PhUSE 2015

Paper PP21

Blinding the un-Blinded

Angelo Tinazzi, Cytel Inc., Geneva, Switzerland

Dean Shults, Cytel Inc., Geneva, Switzerland

ABSTRACT

In blinded or confirmatory open label trials where independent analyses through regular DMC are required, rules should be defined (i.e. SOPs and DMC Charter) so that trial un-blinding is not compromised and bias remains controlled. For these interim activities two separate teams sitting in different locations are recommended.

Usually outputs are developed and validated by a blinded team using dummy treatment codes, then transferred to a separate area handled by the un-blinded team who make sure programs works same way with real treatment codes.

There are circumstances where using dummy treatment code is not enough. For example in an open trial one may collect the relationship to a drug specific to one arm or in a blinded trial the occurrence of a particular AE may lead to guessing the arm.

In order to overcome this issue, Cytel has developed a process and a tool so that data are blinded by applying systematic 'data masking' criteria such as shuffling patients, records and/or variables, free text blinding and dictionary blinding.

This approach can be also useful when developing mock-up outputs when sponsor requires to use test data or a partial set of data; we believe working with a 'data masked' version of real data will expose the programming team to more 'real' data scenarios than working with test data.

DEFINITIONS

Blinded Data: Data from patient(s) in a clinical trial presented with no indication of which treatment the patient(s) are receiving or which group the patient(s) are in.

Blinded Study Team: The Trial (Blinded) Statistician and Sponsor team members who remain blinded throughout the study until database is locked and study un-blinded.

Trial (Blinded) Statistician: The Lead Study Statistician, who for randomized double blind studies is also referred as the blinded statistician.

Confirmatory Trial: An adequate and well controlled trial designed to provide additional or firm evidence of efficacy or safety, specifically where the principal aim it confirmation of hypotheses that have been developed in earlier stage of drug development.

Data Monitoring Committee (DMC): A group of individuals with pertinent expertise that reviews accumulating data from an on-going clinical trial. The DMC advises Sponsor regarding the continuing safety of trial subjects and those yet to be recruited to the trial, as well as the continuing validity and scientific merit of the trial.

Independent (Un-blinded) Statistical Team: The Independent Statistician is often supported by an Independent Programming Team, who will help to produce the statistical outputs delivered to the DMC. The Independent Statistician and Programming Team are both un-blinded, and collectively referred to as the Independent (Un-blinded) Statistical Team.

Independent (Un-blinded) Statistician: The Statistician who prepares data for review by the DMC and who is Independent from the Blinded Study Team conducting the study.

Un-blinded Data: Data from patient(s) in a clinical trial presented with the name of the treatment that the patient(s) are receiving explicitly presented.

INTRODUCTION

In blinded or confirmatory open label trials where independent analyses through regular DMC are required, rules should be defined (e.g. SOPs and DMC Charter) so that trial un-blinding is not compromised and bias remains controlled. One way of handling DMCs is to have two separate teams, blinded versus un-blinded Statistical Study Team, responsible for DMC (blinded) and Final Analysis (un-blinded) respectively. This solution requires additional resources since a separate Statistical Programming Team in addition to the blinded one, needs to be assigned to the project (see option 1 in Figure 1).

An alternative solution it is still to have two separate teams, sitting in different locations, where the Blinded Statistical Team is responsible for both DMC and Final Analysis activities. In this model all outputs are programmed and quality-controlled by the Blinded Statistical Team using blinded data and dummy treatment schedules as applicable. The Trial Blinded Statistical Team will then transfer the final validated programs to the un-blinded Statistical Team. The un-blinded Statistical Team will the rerun the programs using blinded data on the un-blinded study folder, and do a comparison between the provided blinded outputs and the blinded outputs run on the un-blinded platform, in order to confirm that the transferred programs produce the same blinded results. After this consistency check the un-blinded Statistical Team will update flags to use actual randomization assignments, and create the un-blinded outputs. Consistency checks are done to ensure the correct treatment assignments are being used and outputs transferred to the DMC members via a secure file transfer system.

There are circumstances where using dummy treatment codes may not be enough. For example in an open trial one may collect the relationship to a drug specific to one arm or in a blinded trial the occurrence of a particular AE may lead to guessing the arm. In order to overcome this issue Cytel has developed a process and a tool so that data are blinded by applying systematic 'data masking' criteria such as shuffling patients, records and /or variables, blinding free text and medical coding dictionary. In this model, in addition to the use of dummy treatment codes, data goes through an additional blinding, or masking process, where data are first transferred to the un-blinded Statistician who run the blinding / data masking criteria, then the 'manipulated' data are transferred to the blinded Statistical Team (see option 2 in figure 1). The process described before for statistical programs development is then applied.

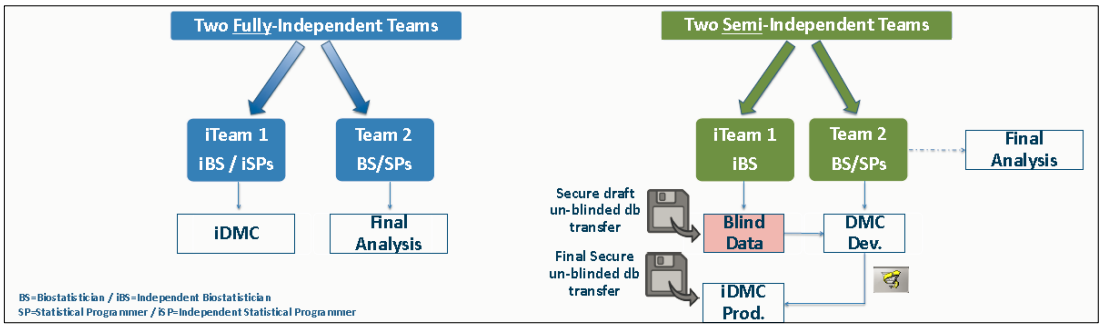


Figure 1: Two Fully-Independent Teams vs. Two Semi-Independent Teams

CYTEL PROCESS FOR BLINDING AND MASKING DATA

As previously described in addition to the use dummy treatment codes a "data masking" process is applied. Different masking criteria can be agreed between the blinded and un-blinded Statistical Team, approved by the Sponsor, and applied by the un-blinded Statistical Team prior to transfer data to the blinded Statistical Team. Data can be limited to a subset of patients (patient cut-off) and/or datasets/variables (datasets and variables hiding); remaining datasets and variables can go through various masking process such as patients and records shuffling, variables (numeric and text variables) and dictionary blinding. The process is then validated through some reports to show the correct implementation of the masking criteria.

PATIENT CUT-OFF

A percentage of patients are randomly selected and all records from unselected patients are removed from all datasets transferred from the un-blinded to the blinded Statistical Team.

DATASETS AND VARIABLES HIDING

If not all datasets/variables are needed and if among those datasets/variables there is a risk of providing un-blinded information, those datasets/variables can be removed from the transfer to the blinded statistical team.

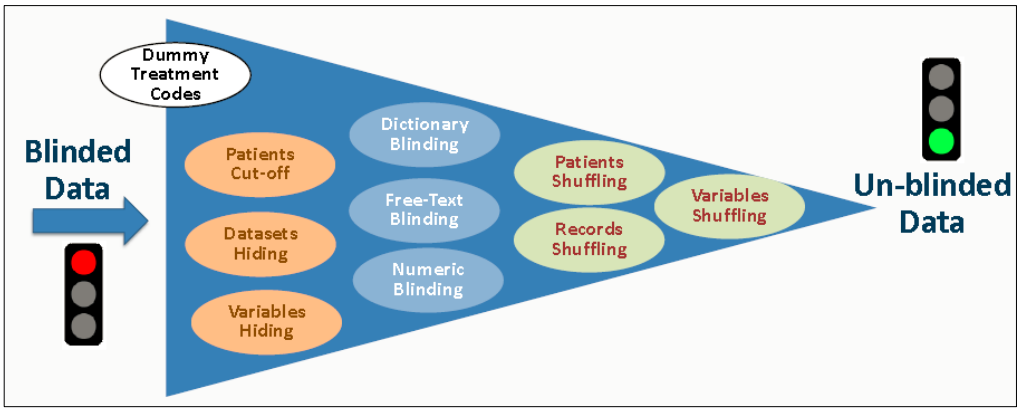


Figure 2: The Blinding and Masking Process

NUMERIC BLINDING

All or a selected set of variables are automatically "hard coded" to a random number. When for example it is applied to exposure, if a particular treatment has a particular dosage, assigning a random number will make treatment undistinguishable thus blinding the experimental treatment.

FREE-TEXT BLINDING

Any free text, including AE term, will be replaced by a random set of characters e.g. ALT increase will be replaced by XXX XXXXXXXX. This will make sure that no specific reference will be made to the experimental treatment (e.g. made by the investigator in the free text field).

DICTIONARY BLINDING

It can be applied whenever a dictionary is applied such as WHO-DD for CM. Method consists of getting unique observed CM Classes and PTs and randomly assign one of the unique terms to each occurred CM. If some CMs are typically related to the experimental treatment, we may see a higher incidence of those particular CMs expected for the experimental treatment thus making some conclusions on the safety profile of the experimental treatment. By blinding the dictionary we will still report the same observed type of CM Class/PT but with a 'modified' incidence rate as shown in the example in figure 3. The same is true for Adverse Events with MedDRA coding.

CM Class	CM PT Term	NBefore	Nafter
		245	
5-HT3 ANTAGONISTS	GRANISETRON HYDROCHLORIDE	12	41
5-HT3 ANTAGONISTS	ONDANSETRON	24	51
ANGIOTENSIN-CONVERTING ENZYME INHIBITORS	RAMIPRIL	18	40
ANTIHISTAMINES	CLEMASTINE	23	26
BETA-ADRENOCEPTOR BLOCKING AGENTS	BISOPROLOL FUMARATE	16	45
CALCIUM CHANNEL BLOCKING AGENTS	AMLODIPINE	16	48
HERBAL HOMEOPATHIC & DIETARY SUPPLEMENTS	IGNATIA	14	35
STATINS	SIMVASTATIN	18	40
STEROIDS	DEXAMETHASONE	16	47
THYROID HORMONES	LEVOTHYROXINE SODIUM	13	42

Figure 3: Effect of Dictionary Blinding

PATIENTS SHUFFLING

In all datasets patient id is systematically shuffled e.g. patient 001 may be assigned records of patient 002, patient 002 records of patient 010, etc. (see example in figure 4). This is done systematically in all records e.g. all records of patient 001 in all datasets will be assigned to patient 002. This way blinding is guaranteed at the patient level e.g. any conclusion / assessment on patient 001, if made, will be wrong.

RECORDS SHUFFLING

A number of records are randomly systematically "exchanged". If for example this is applied to exposure, records of experimental treatment may be assigned to the control arm. By doing this "data manipulation" inconsistencies will be generated.

VARIABLE SHUFFLING

For some selected variables expected values will be randomly assigned. For example, if applied to drug relationship and if expected values are 1 (Not related) and 2 (Related), the original values will be randomly replaced with one of the two values, thus if drug relationship is specifically collected for the experimental treatment the variable will 'lose' its original meaning.

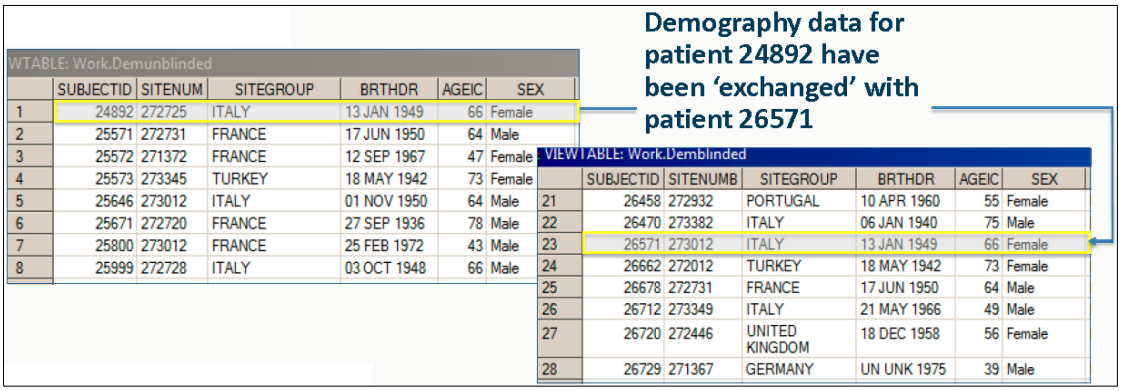


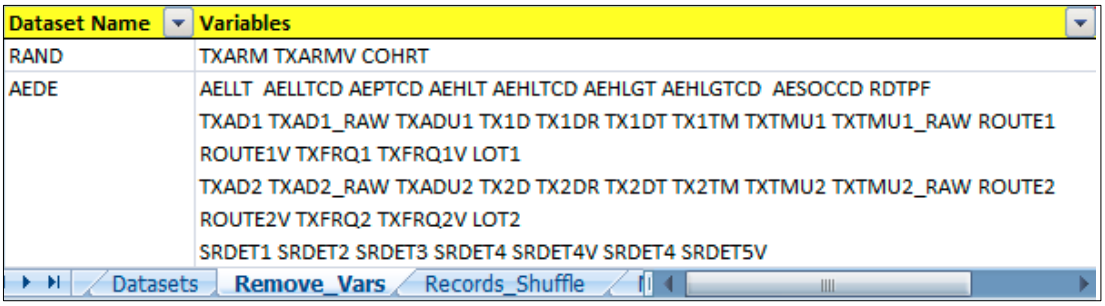
Figure 4: Effect of Patients Shuffling

%BLINDATA

%BLINDATA is a macro developed by the Cytel Statistical Programming Team. It consists of an excel file with several sheets where the user can specify the different masking criteria to apply. The excel file is the processed by the %BLINDATA macro that will first copy un-blinded raw datasets, apply the blinding and masking criteria and provides in outputs the 'hard-coded' blinded raw datasets.

Parameter	Description	Value
RAWUN	Unblided datasets location	..\DMCBlind\Biostatistics\Inputdata\Raw\Unblinde
RAWBLIND	Blinded datasets location. This will contains the artifacted datasets (derived from RAWUN)	..\DMCBlind\Biostatistics\Inputdata\Raw\Blinded
REMOVEDS	List of datasets that regardless of the data artifact process should be not provided to blinded team	See assessment in Datasets sheet
REMOVEVARS	YES/NO (default=NO). If YES specify in **Remove_Vars** sheet variables to be removed from each raw datasets	YES
REMOVEVARSALL	List of variables to be removed from all datasets	SUBJECT SITEID SITE
PCTPTS	% of patients to be included (default=100% so no cut will be applied all extracted patients will be provided to blinded	100
PTSHUFFLE	YES/NO See General Criteria (default=YES)	YES
RECSHUFFLE	YES/NO See General Criteria (default=YES). If YES specify additional criteria in **Records_Shuffle** sheet	NO
NUMBLIND	YES/NO See General Criteria (default=YES). If YES specify additional criteria in **Numeric_Blinding** sheet	NO
TXTBLIND	YES/NO See General Criteria (default=YES). If YES specify additional criteria in **Text_Blinding** sheet	YES
DICTBLIND	YES/NO See General Criteria (default=YES). If YES specify additional criteria in **Dictionary_Blinding** sheet	YES
VARSHUFFLE	YES/NO See General Criteria (default=YES). If YES specify additional criteria in **Variables_Shuffle** sheet	YES
DSPT	Datasets identifying patients e.g. DM	PTID
SUBJID	Subject Id	SUBJECTID

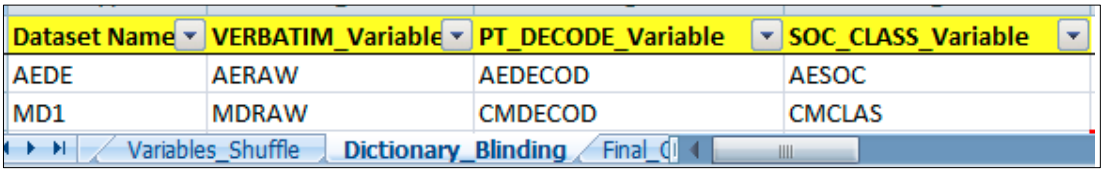
Figure 5: General Specifications



Dataset Name	Variables
RAND	TXARM TXARMV COHRT
AEDE	AELLT AELLTCD AEPTCD AEHLT AEHLTCD AEHLGT AEHLGTCD AESOCCD RDTPF
	TXAD1 TXAD1_RAW TXADU1 TX1D TX1DR TX1DT TX1TM TXTMU1 TXTMU1_RAW ROUTE1
	ROUTE1V TXFRQ1 TXFRQ1V LOT1
	TXAD2 TXAD2_RAW TXADU2 TX2D TX2DR TX2DT TX2TM TXTMU2 TXTMU2_RAW ROUTE2
	ROUTE2V TXFRQ2 TXFRQ2V LOT2
	SRDET1 SRDET2 SRDET3 SRDET4 SRDET4V SRDET4 SRDET5V

Datasets | Remove_Vars | Records_Shuffle

Figure 6: Variables to be removed from Blinded Transfer



Dataset Name	VERBATIM_Variable	PT_DECODE_Variable	SOC_CLASS_Variable
AEDE	AERAW	AEDECOD	AESOC
MD1	MDRAW	CMDECOD	CMCLAS

Variables_Shuffle | Dictionary_Blinding | Final_C

Figure 7: Dictionary Blinding Specifications

Figure 5, 6 and 7 provide a real example where we applied blinding/masking criteria. We first specified the set of criteria to apply (figure 5) and also specified the un-blinded and the blinded datasets location. We then provided additional details such as variables to be removed (figure 6) and criteria for blinding dictionary data (figure 7); for example the specifications in figure 7 tell the %BLINDATA macro to blind coding in AEDE (Adverse Events) and MD1 (Concomitant Medications) datasets by shuffling respectively AESOC/AEDECOD and CMCLAS/CMDECOD variables.

CONCLUSIONS

Although the solution we propose requires to carefully plan and select the appropriate variables to transfer and/or blind, this approach consistently reduced the resources needed (no team duplication) and made easier re-use of DMC Analysis Programs for Final Analysis since the same developing programming team is used, thus owning the programing package for all analysis-tasks.

The %BLINDATA solution only requires the specification of a number of parameters to correctly apply the 'blinding/masking' criteria.

This approach can be also useful when developing mock-up outputs where the use of test/partial data is required by the Sponsor; this way the programming team will be exposed to more 'real' data scenarios and avoid using test data created with different purpose e.g. User Acceptance Tests for eCRF.

REFERENCES

[1] S Ellenberg, TR Fleming, DL DeMets, Data Monitoring Committees in Clinical Trials: A Practical Perspective; Statistics in Practice 2002.

[2] European Medicines Agency, Guideline on Data Monitoring Committee; EMA 2005.

[3] Food and Drug Administration (FDA), Guidance for Clinical Trial Sponsors: Establishment and Operational of Clinical Trial Data Monitoring Committees; FDA 2006.

[4] A Grant et al, Issues in data monitoring and interim analysis of trials; Health Technology Assessment 9(7) 2005.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Angelo Tinazzi
Cytel Inc.
Route de Prè-Bois 20
1215 Geneva – Switzerland
Email: angelo.tinazzi@cytel.com
Web: www.cytel.com