

Columbia–Suicide Severity Rating Scale Scoring and Data Analysis Guide

Contributors: Mary E. Nilsson, Shailaja Suryawanshi, Cristiana Gassmann-Mayer, Sarah Dubrava, Paul McSorley, and Kaihong Jiang.

Introduction

The Columbia–Suicide Severity Rating Scale (C-SSRS) is an assessment tool that evaluates suicidal ideation and behavior. This guide outlines the proposed safety outcomes and statistical analysis strategy for the C-SSRS for an individual clinical trial. The actual tables or combination of tables may vary to be consistent with a sponsor’s standards, or to be consistent with the needs of a particular regulatory agency/division. As such this document serves as a general guideline and is not a proposal for mandatory analyses involving the C-SSRS. When analyzing the C-SSRS across studies, the same safety outcomes can be used. However, different statistical methodology would apply to combining data across studies, and is not discussed in this document. The briefing document for the 13 December 2006 Food and Drug Administration (FDA) Advisory Committee Meeting contains a discussion of meta-analytical methods to consider when analyzing suicide-related outcomes across studies (<http://www.fda.gov/ohrms/dockets/ac/06/briefing/2006-4272b1-index.htm>). The C-SSRS may also be used to assess efficacy, but specific guidance related to efficacy assessment will be a topic for a separate document.

As noted in the August 2012 draft guidance titled “Suicidal Ideation and Behavior: Prospective Assessment of Occurrence in Clinical Trials”, the FDA has adopted the 11 categories defined in the C-SSRS (five subtypes of suicidal ideation, five subtypes of suicidal behavior, and self-injurious behavior without suicidal intent) as their standard. Some data collected on the C-SSRS are not included in these proposed displays for safety analysis (e.g., suicidal behavior lethality and suicidal ideation intensity), but are used for individual clinical management, safety monitoring, or other research purposes.

Analysis Set:

The analysis set should be defined consistently with the sponsor’s standards for the safety analysis set. Frequently, this means including patients who have been exposed to at least one dose of study drug. When analyzing the C-SSRS, it is recommended to include patients having at least 1 post-baseline C-SSRS measurement, regardless of whether they had a baseline C-SSRS measurement. For some analyses of the C-SSRS (e.g., treatment-emergent assessments and shift tables), a pre-treatment C-SSRS measurement is also required.

Outcomes:

There is current debate on whether suicidal ideation and suicidal behavior should be combined and analyzed as a single outcome. Researchers agree on the need for having analyses that keep suicidal ideation and suicidal behavior separate, but disagree on

whether there is value to having any outcome that combines them. This document currently includes the combined outcomes measure.

The following outcomes are C-SSRS categories and have binary responses (yes/no). The categories have been re-ordered from the actual scale to facilitate the definitions of the composite and comparative endpoints, and to enable clarity in the presentation of the results.

Category 1 – Wish to be Dead

Category 2 – Non-specific Active Suicidal Thoughts

Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act

Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan

Category 5 – Active Suicidal Ideation with Specific Plan and Intent

Category 6 – Preparatory Acts or Behavior

Category 7 – Aborted Attempt

Category 8 – Interrupted Attempt

Category 9 – Actual Attempt (non-fatal)

Category 10 – Completed Suicide

Self-injurious behavior without suicidal intent is also a C-SSRS outcome (although not suicide-related) and has a binary response (yes/no).

The following outcome is a numerical score derived from the C-SSRS categories. The score is created at each assessment for each patient and is used for determining treatment emergence.

- **Suicidal Ideation Score:** The maximum suicidal ideation category (1-5 on the C-SSRS) present at the assessment. Assign a score of 0 if no ideation is present.

Endpoints:

Composite endpoints based on the above categories are defined below.

- **Suicidal ideation:** A “yes” answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS.
- **Suicidal behavior:** A “yes” answer at any time during treatment to any one of the five suicidal behavior questions (Categories 6-10) on the C-SSRS.
- **Suicidal ideation or behavior:** A “yes” answer at any time during treatment to any one of the ten suicidal ideation and behavior questions (Categories 1-10) on the C-SSRS.

Comparative endpoints of interest are defined below. “Treatment emergence” is used for outcomes that include events that first emerge or worsen. “Emergence” is used for outcomes that include events that first emerge.

- Treatment-emergent **suicidal ideation** compared to recent history: An increase in the maximum suicidal ideation score during treatment from the maximum suicidal ideation category during a specified pre-treatment period (C-SSRS scales taken during the specified pre-treatment period; excludes “lifetime” scores from the Baseline C-SSRS scale or Baseline/Screening C-SSRS scale).
- Treatment-emergent **serious suicidal ideation** compared to recent history: An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment from not having serious suicidal ideation (scores of 0-3) during a specified pre-treatment period (C-SSRS scales taken during the specified pre-treatment period; excludes “lifetime” scores from the Baseline C-SSRS scale or Baseline/Screening C-SSRS scale).
- Emergence of **serious suicidal ideation** compared to recent history: An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment from no suicidal ideation (scores of 0) during a specified pre-treatment period (C-SSRS scales taken during the specified pre-treatment period; excludes “lifetime” scores from the Baseline C-SSRS scale or Baseline/Screening C-SSRS scale).
- Improvement in suicidal **ideation** at a time point of interest compared to baseline: An improvement in this endpoint can be considered as a decrease in suicidal ideation score at the time point of interest (e.g., the last measurement during treatment) from the baseline measurement (e.g., the measurement taken just prior to treatment. This analysis should only be performed for studies in which a baseline C-SSRS can be defined (i.e., having improvement from the worse event over a lifetime is not clinically meaningful).
- Emergence of **suicidal behavior** compared to all prior history: The occurrence of suicidal behavior (Categories 6-10) during treatment from not having suicidal behavior (Categories 6-10) prior to treatment (includes “lifetime” and/or “screening” scores from the Baseline C-SSRS scale, Screening C-SSRS scale, or Baseline/Screening C-SSRS scale, and any “Since Last Visit” from the Since Last Visit C-SSRS scales taken prior to treatment).

Additional comparative endpoints for consideration are defined below. These are not necessarily recommended for all treatment programs, but if used, should follow the nomenclature.

- Treatment-emergent suicidal **ideation** compared to all prior history: An increase in the maximum suicidal ideation score during treatment from the maximum suicidal ideation score prior to treatment (includes “lifetime” and/or “screening” scores from the Baseline C-SSRS scale, Screening C-SSRS scale, or Baseline/Screening C-SSRS scale, and any “Since Last Visit” from the Since Last Visit C-SSRS scales taken prior to treatment).

- Emergence of serious suicidal ideation compared to all prior history: An increase in the maximum suicidal ideation score to 4 or 5 during treatment from no suicidal ideation (scores of 0) prior to treatment (includes “lifetime” and/or “screening” scores from the Baseline C-SSRS scale, Screening C-SSRS scale, or Baseline/Screening C-SSRS scale, and any “Since Last Visit” from the Since Last Visit C-SSRS scales taken prior to treatment).
- Treatment-emergent **serious suicidal ideation** compared to all prior history: An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment from not having serious suicidal ideation (scores of 0-3) prior to treatment (includes “lifetime” and/or “screening” scores from the Baseline C-SSRS scale, Screening C-SSRS scale, or Baseline/Screening C-SSRS scale, and any “Since Last Visit” from the Since Last Visit C-SSRS scales taken prior to treatment).

Outcomes that can be used for clinical management and safety monitoring or potentially for research purposes are described below.

- Suicidal behavior lethality rating taken directly from the C-SSRS
- Suicidal ideation score
Any score greater than 0 is important and may indicate the need for mental health intervention. The protocol procedures related to clinical care of patients with treatment emergent suicidal ideation and behaviors will then be implemented to ensure proper management of the event and protection of the patient’s safety. For example, a score of 4 (active suicidal ideation with some intent to act) or 5 (active suicidal ideation with specific plan and intent) on suicidal ideation can be used to indicate serious suicidal ideation and can be used to trigger further evaluation and immediate contact with patient’s mental health practitioner (for non-psychiatry trials this may be used to trigger a prompt referral to a mental health professional) and/or possibly the emergency room.
- Suicidal ideation intensity rating
Add the five intensity item scores to create a total score (range 0 to 25) to represent the intensity rating. If the patient did not endorse any suicidal ideation set the intensity rating to 0.

Analytical and Statistical Methods:

The sponsor’s standard approaches to present safety data should be used consistently. Suicidal ideation and suicidal behavior are relatively rare events in individual, double-blind, controlled clinical trials, except possibly for a population particularly susceptible to this safety risk. In general, no formal statistical hypothesis testing is recommended for individual studies as only few events are typically observed and descriptive summaries will then suffice [9].

See the mock tables, the mock listing, and mock figure for suggested summaries and presentation of the data. Note that these mock tables can be easily extended to include multiple treatment groups/comparisons as needed. Tables should be modified to fit the needs of the specific analyses performed.

For very small studies (e.g., Phase 1 studies), or indications where suicide-related events are expected to be rare, it may be sufficient to plan on providing a listing only (see Listing 1). Otherwise, Tables 1 and 2 as shown may be provided.

The specific statistical method (or lack of an inferential method) should be consistent with the sponsor's standards. When inferential analyses are not performed the p-value column should be eliminated. Tables 3-4 (options for a shift table) will likely be used when there are findings to understand in more detail and/or when the expected number of suicide-related events is large either due to having a large study or having a population at high risk. A graphical presentation, for example a stacked bar chart, may also be used when appropriate (see Figure 1).

A summary of pre-treatment data is generally desirable and can be added to existing planned tables for patient characteristics (e.g., percent of subjects with lifetime suicidal ideation, baseline suicidal ideation, lifetime suicidal behavior, baseline suicidal behavior, lifetime self-injurious behavior without suicidal intent, and baseline self-injurious behavior without suicidal intent). Alternatively, if all categories of suicidal ideation and suicidal behavior are of interest for pre-treatment summaries, Table 1 can be repeated for providing a summary of lifetime outcomes and again for baseline outcomes. The specific timeframe used for defining baseline will need to be defined.

When inferential statistics are used, methods for consideration include: the Miettinen and Nurminen method [1], an unconditional, asymptotic method; the Fisher's exact test; 95% confidence intervals (CIs) based on Wilson's score method; and forest plots to present effect across treatment arms or subgroups of interest.

A further consideration for inferential analysis may be the requirement that at least 4 patients in any treatment group exhibit the event [2]. Because many 95% CIs may be provided without adjustment for multiplicity, the CIs should be regarded as a helpful descriptive measure to be used in review, not a formal method for assessing the statistical significance of the between-group differences in adverse experiences. Appropriate cautionary statements on the interpretation of inferential statistics for purposes of summarizing information or signaling trends should be made.

Methods for assessing dose-response relationships should be consistent with the sponsor's standards. Calculating and presenting incidence rates (e.g., n/patient-years) should be considered, especially for studies with long patient exposures and expected differential drop-out rates between treatment groups.

In long-term trials where the suicidality scale is administered frequently, a time-to-event analysis where the event is any suicidal event (either ideation or behavior whichever occurs first) can be performed using methods such as Kaplan-Meier method [3], log-rank

test [4], generalized Wilcoxon test [5] and Cox proportional hazards model [6] adjusting for a limited number of covariates. It is noted that event type-specific analysis (i.e., time-to-first-ideation and time-to-first-behavior analyses) may not be interpretable due to the presence of competing risks [7] and informative censoring [8]. In situations that warrant an assessment by age, gender and/or other risk factors summary statistics of scores on suicidality endpoints of interest can be presented by visit. For studies with longer duration, scores can be summarized by visit or time period (treatment phase, post-treatment phase etc.).

Based on the individual program strategy, active treatment groups may be pooled for comparisons to placebo/active comparator, or for performing all pair-wise dose comparisons. Analysis methods for assessing effects over several trials, subgroups or indications are presented in [9].

Note that missing data should not be imputed.

Table 1: Number of Patients with Suicidal Ideation, Suicidal Behavior, and Self-Injurious Behavior without Suicidal Intent Based on the C-SSRS During Treatment; [Define Analysis Set]

Events during treatment	Drug Name N=xx n (%)	Comparator Name N=xx n (%)	p-values ^a (to compare percentages)
Suicidal Ideation (1-5)	x (%)	x (%)	0.xxx
1) Wish to be dead	x (%)	x (%)	
2) Non-specific active suicidal thoughts	x (%)	x (%)	
3) Active suicidal ideation with any methods (not plan) without intent to act	x (%)	x (%)	
4) Active suicidal ideation with some intent to act, without specific plan	x (%)	x (%)	
5) Active suicidal ideation with specific plan and intent	x (%)	x (%)	
Suicidal Behavior (6-10)	x (%)	x (%)	0.xxx
6) Preparatory acts or behavior	x (%)	x (%)	
7) Aborted attempt	x (%)	x (%)	
8) Interrupted attempt	x (%)	x (%)	
9) Non-fatal suicide attempt	x (%)	x (%)	
10) Completed suicide	x (%)	x (%)	
Suicidal Ideation or Behavior (1-10)	x (%)	x (%)	0.xxx
Self-injurious behavior without suicidal intent	x (%)	x (%)	

NOTE: the p-value column should be eliminated when inferential analyses are not performed depending on the Sponsor's standard approach for safety data presentations.

^a p-values are from [specify test].

Notes: N = number of enrolled patients with at least one post-baseline C-SSRS assessment. In this table, n and (%) refer to the number and percent of patients who experience the event at least once during treatment. For the composite endpoint of suicidal ideation (1-5), n and (%) refer to the number and percent of patients who experience any one of the five suicidal ideation events at least once during treatment. For the composite endpoint of suicidal behavior (6-10), n and (%) refer to the number and percent of patients who experience any one of the five suicidal behavior events at least once during treatment. For the composite endpoint of suicidal ideation or behavior (1-10), n and (%) refer to the number and percent of patients who experience any one of the ten suicidal ideation or behavior events at least once during treatment.

Table 2: Number of Patients with Suicide-Related Treatment-Emergent Events Based on the C-SSRS During Treatment; [Define Analysis Set]

Treatment-emergent (TE) Events	Drug Name		Comparator Name		p-values ^a
	N	n (%)	N	n (%)	
TE suicidal ideation (1-5) compared to recent history ^b	xx	x (%)	xx	x (%)	0.xxx
TE serious suicidal ideation (0-3 to 4-5) compared to recent history ^c	xx	x (%)	xx	x (%)	0.xxx
Emergence of serious suicidal ideation (0 to 4-5) compared to recent history ^d	xx	x (%)	xx	x (%)	0.xxx
Improvement in suicidal ideation at endpoint compared with baseline ^e	xx	x (%)	xx	x (%)	0.xxx
Emergence of suicidal behavior (6-10) compared to all prior history ^f	xx	x (%)	xx	x (%)	0.xxx

Notes: For the composite endpoint of suicidal ideation (1-5), n and (%) refer to the number and percent of patients who experience treatment-emergent suicidal ideation during treatment. For the composite endpoint of suicidal behavior (6-10), n and (%) refer to the number and percent of patients who experience treatment-emergent suicidal behavior during treatment.

^a p-values are from [specify test].

^b N=Number of enrolled patients with at least one post-baseline suicidal ideation score and whose maximum C-SSRS suicidal ideation score during the comparison period is non-missing and <5.

^c N=Number of enrolled patients with at least one post-baseline suicidal ideation score and whose maximum C-SSRS suicidal ideation score during the comparison period is 0-3.

^d N=Number of enrolled patients with at least one post-baseline suicidal ideation score and whose maximum C-SSRS suicidal ideation score during the comparison period is 0.

^e N=Number of enrolled patients whose suicidal ideation score is non-missing and >0 just prior to treatment.

^f N=number of enrolled patients with at least one post-baseline C-SSRS assessment and who did not have suicidal behavior (6-10) prior to treatment.

NOTE: the p-value column should be eliminated when inferential analyses are not performed depending on the Sponsor's standard approach for safety data presentations.

Table 3. Shift-table to Demonstrate Changes in C-SSRS Categories from Baseline During Treatment; [Define Analysis Set]

Treatment	Baseline Category	Treatment Category		
		No suicidal ideation or behavior n (%)	Suicidal ideation n (%)	Suicidal behavior n (%)
Drug Name (N=xxx)	No suicidal ideation or behavior	x (%)	x (%)	x (%)
	Suicidal Ideation	x (%)	x (%)	x (%)
	Suicidal Behavior	x (%)	x (%)	x (%)
Comparator Name (N=xxx)	No suicidal ideation or behavior	x (%)	x (%)	x (%)
	Suicidal Ideation	x (%)	x (%)	x (%)
	Suicidal Behavior	x (%)	x (%)	x (%)

Notes: N = number of patients with a baseline and post-baseline C-SSRS assessment, n = number of patients in category. % = 100*n/N.

Baseline refers to [specify definition]

Suicidal Ideation includes any one of the five suicidal ideation events (Categories 1-5). Suicidal behavior includes any one of the five suicidal behavior events (Categories 6-10).

Each patient is counted in one cell only. Patients with both Suicidal Ideation and Suicidal Behavior are included in the Suicidal Behavior category.

Table 4. Shift-table to Demonstrate Changes in C-SSRS Suicidal Ideation Scores from Baseline During Treatment; [Define Analysis Set]

Treatment	Maximum Baseline Score	Maximum Suicidal Ideation Score During Treatment					
		0 n (%)	1 n (%)	2 n (%)	3 n (%)	4 n (%)	5 n (%)
Drug Name (N=xxx)	0	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	1	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	2	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	3	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	4	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	5	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
Comparator Name (N=xxx)	0	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	1	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	2	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	3	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	4	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)
	5	x (%)	x (%)	x (%)	x (%)	x (%)	x (%)

Notes: N = number of patients with a baseline and post-baseline C-SSRS suicidal ideation score, n = number of patients in category,
% = 100*n/N.

Baseline refers to [specify definition]; Maximum refers to the maximum C-SSRS suicidal ideation score during treatment (0 = least severe, 5 = most severe) where 0=No Suicidal Ideation, 1=Wish to be Dead, 2=Non-specific Active Suicidal Thoughts, 3=Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act, 4=Active Suicidal Ideation with Some Intent to Act, without Specific Plan, and 5=Active Suicidal Ideation with Specific Plan and Intent.

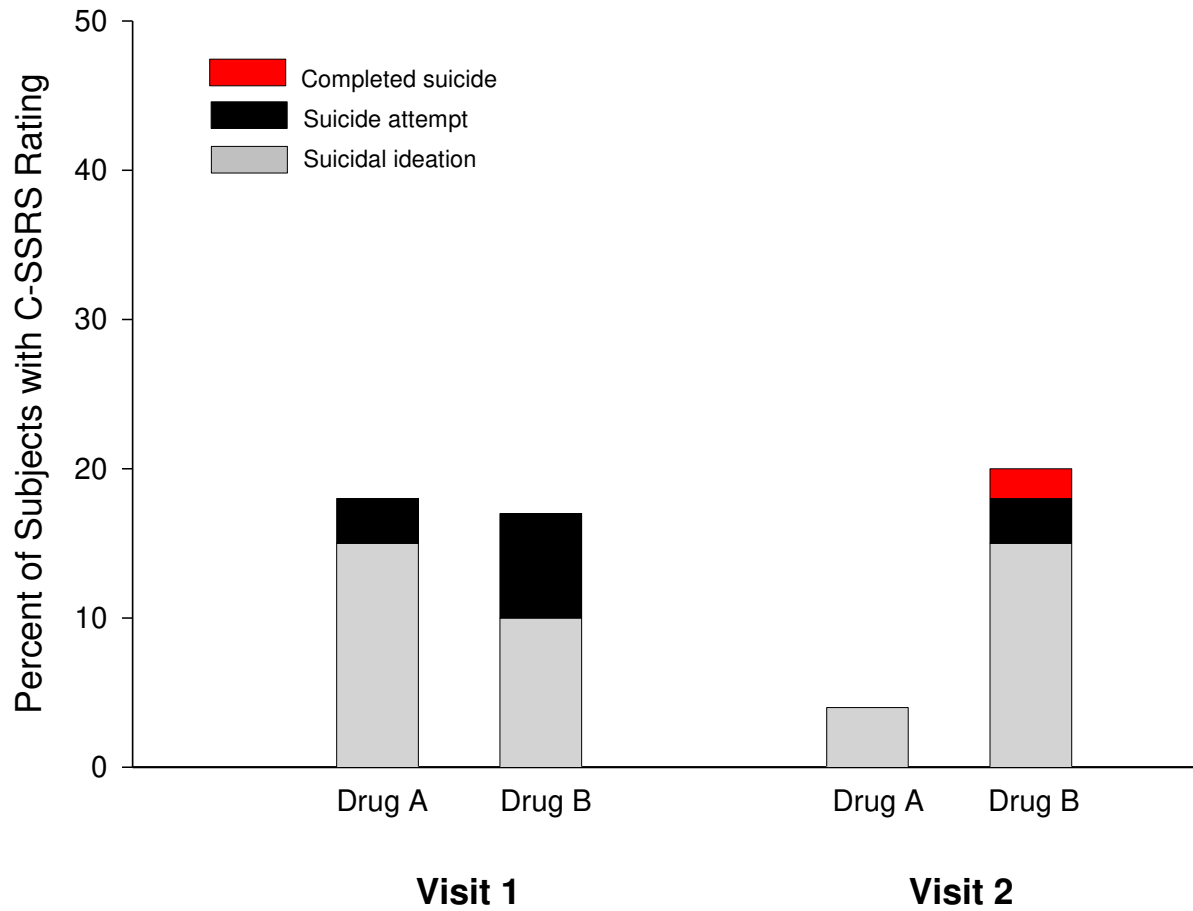
Listing 1: Listing of Subjects with Suicidal Ideation, Suicidal Behavior, or Self-Injurious Behavior without Suicidal Intent Based on the C-SSRS During Treatment^a; [Define Analysis Set]

Patient	Trt	Visit	Suicidal Ideation					Suicidal Behavior					Self-Inj Beh wo SI
			1	2	3	4	5	6	7	8	9	10	
xxxx			Y	Y	Y	N	Y	N	N	N	N	N	N

Note: Only patients with suicidal ideation, suicidal behavior, or self-injurious behavior without suicidal intent are displayed. For patients with suicidal ideation, suicidal behavior, or self-injurious behavior without suicidal intent at any time, data from all visits are displayed. Self-Inj Beh wo SI = Self-injurious Behavior without Suicidal Intent.

^a Key: 1=Wish to be Dead, 2=Non-specific Active Suicidal Thoughts, 3=Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act, 4=Active Suicidal Ideation with Some Intent to Act, without Specific Plan, 5=Active Suicidal Ideation with Specific Plan and Intent, 6=Preparatory Acts or Behavior, 7=Aborted Attempt, 8= Interrupted Attempt, 9=Actual Attempt (non-fatal), 10=Completed Suicide.

Figure 1: Distribution of Subjects With Suicide-Related Events Based on the C-SSRS Scores Over Time [Define Analysis Set]



References:

1. Miettinen O, Nurminen M. Comparative analysis of two rates. *Statistics in Medicine* 1985; 4:213-226.
2. Crowe B, Xia A, Berlin J, Watson D, Shi H, Lin S, Kuebler J, Schriver R, Santanello N, Rochester G, Porter J, Oster M, Mehrotra D, Li Z, King E, Harpur E, Hall D. Recommendations for safety planning, data collection, evaluation and reporting during drug, biologic and vaccine development: a report of the safety planning, evaluation, and reporting team. *Clinical Trials* 2009; 6:430-440.
3. Kaplan EL, Meier P. Nonparametric estimation for incomplete observations. *Am Stat Assoc* 1958; 53: 457-481.
4. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep* 1966; 50(3): 163-170.
5. Gehan EA. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* 1965; 52: 203-223.
6. Cox DR. Regression models and life tables. *J Royal Stat Soc* 1972; 34: 87-220.
7. Pintilie M. *Competing risks: a practical perspective*. John Wiley & sons: West Sussex, England, 2006.
8. Satagopan JM, Ben-Porat L, Berwick M, et al. A note on competing risks in survival data analysis. *Brit J Cancer* 2004; 91: 1229-1235.
9. Gassmann-Mayer C, Jiang K, McSorley P, Arani R, DuBrava S, Suryawanshi S, Webb DM, Nilsson M. Clinical and Statistical Assessment of Suicidal Ideation and Behavior in Pharmaceutical Trials. *Clin Pharmacol Ther* 2011; 90(4): 554-560.