2S-Sub-TITE: An adaptive two-stage time-to-toxicity model for subgroup-specific dose finding in Phase I oncology trials

Alana McGovern[1], Andrew G. Chapple[2], Clement Ma[1,3]

[1]Dana-Farber/Boston Children's Cancer and Blood Disorders Center; [2]Biostatistics Program, Louisiana State University; [3]Harvard Medical School

Introduction

The subgroup-specific time-to-event (Sub-TITE) phase 1 design is a modification of the time-to-event Continual Reassessment Method (TITE-CRM) which identifies the maximum tolerated dose (MTD) separately for 2 or more heterogeneous patient subgroups. This is the dose that has an estimated toxicity probability closest to some target π^*

Sub-TITE allows borrowing strength and dynamic clustering across subgroups from the start of the trial, but delaying the initiation of borrowing and clustering until later in the trial may improve dose-finding accuracy.

We introduce a practical extension of the Sub-TITE design, where borrowing and/or clustering may be initiated at some pre-specified point in the trial, based on patient enrollment.

The 2-Stage Sub-TITE (2S-Sub-TITE) design is a flexible framework that can allow more DLT information to be gathered before determining whether subgroups should be clustered.

Our goal is to optimize operating characteristics across many scenarios based on
(1) Timing of initiation of borrowing and/or clustering
(2) Prior assumptions of subgroup heterogeneity
(3) Hyperparameters based on physician elicited dose-specific toxicity probabilities.

2S-SubTITE Design

Stage 1: No Borrowing or Clustering

- Each patient's observation time and DLT status, as well as the physician elicited prior, contributes to the working likelihood estimates for their subgroup
- These estimates are then used to determine the probability of toxicity at each dose, for that subgroup, via the logit function
- The dose with an estimated probability of toxicity closest to the target toxicity probability, π^*, is the recommended dose for the next patient enrolled into that subgroup

We will continue in this manner until a pre-specified point of patient accrual.

Stage 2: Adaptive Borrowing and Clustering

- Each patient's observation time and DLT status, as well as the physician elicited prior, contributes to a *joint* working likelihood which provides effect estimates for both subgroups. We allow for the dynamic clustering of subgroups by introducing random latent subgroup membership variables, whose distributions are dictated by a user-specified *P_hetero,* which controls the probability that a subgroup g will *not* be dynamically clustered on other subgroups.
- The following 2 steps proceed similarly to Stage 1

Simulation Design and Conduct

Using the SubTite R package, we performed extensive simulations to assess the operating characteristics of different versions of our 2S-Sub-TITE design. The varied parameters in our simulations are detailed as follows:

- *True toxicity probability scenarios*: we randomly generated 1000 unique simulation scenarios.
- *Prior toxicity probabilities: we* selected 18 different sets of prior toxicity probabilities with varied characteristics.
- *Escalation* rules*:* in our primary set of simulations we use conservative escalation rules (requires that three patients be fully evaluated at a dose before escalating to the next, previously untested, dose) and in a secondary set of simulations we use aggressive escalation rules (requires only one patient to be evaluated at a dose before escalating).
- *Sample size:* in our primary set of simulations, the total sample size was 40 patients. In a secondary set of simulations the total sample size was 64.
- *Methods of borrowing and clustering:* we evaluated 18 methods with varying combinations of when the borrowing and/or clustering begins and P_hetero, as shown below

	Method*	Start borrowing at	Start clustering at	P_hetero
A	Brrw&Clstr@0%|Phet=0	0% Accrual	0% Accrual	0
B	Brrw&Clstr@0%|Phet=0.5	0% Accrual	0% Accrual	0.5
C	Brrw&Clstr@0%|Phet=0.7	0% Accrual	0% Accrual	0.7
D	Brrw&Clstr@0%|Phet=0.9	0% Accrual	0% Accrual	0.9
E	Brrw&Clstr@25%|Phet=0.5	25% Accrual	25% Accrual	0.5
F	Brrw&Clstr@25%|Phet=0.7	25% Accrual	25% Accrual	0.7
G	Brrw&Clstr@25%|Phet=0.9	25% Accrual	25% Accrual	0.9
H	Brrw&Clstr@50%|Phet=0.5	50% Accrual	50% Accrual	0.5
I	Brrw&Clstr@50%|Phet=0.7	50% Accrual	50% Accrual	0.7
J	Brrw&Clstr@50%|Phet=0.9	50% Accrual	50% Accrual	0.9
K	Brrw&Clstr@75%|Phet=0.5	75% Accrual	75% Accrual	0.5
L	Brrw&Clstr@75%|Phet=0.7	75% Accrual	75% Accrual	0.7
M	Brrw&Clstr@75%|Phet=0.9	75% Accrual	75% Accrual	0.9
N	Brrw@0%|NoClstr|Phet=1	0% Accrual	--	1
O	Brrw@25%|NoClstr|Phet=1	25% Accrual	--	1
P	Brrw@50%|NoClstr|Phet=1	50% Accrual	--	1
Q	Brrw@75%|NoClstr|Phet=1	75% Accrual	--	1
R	NoBrrw|NoClstr	--	--	1

For our primary set of simulations, 100 simulated trials were conducted for each combination of the 18 priors, 18 methods, and 1000 scenarios, fixing sample size at 40 and using conservative escalation.

We also assumed 2 evenly distributed subgroups, 5 possible doses, a DLT observation period of 1 month, and a target toxicity probability of 0.2.

The primary endpoint of interest in each scenario is the probability of selecting the optimal dose, denoted *PCS*.

Results

We compared the mean *PCS* of each 2S-Sub-TITE method separated by each of the tested varied parameters (figure not shown) and concluded:

- Trials with some priors perform better than others (yielding higher PCS), but the effect of this is not modified by choice of method
- Trials using conservative escalation perform uniformly slightly better (1.1%) than those using aggressive escalation, but, similarly the effect is not modified by choice of method
- Trials with a sample of size of 64 perform uniformly better (35%) than those with a sample size of 40, but this effect is also not modified by choice of method
- Trials with true subgroup MTDs which are closer together (or even equal), perform better than those with true subgroup MTDs which are further apart. This effect, however, is modified by choice of method.

These findings highlight the importance of considering the difference between MTDs when choosing a 2S-Sub-TITE method, compared to other factors.

Method Recommendations and Other Considerations

If you are not confident in your prior, consider using the most robust method, Method I: Allow borrowing & clustering starting **at 50% accrual**; ***p_hetero*=0.7**

Otherwise, if your prior indicates the difference between subgroup MTDs to be:

0

Consider using the original TITE-CRM model, however, if you want to allow for different subgroups, consider using:
- Method C: Allow borrowing & clustering starting **at 0% accrual**; ***p_hetero*=0.7**
- Method J: Allow borrowing & clustering starting **at 50% accrual**; ***p_hetero*=0.9**
- Method P: Allow borrowing starting **at 50% accrual**; **no clustering**

Avoid using:
- Method R: **No borrowing or clustering**

1*

Consider using:
- Method E: Allow borrowing & clustering starting **at 25% accrual**; ***p_hetero*=0.5**
- Method I: Allow borrowing & clustering starting **at 50% accrual**; ***p_hetero*=0.7**
- Method L: Allow borrowing & clustering starting **at 75% accrual**; ***p_hetero*=0.7**
- Method N: Allow borrowing starting **at 0% accrual**; **no clustering**

Avoid using:
- Method A: Allow borrowing & clustering starting **at 0% accrual**; ***p_hetero*=0**
- Method R: **No borrowing or clustering**

≥ 2

Consider using separate TITE-CRM models for each subgroup, however, if sample size is prohibitive, consider using:
- Method F: Allow borrowing & clustering starting **at 25% accrual**; ***p_hetero*=0.7**
- Method O: Allow borrowing starting **at 25% accrual**; **no clustering**
- Method M: Allow borrowing & clustering starting **at 75% accrual**; ***p_hetero*=0.9**

Avoid using:
- Method A: Allow borrowing & clustering starting **at 0% accrual; p_hetero=0**
- Method C: Allow borrowing & clustering starting **at 0% accrual**; ***p_hetero*=0.7**
- Method J: Allow borrowing & clustering starting **at 50% accrual**; ***p_hetero*=0.9**

Sample Size
As with many trial designs, increasing sample size will increase dose-finding accuracy. However, picking the right borrowing & clustering method can increase dose-finding accuracy by just as many percentage points, sometimes more. (Figure 1D)

Escalation Type
Choosing a conservative escalation approach, waiting for 3 patients to be fully evaluated at the previous dose before escalating to the next dose, will not significantly affect the accuracy of dose-finding. (Figure 1C)

Choice of Prior
In agreement with previous research findings, our results show that choice of prior does not have a significant effect on dose-finding accuracy. The chosen prior has most influence in the model at the beginning of the trial and loses influence as more data is gathered. If an investigator is unsure of their chosen prior, they should consider choosing a method which does not allow for clustering until later in the trial, so that more data can be gathered before clustering decisions are made. (Figure 1A)

*When the difference in subgroup MTDs is equal to 1, the affect of method choice on dose-finding is not as large. Thus, if the investigator has an inclination that the difference in subgroup MTDs is more likely to be 0 than ≥ 2, or vice versa, then the investigator should take this into account when choosing a method.

Reference

O'Quigley J, Pepe M and Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. *Biometrics* 1990; 46: 33-48. 1990/03/01.

Chapple AG and Thall PF. Subgroup-specific dose finding in phase I clinical trials based on time to toxicity allowing adaptive subgroup combination. *Pharm Stat* 2018; 17: 734-749. 2018/08/17. DOI: 10.1002/pst.1891.

Chapple AG. SubTite: Subgroup Specific Optimal Dose Assignment. R Package Version 4.0.0. R package version 4.0.1 ed. 2020.

Acknowledgements

This research was done using resources provided by the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science [31 32]

We give special thanks to Dr. Karen Wright and Dr. Wendy London for their thoughtful discussion and contributions.