

Aristotle University of Thessaloniki
Department of Mathematics
Postgraduate Program in Statistics and
Mathematical Modelling



Statistical Design of Phase II and Phase III Clinical Trials

Evripidis Kapanidis
Supervisor:Assistant Professor Georgios Afendras

July 27, 2022

Statistical Design of Phase II and Phase III Clinical Trials



A thesis submitted for the degree of
MASTER OF SCIENCE

Examination Committee:

Assistant Professor Georgios Afendras (Supervisor)

Research Personell-Lecturer Vassilis Karagiannis

Professor Georgios Tsaklidis

July 27, 2022

Declaration of Authorship

I, Kapanidis Evripidis, declare that this thesis titled, Statistical Design of Phase II and Phase III Clinical Trials, and the work presented in it are my own. I confirm that:

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this university or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Abstract

In this thesis we will present statistical methodology for the design and analysis of phase II and phase III clinical trials. The statistical methods presented here are mainly found in the therapeutic area of oncology. Initially, we describe and present fundamental terminology commonly found in clinical trials, the general structure of clinical trials and the relative statistical considerations. Afterwards we study widely used phase II oncology designs. Particularly we present designs that terminate early for futility, designs that compare multiple treatments and inferential methodologies for phase II trial data (estimation of response rates). Furthermore, we present suitable designs for the special cases of heterogeneous populations (that is, when there are multiple subpopulations that have a significantly different response to a treatment) and non-existent historical data. Additionally, for phase III trials we introduce and describe categories of designs that ask different questions of clinical interest (superiority,non-inferiority and equivalence trials) and utilize different endpoints (quantitative,dichotomous and survival endpoints). First we present the main theory for survival endpoints. We then examine the sample size needed for each clinical question and endpoint to acquire proper power and also present methods to determine the duration of phase III clinical trials when survival data is used. Moreover, we examine the methodology and assumptions in group sequential trials as well as the different terminating boundaries that may be used for a group sequential design. Afterwards, we perform simulation studies for both Phase II and Phase III trials. Through the use of simulations we compare different designs for the same hypotheses and examine how the sample size and expected sample size changes. We also compare the distributions of different estimators for the response rate and perform power analyses for superiority, non-inferiority and equivalence phase III trials. Lastly, we analyze real survival data and use the results of this analysis to design a hypothetical future clinical trial.

Περίληψη

Στην παρούσα εργασία θα ασχοληθούμε με μεθόδους σχεδιασμού και ανάλυσης κλινικών δοκιμών φάσης II και φάσης III που συναντώνται στην ογκολογία. Στην αρχή, παρουσιάζουμε βασικές ορολογίες που χρησιμοποιούνται στις κλινικές δοκιμές, τη γενικευμένη δομή τους καθώς και τον ρόλο της στατιστικής για τον κατάλληλο σχεδιασμό και ανάλυση τους. Εν συνεχείᾳ, παρουσιάζουμε τους πιο γνωστούς σχεδιασμούς για κλινικές δοκιμές φάσης II στην ογκολογία. Συγκεκριμένα θα παρουσιάσουμε σχεδιασμούς που τερματίζονται νωρίς όταν τα πρώτα αποτελέσματα από την κλινική δοκιμή υποδειχνύουν ότι η πειραματική θεραπεία δεν έχει κάποιο όφελος, σχεδιασμούς που συγχρίνουν πολλά φάρμακα και μεθοδολογίες εκτίμησης των response rates κατόπιν της κλινικής δοκιμής. Παρουσιάζουμε επίσης σχεδιασμούς για ειδικές περιστάσεις, συγκεκριμένα σχεδιασμούς για ετερογενείς πληθυσμούς και για όταν συναντάται έλλειψη ιστορικών δεδομένων. Στο κεφάλαιο 3 θα παρουσιάσουμε και για περιγράψουμε σχεδιασμούς που θέτουν διαφορετικές ερωτήσεις κλινικού ενδιαφέροντος (ανωτερότητας, μη-κατωτερότητας, ισοδυναμίας) και που χρησιμοποιούν διαφορετικούς τύπους δεδομένων (ποσοτικά, διχοτομικά και δεδομένα επιβίωσης). Πρώτα παρουσιάζεται η θεωρία που χρησιμοποιείται για δεδομένα επιβίωσης. Έπειτα μελετάμε το μέγεθος δείγματος που χρειάζεται για να έχουμε κατάλληλη ισχύ για κάθε κλινική ερώτηση και τύπο δεδομένου και παρουσιάζουμε μεθοδολογία βάσει της οποίας αποφασίζουμε τη διάρκεια της μελέτης όταν χρησιμοποιούνται δεδομένα επιβίωσης. Τέλος εξετάζουμε τους ακολουθιακούς σχεδιασμούς και τις παραδοχές τους, καθώς και τους διαφορετικούς κανόνες τερματισμού. Κατόπιν, πραγματοποιούμε μελέτες προσομοίωσης. Μέσω των προσομοιώσεων συγχρίνουμε διαφορετικούς σχεδιασμούς και προσεγγίσεις για τις ίδιες υποθέσεις και εξετάζουμε πιας το απαιτούμενο δείγμα και το αναμενόμενο δείγμα μεταβάλλονται. Συγχρίνουμε επίσης τις κατανομές εκτιμητριών των response rates και πραγματοποιούμε ανάλυση ισχύος για δοκιμές ανωτερότητας, μη-κατωτερότητας και ισοδυναμίας. Στο τέλος αναλύουμε πραγματικά δεδομένα επιβίωσης και αξιοποιούμε τα αποτελέσματα της ανάλυσης για να σχεδιάσουμε κατάλληλα μια υποθετική μελλοντική κλινική δοκιμή.

Acknowledgements

I would like to thank Dr. Afendras for his precious advice and guidance throughout the preparation of this thesis. I am also grateful to my family for supporting me during my studies.

Contents

1	Introduction	1
1.1	Definition of Clinical Trials	1
1.1.1	Subtypes of Clinical Trials	2
1.1.2	Organizational Structure of Clinical Trials	2
1.1.3	Phases of Clinical Trials	3
1.2	The Statistical Viewpoint of Clinical Trials	5
1.2.1	General Methodology in Phases of Clinical Trials	7
1.2.2	Statistical Issues in Phase II and III Designs	8
1.3	Structure of this Thesis	11
2	Phase II Trials	13
2.1	Single Arm Designs	13
2.1.1	Simon's two stage design	14
2.1.2	Heterogeneous designs	16
2.1.3	Designs For Early Termination	20
2.1.4	UMVU Estimator of Reponse Rate	22
2.2	Multiple Arm Designs	26
2.2.1	Simon's Randomized Designs	26
2.2.2	Jung's Randomized Phase II Designs	31
2.3	Discussion of Methods	34
3	Phase III trials	37
3.1	Classification of Phase III trials: Discussion	37
3.2	Endpoints In Phase III Trials	41
3.2.1	Survival Endpoints	41
3.2.1.1	Survival Functions	43
3.2.1.2	Estimation of Survival Curves	44
3.2.1.3	Cox's Proportional Hazards Model	47

3.2.1.4	Residuals and Diagnostics For Cox's Model	54
3.2.1.5	Log Rank Test	57
3.2.2	Quantitative and Dichotomous Endpoints	60
3.3	Sample size estimation	67
3.4	Group Sequential Designs	72
3.4.1	Assumptions and Power in Group Sequential Designs	75
3.4.2	Error Rates Probability Computation	76
3.4.3	Choice of Boundaries	80
4	Simulations and Data Analysis	83
4.1	Designing Simon's Two Stage Designs	83
4.2	Designing Simon's Randomized Phase II Designs	86
4.3	Distribution of UMVUE vs Distribution of MLE	87
4.4	Power Analysis For Phase III Trials	93
4.5	Group Sequential Trials	96
4.6	Survival Trial Design Using Historical Data	99
A	R Code	107
B	UMVUE for Two Stages	126
	Bibliography	128