[a]Division of Biostatistics and Epidemiology, RTI International, Research Triangle Park, NC, USA; [b]Collaborative Studies Coordinating Center, Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA; [c]Division of Biometrics IV, Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA; [d]Department of Psychiatry, Dell Medical School, University of Texas, Austin, TX, USA

ABSTRACT

A Data Monitoring Committee (DMC) evaluates patient safety in a clinical trial of an investigational intervention through periodic review of adverse events (AEs) and clinical safety assessments. Our aim was to construct DMC report displays to enhance the DMC safety review through use of graphics and clear identification and adjustment for missing data caused by early discontinuations and ongoing study participation. Suggested displays include a study snapshot graph, enhanced adverse event incidence tables including the incidence density and plotted incidence proportions, line graphs in place of by-patient listings, and trend plots in place of tables for continuous assessments.

ARTICLE HISTORY
Received 21 August 2020
Accepted 21 August 2020

KEYWORDS
Data monitoring committee; data and safety monitoring board; graphical displays; adverse event; incidence density; patient drop-out; clinical trial; safety data

1. Introduction

A Data Monitoring Committee (DMC) also called Data and Safety Monitoring Board (DSMB) for a clinical trial is charged with unmasked periodic review of accumulating data regarding patient recruitment, study completion, adverse events/side effects and other safety assessments (e.g., laboratory, electrocardiogram, vital signs). DMCs are most often comprised of academic experts in the medical condition under study or known potential adverse events, clinical trials, biostatistics, and sometimes ethicists or patient representatives. Three types of DMCs are (1) those for an industry-sponsored study that will be part of a regulatory submission, (2) those for a multicenter research trial sponsored by National Institute of Health or similar organization, and (3) standing institutional DMCs designed to oversee smaller single-site trials within their institution. There are guidelines and recommendations for conduct and roles of each type of DMC (CTSA Collaborative DSMB Workgroup 2018; European Medicines Agency 2005; Office of Inspector General Department of Health and Human Services 2013; US Food and Drug Administration 2006). Committee members take on a substantial responsibility for the safety of participants in the study, those yet to be enrolled, and even those who might someday receive the intervention. Numerous challenges associated with this responsibility have been recently described, including the need for training for DMC members and clear DMC reports (Calis et al. 2017b, 2017a; Davis et al. 2018; DeMets and Ellenberg 2016; DeMets and Fleming 2004; Ellenberg and Ellenberg 2018; Ellenberg et al. 2002; Fleming et al. 2017, 2018; Lewis et al. 2016; Neaton et al. 2018). DMC reports are produced by an independent unmasked statistical team. The masked project team and independent

CONTACT Sonia M. Thomas Smthomas@rti.org Division of Biostatistics and Epidemiology, RTI International, Research Triangle Park, NC, USA

*Disclaimer: This paper reflects the views of the author and should not be construed to represent FDA's views or policies. Dr. Thomas (formerly Davis) and Dr. Sun were affiliated with the University of North Carolina during development of this research.

unmasked team together have responsibility to provide an accurate report to the DMC. While best practice for avoiding accidental unblinding of masked individuals through mis-delivery or misplacement of a report is to display treatment groups by codes (such as A and B), DMC members should be unmasked to the actual treatment assignment from the very beginning of the trial so they are able to effectively assess patient risk-benefit (US Food and Drug Administration 2006; Fleming et al. 2018; B. Davis et al. 2018; Meinert 1998).

Recommended content of a DMC report has been described (Harrell 2017a, 2017b; Neaton et al. 2018) For industry trials, it is common for the DMC report to comprise a subset of tables and listings planned for the final Clinical Study Report for submission to the US Food and Drug Administration (FDA). However, these displays do not adequately differentiate between patients who have discontinued the study or are still ongoing, and the comprehensive tables and listings designed for a regulatory submission are often too detailed and not helpful for a DMC, who have limited time to review the data.

Displays for a DMC must be efficient and easy to review, and graphics provide an excellent improvement. Proposed graphical displays for safety data have been published and open-source programming code is available for both static images and dynamic graphs with interactive options and drilldown capabilities (Amit et al. 2008; CTSpedia Clinical Trials Safety Graphics Home Page; Duke et al. 2015; Harrell 2014, 2016, 2017a, 2017b, 2018; Ma et al. 2015; PhUSE Computational Science Standard Analyses and Code Sharing Working Group Analysis and Display White Papers Project Team 2013; 2015, 2017; Pocock et al. 2007; Soukup 2011; Swihart et al. 2010; Torsvik et al. 2013; Wildfire et al. 2018; Zink et al. 2013), yet most DMC reports continue to rely solely on tables and listings. In this paper, we describe a set of static graphical displays that our DMC found both useful and easy to review. Our objective is to demonstrate the content of an effective graphics-based report to statisticians and DMC members in order to promote more widespread use of these and similar methods for DMC reports.

A complexity for DMC reports is that the data is a snapshot of participants who have completed the study, discontinued the study early, or are ongoing. Later study visits may have substantially less data available than earlier visits. The periodic DMC report thus has a mixture of data that is missing at random (MAR, Little and Rubin 2014) due to ongoing patients who have not yet attended the next scheduled visit plus data that is not missing at random (NMAR) due to early study discontinuation in one or more treatment arms. The DMC must be able to easily identify whether patients have differentially discontinued among treatment groups, discern how this discontinuation impacts other safety data, and interpret safety signals in spite of differential discontinuation.

We present consolidated suggestions for graphical DMC report displays that provide efficient transparency of safety signals, relative sample sizes, and appropriately account for the potential impact of different discontinuation rates between treatment groups.

2. Materials and methods

Our displays were developed by an independent unmasked statistical team for periodic DMC reports of nine Phase 2 or Phase 3 clinical trials. Multiple trials were ongoing concurrently, with different study durations, trial designs, and patient populations. The DMC, chaired by Dr. Strakowski and comprised of clinicians and biostatisticians, was consulted for their input and preferences for data displays in a planning meeting prior to the first report and in an ongoing fashion. Graphical safety data displays were planned based on published recommendations and were enhanced throughout the project based on feedback from the committee (Amit et al. 2008; CTSpedia Clinical Trials Safety Graphics Home Page; Harrell 2014, 2017a, 2017b; Soukup 2011; PhUSE Computational Science Development of Standard Scripts for Analysis and Programming Working Group Analysis and Display White Papers Project Team 2013; 2017; Wildfire et al. 2018). In addition to the graphical safety displays, the reports contained tables for participant disposition and demographics and supportive listings of serious adverse events.

2.1. Report displays

All graphs were produced using SAS/GRAPH and the Graph Template Language, version 9.4 of the SAS® System for Windows, as described by Jung (2015). Reports in Portable Document Format (PDF) were delivered to DMC members through secure electronic transmission. Supporting data listings of serious adverse events were provided in a separate PDF file. Figures in this manuscript are based on simulated data from a fictitious study.

2.2. Incidence density

Participants from two treatment groups may discontinue from a study at unequal rates due to varying levels of intolerability or lack of efficacy. In an active treatment arm with high levels of early discontinuation due to intolerability (such as constipation), it is possible for the incidence proportion (the number of people with an event divided by the number exposed) for a rare event that could occur at any time during exposure (such as a stroke) to be lower than the event probability during the study period because patients who discontinue early will have less exposure, and thus less opportunity to experience the event. Similarly, a placebo or ineffective treatment group with a high level of dropout due to lack of efficacy may demonstrate a reduced incidence proportion relative to the true event probability during the study period for events that are unrelated to the treatment (such as headache). In the first case, the active versus placebo difference in risk of the rare treatment-related event may be under-estimated by the incidence proportion, and in the second case, the active versus placebo risk difference of the event unrelated to treatment may be over-estimated. As an alternative measure, the incidence density, defined as the number of participants with an event divided by the total person-time of exposure to treatment, adjusts for differences in exposure time. The incidence density can be interpreted as the estimate of a constant survival hazard rate.

Person-time of exposure is calculated as the sum over individuals in a treatment group of the duration from study start until either study completion, early discontinuation, first occurrence of the event, or the DMC review database cutoff date (He et al. 2015; Koch et al. 1993; Liu et al. 2006; Tangen and Koch 2000). For example, for a 1-year study of N = 100 participants with no early discontinuations, if n = 20 people experience the event, then the incidence proportion is n/N = 20/100, or 20%. If the 20 events were equally distributed across the year, the average person-years of exposure until the event would be 0.5 years, and the incidence density is n/(sum of exposure time) = 20/(20*0.5 years + (100–20)*1 year) = 0.22 events/person-year of exposure. For a safety review, even when there is no differential exposure due to drop-out, it can be helpful to evaluate both the incidence density and the incidence proportion. For example, if constipation was present in two groups at an equal incidence but occurred much sooner in one group than another, then the incidence proportions will be the same but the incidence density in the arm occurring sooner will be higher. If the DMC were focused on a recurring AE instead of a rare event, an additional calculation of interest could be the prevalence rate: total number of events divided by the total exposure time.

We performed a simulation study to demonstrate the potential impact of differential participant discontinuation on adverse event incidence proportions. We simulated a single iteration of a clinical trial with two treatment arms each with 100 participants and with a 1-year study duration. In our simulation, arm A had a 5% study discontinuation rate, and arm B had a 25% rate. Consistent with the common trend of higher rates of study discontinuation early in a clinical trial, we simulated 50% of the discontinuations to occur uniformly over the first quarter of the study, 33% to occur uniformly over the 2nd quarter, and 17% to occur uniformly over the last half of the study. We assumed that subjects could experience one adverse event with equal probability and simulated the time-to-onset of the event to be uniformly distributed over the study period. Time to study discontinuation and time to adverse event were simulated independently, so that a participant's AE might occur after their study discontinuation, representing an event that might have happened but would not be observed due to early study discontinuation. We varied the probability of event during the year for both treatment

arms from 5% to 50% and calculated the incidence proportion and the incidence density for the simulated observed events that occurred before study discontinuation.

3. Results

Safety DMC reports generally present the following types of data by treatment group: participant disposition and treatment exposure, demographics, adverse event rates, details about individual adverse events of special interest or concern, and both trends over time and incidence of abnormal values for laboratory results, electrocardiograms, vital signs, or other safety assessments. Graphical displays are demonstrated for each type of data except for demographics which is well displayed in a table.

3.1. Patient disposition and treatment exposure

Since periodic DMC reports present accumulating safety data, at each report the number and percent of participants that have completed the study or completed each visit within the study changes. It is helpful for a DMC report to start with a display typical for a ***CONSORT diagram***, including counts and percentage of discontinuations by reason for discontinuation (Moher et al. 2001). In addition, we found it very helpful to present a stacked bar ***study snapshot graph*** of the number of participants that were randomized, attended each visit, discontinued the study after each visit, and total completed or discontinued from the study, as in Figure 1. This display clearly shows the DMC the pattern of discontinuation by treatment group and how much data are available at later study visits relative to the number randomized. For example, in Figure 1, it is easy to discern that (a) Arm B has a higher discontinuation rate than Arm A after each visit, (b) approximately half of randomized patients are either completed or discontinued, (c) later study visits have a much smaller sample size than earlier visits, and (d) Arm A has more data than Arm B at later visits because of the higher early discontinuation rate of Arm B. In addition, a ***Kaplan-Meier plot*** of time to study discontinuation by treatment

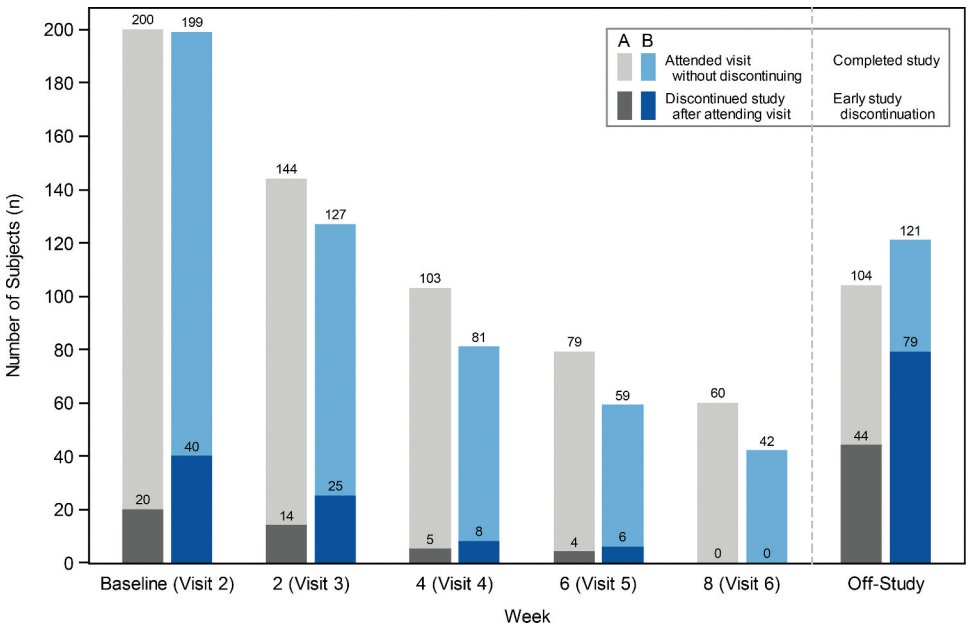


Figure 1. Study snapshot graph. Sample size is graphed by treatment group for each visit and for participants who have completed or discontinued the study. Participants who discontinued early are depicted by shading at their last attended visit. Simulated data from a fictitious study.

group provides the DMC a clear and informative group comparison of study discontinuation over time (S. Davis et al. 2011). A further refinement of the study snapshot graph can display missed visits or the reasons for discontinuation with different types of shading. Finally, if the study has dose titration or allows variable dosing, in which each participant is titrated until they achieve satisfactory efficacy with manageable side effects, the pattern of dose titration over visits per treatment group can be shown in a separate stacked bar graph, with the percentages of participants receiving each dose level at each visit indicated by different types of shading.

3.2. Adverse event rates

One of the most important components of the DMC report is the presentation of adverse events, including rates of any AE, deaths, serious adverse events (SAEs), and AEs/SAEs grouped by Medical Dictionary for Regulatory Activities (MedDRA, Mozzicato 2009) System-Organ Class (such as gastrointestinal disorders) and Preferred Term (such as nausea).

Tables of AE incidence proportions can be many pages long, and DMC members often express frustration with the challenge of identifying trends in treatment group differences when provided AE incidence tables. A solution often requested or offered is to add a column of p-values for each row of the table comparing the treatment group proportions with a chi-square or Fisher's exact test. The p-values are commonly used by reviewers as a descriptive aid to scan down the page and look more closely at rows with smaller p-values. While p-values are applicable to pre-specified comparisons and can be a helpful descriptive summary, they are not an appropriate tool for DMC members to identify safety trends, since a p-value may not highlight an important difference in a report with small sample sizes and/or small event rates, or conversely may overly highlight small differences in larger studies. Further, one should expect 1 in 20 rows of an AE table comparing identical treatments to show a p-value < 0.05 due to random chance. An alternative strategy which our DMC members found highly effective is the addition of a plot of the AE proportions per treatment group on each row of an AE table (Harrell 2014, 2016, 2017a; PhUSE Computational Science Standard Analyses and Code Sharing Working Group Analysis and Display White Papers Project Team 2017; Wildfire et al. 2018).The ***incidence table with plotted proportions*** is shown in Figure 2. The column of plotted incidence proportions allows the DMC reviewer to quickly identify trends, particularly when the display is sorted by prevalence of the AE term. Further, uncertainty of the incidence proportion can be displayed by adding a confidence interval around each point, or by adding a second plot showing the difference or ratio of proportions with corresponding confidence interval (Harrell 2014, 2017a; Wildfire et al. 2018).

Another problem with traditional AE tables is the exclusive reliance on the AE incidence proportion, which provides the proportion of all patients experiencing an event but does not account for differential levels of study exposure between treatment arms. Results of our demonstration simulation comparing study arms with a 5% versus a 25% early discontinuation rate are shown in Figure 3. The incidence density (events per person-years of exposure) is shown to be equal between the groups for any AE event probability, while the incidence proportion is lower for the treatment arm with a 25% discontinuation rate compared to the arm with a 5% discontinuation rate, and this difference increases with the underlying probability of event. For example, at an exaggerated event probability of 50% during the study period, the arm with 5% early discontinuation demonstrated an incidence proportion of approximately 50%, whereas the arm with 25% early discontinuation demonstrated a much lower incidence proportion of approximately 42%.

Adding a column for the incidence density to AE displays allows for the comparison between treatment groups in situations where the discontinuation rates vary between groups. For example, in Figure 2, preferred term "1", has the same incidence proportion in the two groups but a somewhat smaller incidence density in Group B compared to Group A, indicative of less exposure and a greater rate of dropout in Group B, as identified in Figure 1. Also, perhaps more importantly, incidence

System Organ Class / *Preferred Term*	*A* (N=200) n (%)	*A* (N=200) *Incidence density**	*B* (N=199) n (%)	*B* (N=199) *Incidence density**
Subjects with any Treatment-Emergent Adverse Event	100 (50)	6.3	150 (75)	17.3
System Organ Class 1	61 (31)	2.7	85 (43)	4.7
Preferred Term 1	38 (19)	1.4	36 (18)	1.1
Preferred Term 2	15 (8)	0.5	36 (18)	1.3
Preferred Term 3	12 (6)	0.4	25 (13)	0.8
Preferred Term 4	9 (5)	0.3	17 (9)	0.5
Preferred Term 5	5 (3)	0.2	8 (4)	0.2
System Organ Class 2	38 (19)	1.4	100 (50)	6.1
Preferred Term 6	17 (9)	0.6	60 (30)	2.4
Preferred Term 7	13 (7)	0.4	50 (25)	1.9
Preferred Term 8	10 (5)	0.3	37 (19)	1.3
Preferred Term 9	5 (3)	0.2	19 (10)	0.6

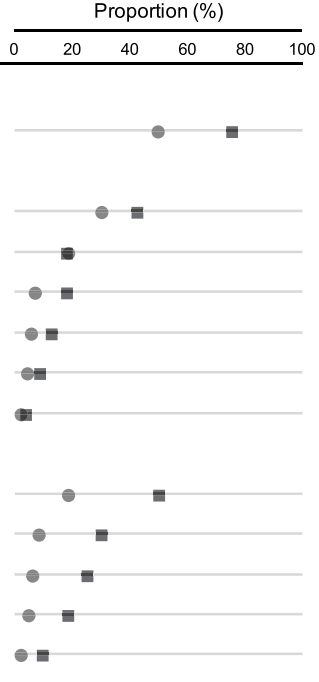


Figure 2. Adverse event table with incidence density and plotted proportions. The incidence density (events per person-time of exposure) is shown in addition to the incidence proportion. The proportion of each event is graphed. Simulated data from a fictitious study.

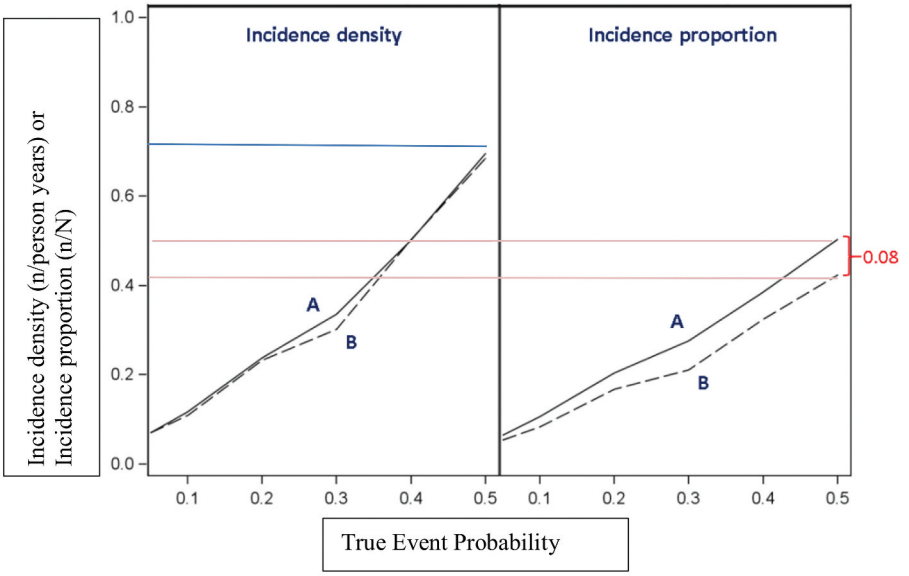


Figure 3. Simulated example showing impact of differential discontinuation on adverse event incidence density and incidence proportion. Discontinuation rate is 5% for group A and 25% for group B. N = 100 patients per arm, and study duration = 1 year. The true AE event probability is uniform. Incidence density is similar across treatment groups, while the observed incidence proportion is under-represented in the group with higher discontinuation rate. On the left of the graph, we see that as the true event probability increases, the incidence density (events per person-years of exposure) increases equally across the 2 treatment groups. When the event probability is 50%, both groups regardless of discontinuation rate have approximately 0.75 events per person-year of exposure. On the right of the graph, the incidence proportion is approximately 0.5 for Arm A, (with 5% early discontinuation), but is substantially lower (approximately 0.42) for Arm B (with 25% early discontinuation).

densities allow the DMC to compare rates of rare AEs across studies with widely different durations of treatment exposure.

3.3. Individual adverse events

In addition to overall incidence of adverse events, DMCs often need to review details of specific AEs of special interest or concern, including the time course and severity of the event, whether the participant discontinued the study because of the event, other AEs experienced by the participant, and to evaluate similarities or differences of rare events across cases. Graphical displays supplemented by back-up listings offer improvements over traditional listings alone since they help the DMC evaluate patterns of events within and between patients in ways that textual displays cannot. An overall summary of individual occurrences of specific AEs of interest across all patients in a small- to moderate-sized study can be visualized through a ***pyramid plot*** (Figure 4). Treatment groups are displayed side by side. The X axis represents duration in the study, and each participant is depicted by a single line, the length of which represents the duration of their study participation. Participants who discontinued the study are represented by one colored line (such as red) while those who are ongoing or completed are another color (such as black). Along the line, each event is indicated by a dot, the diameter of which indicates its severity (mild, moderate or severe). AEs directly at the end of a red line indicates that the patient discontinued directly after the event. Two or more different types of event can be displayed using different colored dots. More than two treatment groups can be displayed with side-by-side pyramids or half pyramids all aligned in the same direction. Figure 4 shows that (1) adverse events represented by Preferred Term "1" tend to occur earlier in the study, but at

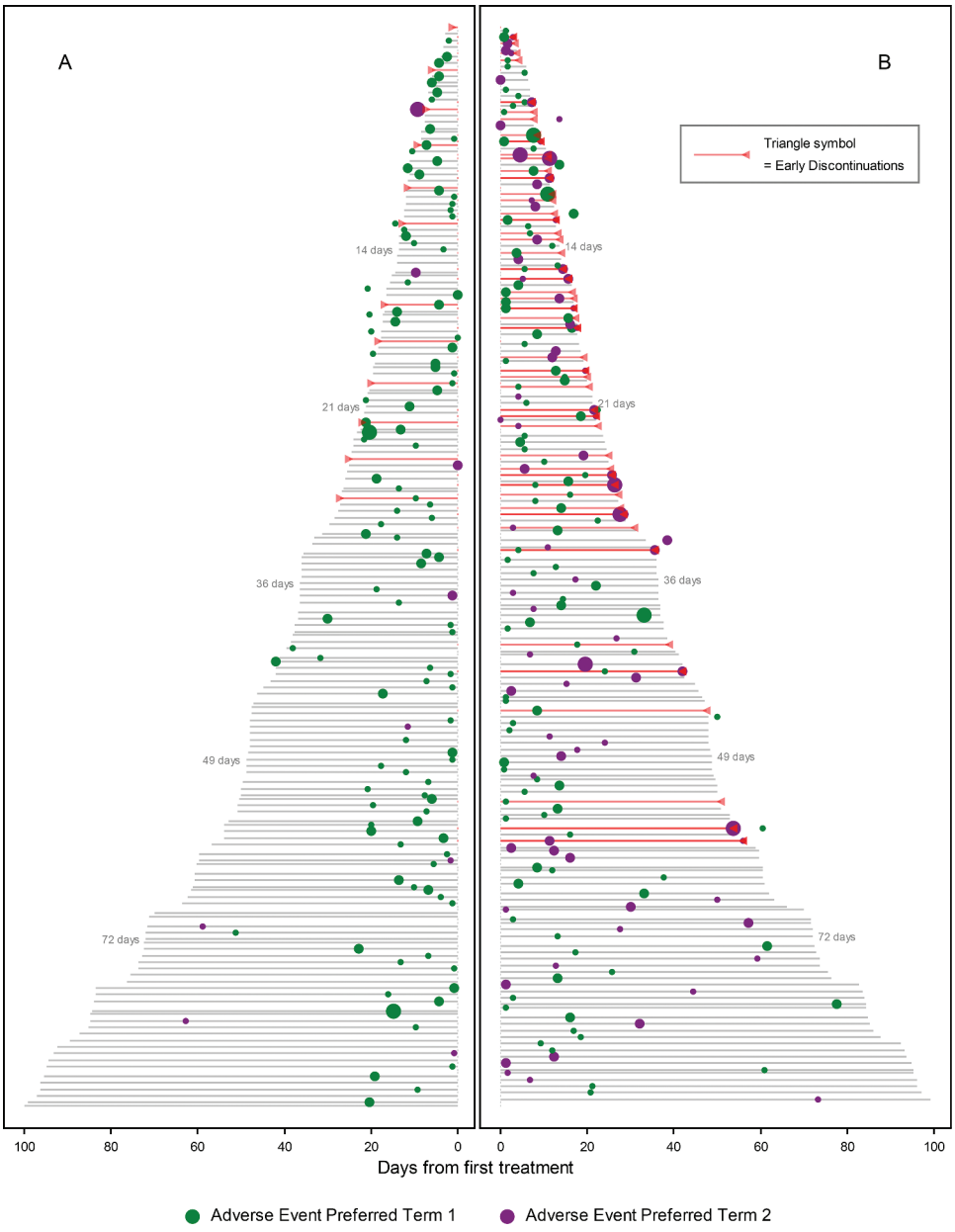


Figure 4. Pyramid plot of selected adverse events and study duration for each participant. Study duration for each participant per treatment group is sorted and displayed as a timeline. A red timeline ending with a symbol indicates early study discontinuation. Dots represent adverse event start dates. Size of the dot represents adverse event severity (mild, moderate, severe, where severe is the largest dot), color of the dot represents adverse event term. Simulated data from a fictitious study.

similar rates for both treatment groups, and rarely occur directly before an early study discontinuation, identified by a triangle at the end of a red line; (2) Arm B has more early study discontinuations (more patients identified by a red line ending in a triangle) than Arm A; and (3), Arm B has more adverse events represented by Preferred Term "2" than Arm A, and several of these adverse events occur shortly before early study discontinuation.

To evaluate individual cases in detail, a ***profile line graph*** provides most of the information contained in a traditional AE listing, yet with visual cues allowing efficient evaluation and comparison

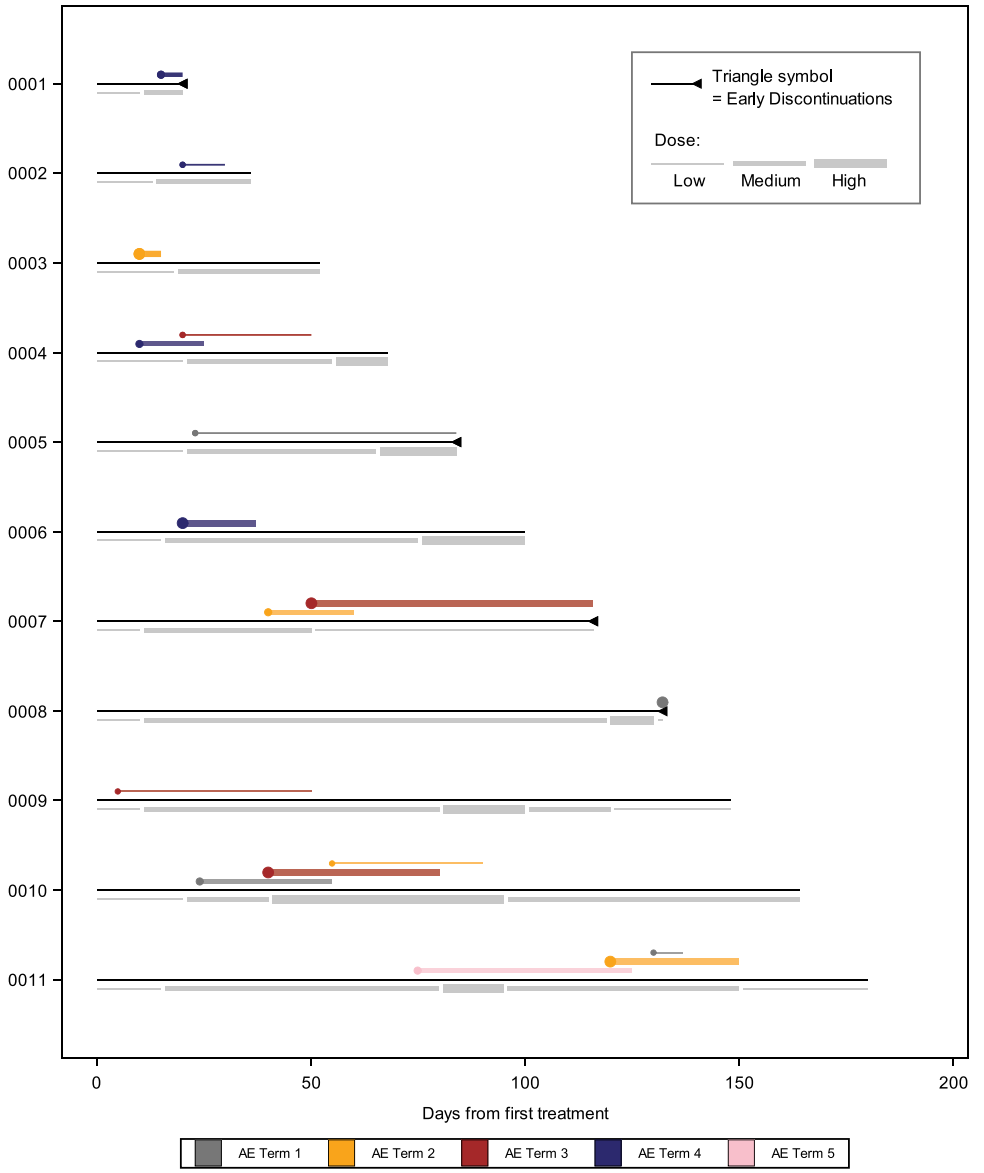


Figure 5. Adverse event by-participant profile line graph. Specific adverse events (AEs) are depicted by a colored line, plotted above the participant's study duration timeline. Symbol at the end of the timeline indicates early study discontinuation. Dots represent adverse event start dates and lines represent adverse event duration. Size of the dot and thickness of the line indicates the severity of the adverse event (mild, moderate, severe, where severe is the largest dot and thickest line). Color indicates the adverse event term. For titration or variable dose studies, dose is indicated by shading underneath the study duration timeline. Treatment groups can be displayed in side-by-side graphs. Simulated data from a fictitious study.

of events not obtainable from a traditional listing. Figure 5 depicts a profile line graph showing AE start relative to treatment start, AE duration, and treatment stop (Amit et al. 2008). Early study discontinuation is differentiated from ongoing patients by a symbol at the end of the treatment duration line. AE severity (mild, moderate, severe) is indicated by the size of the dot at AE start and thickness of the AE line, with severe AEs having the largest dot and thickest line. For variable dose and titration studies, dose of study drug is indicated by thickness of a secondary line, easily identifying whether the AE occurred after a dose escalation, whether the dose was reduced after the AE started,

and whether the dose was subsequently increased. Numerous AEs are plotted one below the other for each participant to show the temporal relationship of overlapping events. Lines for multiple participants fit on the same page of the graph, and graphs for each treatment group can be displayed either separately or side-by-side to assist the DMC in comparing the number and characteristics of individual events between treatment groups.

3.4. Continuous safety parameters

A comprehensive DMC evaluation of the change from baseline in safety assessments such as blood pressure, clinical laboratory values or electrocardiogram parameters includes (1) trends over time for all participants by treatment group, (2) incidence of treatment-emergent abnormal values by treatment group, and (3) trajectories of values over time for participants with at least one abnormal occurrence. Although regulatory study reports for biopharmaceutical products may contain shift tables of normal/abnormal status from baseline to post-baseline visits, our experience is that DMCs uniformly do not find shift tables interpretable and so such tables may be of little use for DMC reports.

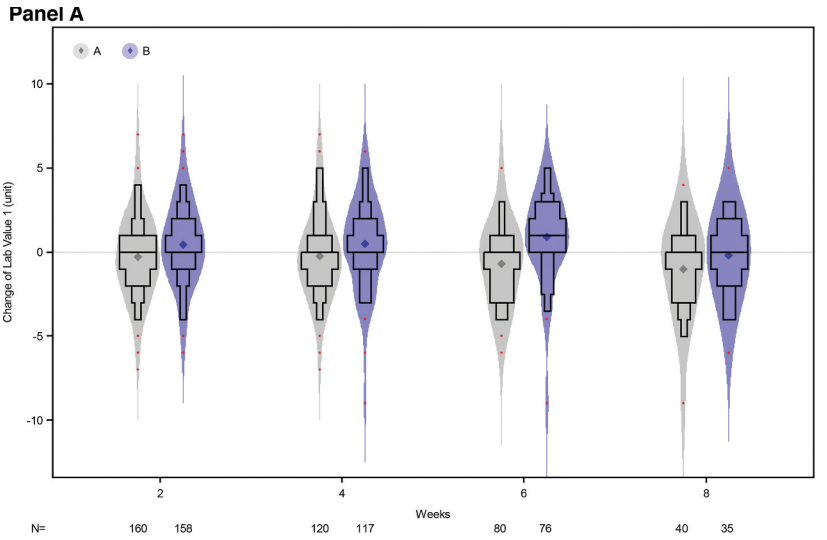
For evaluation of trends over time, a by-visit side-by-side box-plot is substantially enhanced by depicting the density estimate overlaid by additional percentiles of the distribution in a ***violin plot***, in Figure 6, panel A, as recommended by Harrell (2014, 2017a). For a DMC report, plots should be provided in place of traditional tables of descriptive statistics, rather than in addition to them. A DMC may prefer to see plots of the change from baseline to easily compare the mean change to zero on the X axis, or to evaluate lab data on the observed scale, including at baseline. Decreasing sample size at later study visits can also be visually displayed by decreasing the intensity of the graph proportional to the sample size (Harrell 2017a).

For some other continuous assessments, such as total scores from patient-reported scales, the DMC might be concerned only with the mean changes over time rather than the full distribution. In such cases, trends over time by treatment group could instead be presented using a segmented line plot of mean change from baseline with 95% confidence intervals at each visit, which is a familiar and easy to interpret display often used to depict longitudinal efficacy data. Whenever possible, for either display format, the Y axis for a parameter should remain the same across all displays within a report, across subsequent DMC reports, and between concurrent studies, allowing for direct visual comparisons.

Rates of treatment-emergent abnormal vitals, blood laboratory results, or electrocardiogram parameters can be effectively displayed for the DMC using the ***incidence table with plotted proportions*** as shown in Figure 6 panel B, which matches the style of the graphical AE incidence table. Individual trajectories of a parameter for participants with an abnormal value can be efficiently reviewed by DMCs using a segmented line ***spaghetti plot*** of values over time as in Figure 6 panel C (Amit et al. 2008). Horizontal lines on the graph identify the abnormality thresholds, and only participants with an abnormal value are plotted. Participants who discontinue the study are differentiated from ongoing participants by a symbol plotted at their last assessment point. Spaghetti plots efficiently show the DMC whether participants with an abnormal value have been discontinued from the study, whether there is a pattern for continued worsening of the laboratory value over time, or whether participants return to within normal limits at the following assessment. Such patterns are not easily discerned from traditional by-patient listings.

4. Discussion

We have demonstrated some Data Monitoring Committee Report graphical displays for participant disposition, adverse event and other safety data which aim to clearly and efficiently display safety signals and early study discontinuation across study visits and between treatment groups. Most of the figures can easily accommodate three or more treatment groups. We also described the benefits of providing the incidence density in addition to the incidence proportion for evaluating occurrence rate of adverse events. Incidence densities not only adjust for differential discontinuation between



Panel B

		A		B		Proportion (%)
Parameter		*N*	*n (%)*	*N*	*n (%)*	0 5 10 15 20 25
Lab Parameter 1	Low (xxx units)	159	4 (2.5)	156	3 (1.9)	
	High (xxx units)	159	2 (1.3)	156	1 (0.6)	
Lab Parameter 2	Low (xxx units)	160	3 (1.9)	158	4 (2.5)	
	High (xxx units)	160	1 (0.6)	158	2 (1.3)	
Lab Parameter 3	Low (xxx units)	160	2 (1.3)	158	0	
	High (xxx units)	160	3 (1.9)	158	2 (1.3)	

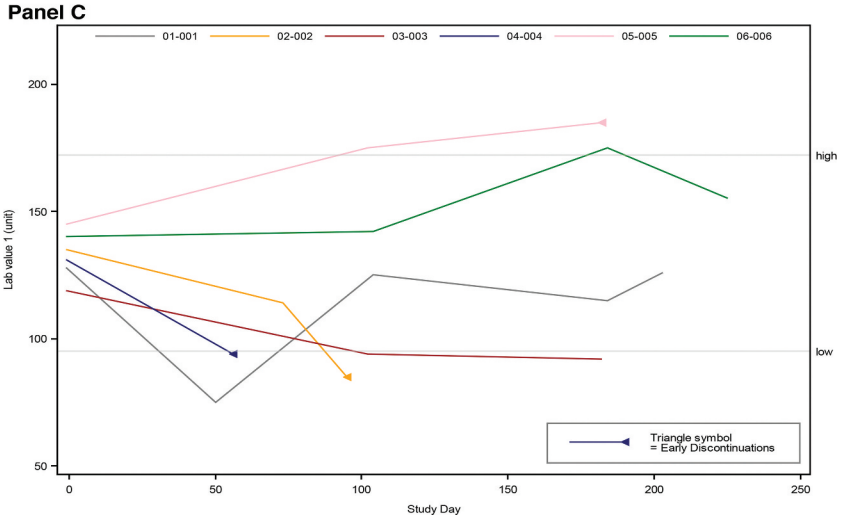


Figure 6. Continuous laboratory parameter trend plot, incidence table, and line graph of participants with abnormal values. Panel A: Side-by-side Violin plots of change from baseline. The smoothed distribution is represented by the shaded area. An extended box plot is overlaid, with steps representing the 5%, 12.5%, 25%, 37.5%, 50%, 62.5% 75%, 87.5%, 95% percentiles, and with outliers represented by a dot. The mean is represented by the diamond. Panel B: Incidence table with plotted proportions displays abnormal laboratory values in a similar format as adverse event displays. Panel C: Spaghetti plot for participants with abnormal values. A symbol at the end of the final segment for a participant indicates early study discontinuation, while a line ending without a symbol indicates an ongoing participant. Normal reference ranges are shown by horizontal lines. Treatment groups can be displayed for comparison purposes in side-by-side figures with identical axes. Simulated data from a fictitious study.

groups in a study, but they also support the comparison of event rates between studies of varying durations.

These displays were generated and refined as part of a DMC for a series of Phase 2 and Phase 3 clinical trials of a pharmaceutical agent in which reports for multiple studies were reviewed concurrently. By replacing many of the commonly produced tables and listings with graphical displays, we kept the DMC reports covering a complete set of safety data (including labs, electrocardiograms, vital signs and patient reported scales) for a typical moderate-sized Phase 3 pharmaceutical study to well under 150 pages. Supportive adverse event listings were provided in a separate electronic document. The displays allowed efficient review of safety data by the DMC and were enthusiastically received. Once initially programmed, graphs were produced by the independent data analysis center as efficiently as traditional tables.

Limitations: Displays of efficacy data and data quality metrics (reported for Institute-funded studies) have not been discussed, although some of the demonstrated graphs can be applied to them. The gain in efficiency and quality of DMC review from graphical displays might be less applicable when applied to a very small study with less data quantity, such as a phase 1 study, yet even in the small sample setting, visual demonstration of distribution variability is very helpful for the DMC in their review. Conversely, the by-patient graphical displays such as the AE pyramid plot, AE profile line graph and clinical laboratory spaghetti plot may require modification or may not be appropriate for very large studies with thousands of participants. Further, a graph that efficiency displays recurrence of adverse events would be of benefit to the DMC yet is not addressed here, other than the AE pyramid plot which can demonstrate recurrence of a small number of selected AEs.

A barrier to implementing graphical displays may be that figures have traditionally been more time intensive to prepare than tables and listings and limited by statistical software functionality. We found that graphing functionality with SAS software was more customizable with SAS Graph Template Language, which was used for several of the displays in this project. R graphics applications could also be used. The graphs discussed here are static images and could be further enhanced using interactive functionality which would allow DMC members to explore data trends or drill down to individual data.

A challenge we found is that DMC reviewers may have visual impairment or be color blind. To maximize reviewability, we used contrasting graph colors, made axis labels and text components of figures as large as possible, and tested the resolution quality of the final report document to ensure that figures remained clear when reviewers increase the display size through projection or magnification. Nine percent of men are red/green color deficient, and we learned that red/green color schemes for graphics should be avoided to facilitate review by color deficient individuals, but that blue/red color schemes were discernable.

The recommendations presented here and in other publications of safety data graphics should be further assessed and refined by statistical teams and data monitoring committees, so that benefits of graphics-based displays become more widely adopted in Data Monitoring Committee reports.

Acknowledgments

We wish to thank the members of the Data Monitoring Committee for a Phase II/III clinical trial program of a pharmaceutical compound sponsored by Sunovion Pharmaceuticals, Inc. for their insights, recommendations, and feedback. The committee (chaired by Stephen Strakowski) included Alexandru (Alex) Costea, MD, University of Cincinnati, Cincinnati Ohio; Oren Mason, MD, Grand Rapids Michigan; Jeffrey Newcorn, MD, Mt Sinai Hospital, New York, NY; Paul J. Rathouz, PhD, University of Texas at Austin, Austin, TX; and Lee-Jen Wei, PhD, Harvard University, Boston, MA. We also thank Sunovion Pharmaceuticals, Inc, and in particular Jane Xu, PhD, Head of Global Data Science, for endorsing and encouraging our use of innovative displays.

Disclosure statement

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Partially supported by a contract from Sunovion, Inc.

ORCID

Sonia M. Thomas http://orcid.org/0000-0003-1362-5627

References

Amit, O., R. Heiberger, and P. Lane. 2008. Graphical approaches to the analysis of safety data from clinical trials. *Pharmaceutical Statistics* 7 (1):20–35. doi:10.1002/pst.254.

Calis, K. A., P. Archdeacon, R. P. Bain, D. DeMets, M. Donohue, M. K. Elzarrad, A. Forrest, J. McEachern, M. J. Pencina, J. Perlmutter, et al. 2017a. Recommendations for data monitoring committees from the clinical trials transformation initiative. *Clinical Trials* 14 (4):342–348. doi:10.1177/1740774517707743.

Calis, K. A., P. Archdeacon, R. P. Bain, A. Forrest, J. Perlmutter, and D. DeMets. 2017b. Understanding the functions and operations of data monitoring committees: Survey and focus group findings. *Clinical Trials: Journal of the Society for Clinical Trials* 14 (1):59–66. doi:10.1177/1740774516679665.

CTSA Collaborative DSMB Workgroup. 2018. DSMB training manual. Medford, MA: Tufts Digital Library. May. Accessed September 03, 2018. https://www.tuftsctsi.org/research-services/regulatory/data-and-safety-monitoring-board-training-manual-for-investigator-initiated-studies/

CTSpedia Clinical Trials Safety Graphics Home Page. Accessed September 03, 2018. https://www.ctspedia.org/do/view/CTSpedia/StatGraphHome

Davis, B., D. Kerr, M. Maguire, C. Sanders, S. Snappon, and J. Wittes. 2018. University of Pennsylvania 10th annual conference on statistical issues in clinical trials: Current issues regarding data and safety monitoring committees in clinical trials (morning panel session). *Clinical Trials* 15:335–351. doi:10.1177/1740774518780434.

Davis, S. M., T. S. Stroup, G. G. Koch, C. E. Davis, R. A. Rosenheck, and J. A. Lieberman. 2011. Time to all-cause treatment discontinuation as the primary outcome in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia study. *Statistics in Biopharmaceutical Research* 3 (2):253–265. doi:10.1198/sbr.2011.10013.

DeMets, D. L., and S. S. Ellenberg. 2016. Data monitoring committees—expect the unexpected. *New England Journal of Medicine* 375:1365–1371. doi:10.1056/NEJMra1510066.

DeMets, D. L., and T. R. Fleming. 2004. The independent statistician for data monitoring committees. *Statistics in Medicine* 23 (10):1513–1517. doi:10.1002/sim.1786.

Duke, S. P., F. Bancken, B. Crowe, M. Soukup, T. Botis, and R. Forshee. 2015. Seeing is believing: Good graphic design principles for medical research. *Statistics in Medicine* 34:3040–3059. doi:10.1002/sim.6549.

Ellenberg, S. S., and J. H. Ellenberg. 2018. Proceedings of the University of Pennsylvania 10th annual conference on statistical issues in clinical trials: Current issues regarding data and safety monitoring committees in clinical trials. *Clinical Trials* 15 (4):319–320. doi:10.1177/1740774518781817.

Ellenberg, S. S., T. R. Fleming, and D. L. DeMets. 2002. *Data monitoring committees in clinical trials*. Hoboken, NJ: John Wiley & Sons.

European Medicines Agency. 2005. Committee for medicinal products for human use: Guideline on data monitoring committees. July 27. Accessed September 03, 2018. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003635.pdf

Fleming, T. R., D. L. DeMets, and M. Roe. 2017. Data monitoring committees: Promoting best practices to address emerging challenges. *Clinical Trials: Journal of the Society for Clinical Trials* 14:115–123. doi:10.1177/1740774516688915.

Fleming, T. R., S. S. Ellenberg, and D. L. DeMets. 2018. Data monitoring committees: Current issues. *Clinical Trials* 15 (4):321–328. doi:10.1177/1740774518764855.

Harrell, F. E. 2014. Principles of graph construction. Vanderbilt University School of Medicine, Department of Biostatistics. Accessed April 21, 2019. http://biostat.mc.vanderbilt.edu/wiki/pub/Main/StatGraphCourse/graphscourse.pdf

Harrell, F. E. 2016. Greport package: Graphical reporting for clinical trials. December 19. Accessed September 03, 2018. http://biostat.mc.vanderbilt.edu/wiki/Main/Greport

Harrell, F. E. 2017a. Example closed meeting data monitoring committee report. April 6. Accessed September 03, 2018. https://biostatdata.app.vumc.org/fh/talks/RCTGraphics/greportEx2.pdf

Harrell, F. E. 2017b. DMSB report for EXAMPLE trial. April 22. Accessed September 03, 2018. https://biostatdata.app.vumc.org/fh/talks/RCTGraphics/greportEx1.pdf

Harrell, F. E. 2018. Graphics for clinical trials. September 18. Accessed October 03, 2018. http://biostat.mc.vanderbilt.edu/wiki/Main/RCTGraphics

He, X., L. Chen, L. Lei, H. A. Xia, and M. T. Lee. 2015. A simple method for estimating confidence intervals for exposure adjusted incidence rate and its applications to clinical trials. *Journal of Biometrics & Biostatistics* 6:238. doi:10.4172/2155-6180.1000238.

Jung, K. 2015. Graphic displays of adverse events using the SAS graph template language. Presented at PhUSE single Day Event: Visualizing Clinical Data, A Picture is Worth a Thousand Words, Chapel Hill, NC, June 4. Accessed September 03, 2018. https://www.phusewiki.org/docs/2015_ChapelHill_SDE/2.PhUSE_Jung_Final.pdf

Koch, G. G., J. E. Schmid, J. M. Begun, and W. C. Maier. 1993. Meta-analysis of drug safety data. *Drug Safety Assessment in Clinical Trials*, 279–304. New York: Gilbert GS Marcel Dekker.

Lewis, R. J., K. A. Calis, and D. L. DeMets. 2016. Enhancing the scientific integrity and safety of clinical trials: Recommendations for data monitoring committees. *JAMA* 316 (22):2359–2360. doi:10.1001/jama.2016.16070.

Little, R. A., and D. M. Rubin. 2014. *Statistical analysis with missing data*, 11–13. 2nd ed. New York, NY: John Wiley and sons.

Liu, G. F., J. Wang, K. Liu, and D. B. Snavely. 2006. Confidence intervals for an exposure adjusted incidence rate difference with applications to clinical trials. *Statistics in Medicine* 25 (8):1275–1286. doi:10.1002/sim.2335.

Ma, H., C. Ke, Q. Jiang, and S. Snappin. 2015. Statistical considerations on the evaluation of imbalances of adverse events in randomized clinical trials. *Therapeutic Innovation & Regulatory Science* 49 (6):957–965. doi:10.1177/2168479015587363.

Meinert, C. L. 1998. Masked monitoring in clinical trials – Blind stupidity? *New England Journal of Medicine* 338:1381–1382. doi:10.1056/NEJM199805073381911.

Moher, D., K. F. Schulz, and D. Altman; for the CONSORT Group. 2001. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 285 (15):1987–1991. doi:10.1001/jama.285.15.1987.

Mozzicato, P. 2009. MedDRA: An overview of the medical dictionary for regulatory activities. *Pharmaceutical Medicine* 23 (2):65–75. doi:10.1007/BF03256752.

Neaton, J. D., B. Grund, and D. Wentworth. 2018. How to construct an optimal interim report: What the data monitoring committee does and doesn't need to know. *Clinical Trials* 15 (4):359–365. doi:10.1177/1740774518764449.

Office of Inspector General Department of Health and Human Services. 2013. Data and safety monitoring boards in NIH clinical trials: Meeting guidance, but facing some issues (OEI-12-11-00070). June. Accessed September 03, 2018. https://oig.hhs.gov/oei/reports/oei-12-11-00070.asp

PhUSE Computational Science Development of Standard Scripts for Analysis and Programming Working Group Analysis and Display White Papers Project Team. 2013. Accessed September 03, 2018. Analyses and displays associated with measures of central tendency – focus on vital sign, electrocardiogram, and laboratory analyte measurements in phase 2-4 clinical trials and integrated submission documents. https://www.phuse.eu/documents//working-groups/deliverables/analyses-displays-associated-with-measures-of-central-tendancy-version-10-10-oct-13-11816.pdf

PhUSE Computational Science Development of Standard Scripts for Analysis and Programming Working Group Analysis and Display White Papers Project Team. 2015. Analyses and displays associated with outliers or shifts from normal to abnormal: Focus on vital signs, electrocardiogram, and laboratory analyte measurements in phase 2-4 clinical trials and integrated summary documents. Accessed September 03, 2018. https://www.phuse.eu/documents//working-groups/deliverables/analyses-displays-associated-with-outliers-or-shifts-from-normal-to-abnormal-version-10-10-sept-15-11815.pdf

PhUSE Computational Science Standard Analyses and Code Sharing Working Group Analysis and Display White Papers Project Team. 2017. Analysis and displays associated with adverse events: Focus on adverse events in phase 2-4 clinical trials and integrated summary documents. Accessed September 03, 2018. https://www.phuse.eu/documents//working-groups/deliverables/adverse-events-white-paper-version-10-03-feb-17-11812.pdf

Pocock, S. J., T. G. Travison, and L. M. Wruck. 2007. Figures in clinical trial reports: Current practice & scope for improvement. *Trials* 8 (1):36–53. doi:10.1186/1745-6215-8-36.

Soukup, M. 2011. FDA/industry/academia safety graphic working group charter. Accessed April 23, 2019. https://www.ctspedia.org/do/view/CTSpedia/SafetyGraphicsCharter

Swihart, B. J., B. Caffo, B. D. James, M. Strand, B. S. Schwartz, and N. M. Punjabi. 2010. Lasagna plots: A saucy alternative to spaghetti plots. *Epidemiology* 21 (5):621–625. doi:10.1097/EDE.0b013e3181e5b06a.

Tangen, C. M., and G. G. Koch. 2000. Non-parametric covariance methods for incidence density analyses of time-to-event data from a randomized clinical trial and their complementary roles to proportional hazards regression.

Statistics in Medicine 19 (8):1039–1058. doi:10.1002/(SICI)1097-0258(20000430)19:8<1039::AID-SIM417>3.0.CO;2-O.

Torsvik, T., B. Lillebo, and G. Mikkelsen. 2013. Presentation of clinical laboratory results: An experimental comparison of four visualization techniques. *Journal of the American Medical Informatics Association20* 20 (2):325–331. doi:10.1136/amiajnl-2012-001147.

US Food and Drug Administration. 2006. Guidance for clinical trial sponsors. Establishment and operation of clinical trial data monitoring committees. Rockville, MD: FDA. Accessed September 03, 2018. https://www.fda.gov/downloads/regulatoryinformation/guidances/ucm127073.pdf

Wildfire, J., R. Bailey, R. Z. Krouse, S. Childress, B. Sikora, N. Bryant, S. Rosanbalm, E. Wilson, and J. G. Modell. 2018. The safety explorer suite: Interactive safety monitoring for clinical trials. *Therapeutic Innovation & Regulatory Science* 52 (6):696–700. doi:10.1177/2168479018754846.

Zink, R. C., R. D. Wolfinger, and G. Mann. 2013. Summarizing the incidence of adverse events using volcano plots and time intervals. *Clinical Trials: Journal of the Society for Clinical Trials* 10 (3):398–406. doi:10.1177/1740774513485311.