Case Study: Phase II/III Adaptive Design

Introduction

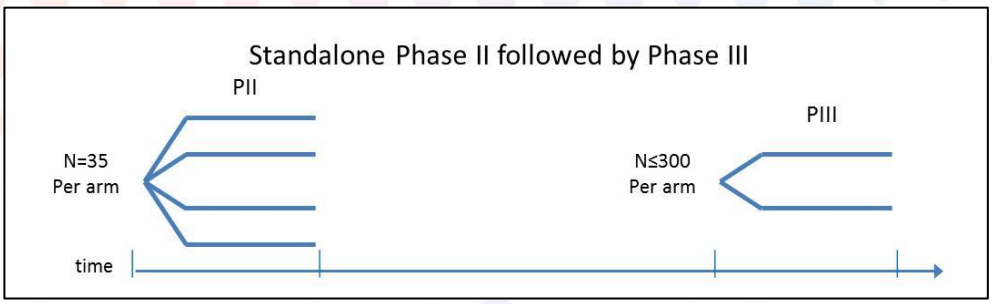
Pharmaceutical drug development continues to very challenging. At an estimated $800m-$2b[1] per development, bringing a new drug to patients has never been more expensive. R&D productivity remains lower than desired across the industry as a whole, with reported failure rates of 80% in Phase II and 50% in Phase III[2,3]. And two thirds of those Phase III failures are reported as due to not demonstrating a positive treatment effect, reflecting poorly on the quality of Phase II design and decision making[3].

In the face of escalating costs and high failure rates, novel statistical methods, including flexible and adaptive designs, offer the opportunity to improve decision making, accelerate development times and enhance the overall chances of success.

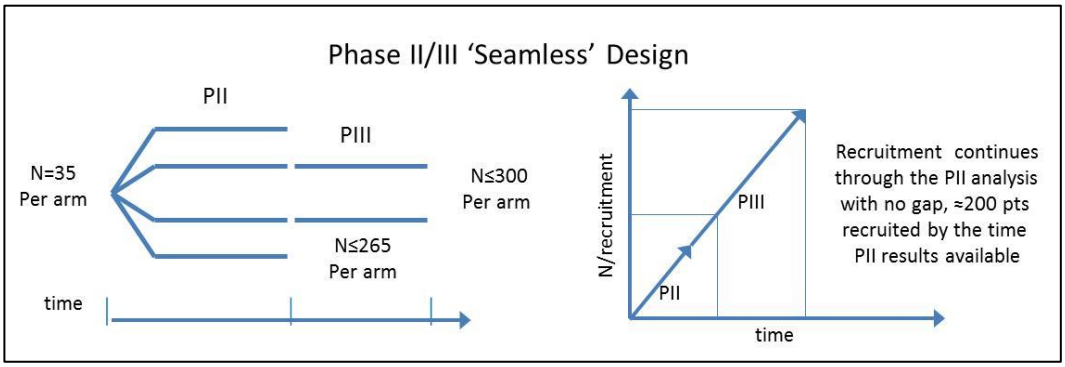
Problem

A client wished to employ an efficient adaptive design to cost-effectively expedite development. The basic design concept was a Phase II/Phase III 'seamless' study. Phase II involved placebo and 3 doses of drug, one of which would be selected to go forward into Phase III. Given uncertainty regarding the primary endpoint, sample size re-estimation was desired at the time the dose selection was made. Finally, precedent and feasibility meant that at most around N=35-40 patients could be recruited per arm in Phase II, expanding to a maximum of around 300 per arm in Phase III.

The proposed design was evaluated statistically and strategically. Modifications and/or alternatives that might better serve the clients' needs and objectives were offered.



Since recruitment was planned to continue through the end of Phase II and given the time required to complete the data collection and analysis, up to approx. 200 patients, i.e an overage of approx. 60 patients, would have been entered by the time the dose decision was made. Around half of these patients would effectively be 'lost' as only one dose of drug and placebo were to be carried forward in to Phase III.



Evaluation

The design was evaluated using multiple simulations assuming treatment effect of 2 units with a SD of 6 units. A target of N=150 per arm was set for Phase III (so an additional N=115 per arm) to deliver a nominal 82.5% power. The dose with best response was to be selected for Phase III and sample size increased to a maximum of N=300 per arm. Although no threshold for Conditional Power was stipulated in the protocol to trigger a sample size increase, the design was evaluated via various options using an approach similar to that described by Mehta and Pocock4.

Table 1: design evaluation via multiple simulation

Design Option	Increase in N if conditional power in range	under the null		under the alternative		Final p-value	Type I error	Power
		% repeat trials with an increase in N after interim	Median increase in N	% repeat trials with an increase in N after interim	Median increase in N			
Protocol	0-80%	78%	150	20%	123	0.025	0.037	94%
Option #1	25%-80%	25%	102	14%	82	0.025	0.047	90%
Option #2	33%-80%	21%	90	15%	77	0.025	0.044	92%
Option #3	50%-80%	14%	75	10%	70	0.025	0.039	91%
Option #4	No increase	0%	0	0%	0	0.025	0.046	89%

Simulations indicated large increases in N were expected with high probability if no or too low a Conditional Power threshold was set. Strategically, and in terms of time and cost, options #1 and #2 were therefore to be avoided. Also, the increase in Phase III power associated with selection of the best performing dose in Phase II was observed to be gained at the price of an inflated Type I error.

Further, simulations suggested setting an adjusted final alpha 1-sided level of 1.15% or less to control the overall Type I error at approximately 2.5%. In so doing, power was observed to be 82-83%, in line with the nominal 82.5% targeted with N=150 per arm. Interestingly, this level of power was reached with or without sample size re-estimation which would be expected to increase N by 70 per arm.

Table 2: design performance with an adjusted final p-value

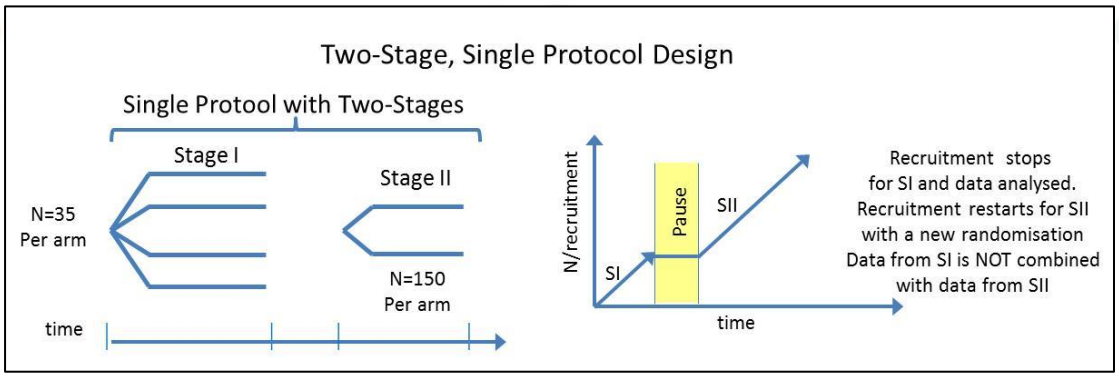
Design Option	Increase in N if conditional power in range	under the null: % repeat trials with an increase in N after interim	under the null: Median increase in N	under the alternative: % repeat trials with an increase in N after interim	under the alternative: Median increase in N	Final p-value	Type I error	Power
Option #3	50%-80%	17%	69	10%	68	0.0115	0.024	83%
Option #4	No increase	0%	0	0%	0	0.0115	0.026	82%

An Alternative Approach

Given the final trial results were intended as pivotal evidence to support a product licence application and associated labelling, an important complication identified was the best practice need to employ an Independent Data Monitoring Committee to oversee the Phase II interim analysis and unblinding. In line with regulatory guidance, this was judged necessary to assure and protect the integrity of study in its continuance into Phase III. Further, and In line with both FDA and EMA guidance, possible adaptations, including the algorithm for de-selection of 2 doses of drug and methodology for sample size re-estimation, needed to be pre-specified at the outset and the Sponsor kept blinded throughout the process to minimise the risk of influence and bias. Hence, and without the construction of complex firewalls and procedures, the Sponsor typically would not have sight of the critical unblinded interim Phase II data.

These complications led to a suggestion for a new, alternative flexible design and strategy, one that would still incorporate a Phase III sample size reassessment and yet would allow the Sponsor access to unblinded Phase II data and avoid the Type I error inflation association with the 'seamless' design.

Table 3. Suggested alternative Two Stage design option



With a Two Stage, Single Protocol design, Phase II dose selection and Phase III confirmation are still retained in a single protocol expediting development time and necessitating just one set of IRB approvals and one set of trial/centre initiations and costs. Recruitment in Phase II would not continue through the analysis as in the 'seamless' design, but would pause for a short period while the Phase II data were analysed and

reviewed. Importantly, this means that the Sponsor could now participate directly in the Phase II data review and critical dose selection since, in this design option, patients from

Phase II are not subsumed into Phase III but are evaluated independently. In addition to providing the Sponsor with sight of the Phase II data, a further benefit of the Two Stage design is that the Phase II portion would provide important, independent supporting evidence to the pivotal Phase III results, a benefit that did not exist with the 'seamless' design as Phase II patients were rolled over and absorbed into Phase III. Once Phase II results are known in the Two Stage design, the dose of drug to be evaluated in Phase III would be selected and recruitment into Phase III commence with initiation of further centres. Phase III would then be independently analysed from Phase II, avoiding Type I error inflation.

The Pros and Cons of the two approaches were summarised:

	Pros	Cons
Phase II/III 'seamless' design	• Removes 'white space' between Phase II and Phase III evaluation • Single Protocol • PII patients 'rolled over' and subsumed into Phase III, providing some savings in N required versus a regular Phase III study • Sample size re-assessment possible	• In continuing recruitment through Phase II, over-recruits patients to arms that will de-selected for Phase III • Phase II to Phase III adaptations need to be defined in advance • Rolling over Phase II patients into Phase III means phases are correlated and no longer independent - therefore Phase II becomes an interim analysis of a larger study and so no longer provides independent supporting evidence to Phase III • To maintain integrity of the overall study and support usage of final results as pivotal evidence, IDMC required to oversee the unblinded Phase II data analysis and review • Sponsor blinded to interim Phase II data and dose selection decision • Type I error inflation – requires alpha adjustment offsetting gains in power or necessitating an increase in N
Two Stage, Single Protocol design	• Phase II and Phase III independent • Limits 'white space' between Phase II and Phase III evaluation • Phase II provides independent, supportive evidence to Phase III • Sponsor has sight of unblinded Phase II data and participates directly in Phase II dose selection and decision making • Sample size reassessment possible • No Type I error inflation	• Does not rollover and re-use Phase II patients in Phase III • Pauses recruitment between Phase II and Phase III • Adds data review and think time between Phase II analysis and Phase III recruitment initiation

Regulatory Feedback

Regulatory authorities reviewed the initial 'seamless' design and raised the concerns identified in advance by KJC Statistics, concerns regarding potential for bias and effective Type I error control. The proposed alternative Two Stage design was then offered and discussed, and was preferred by the authorities as a vehicle for providing pivotal evidence in support of a product license application and labelling.

References

1. Arrowsmith. Phase II failures: 2008–2010. *Nature Rev. Drug Discov*. (2011) 10: 328–329.
2. Arrowsmith. Phase III and submission failures: 2007–2010. *Nature Rev. Drug Discov*. (2011) 10: 87-88.
3. Mehta CR and Pocock SJ. Adaptive Increase in Sample Size when Interim Results are Promising: A Practical Guide with Examples. *Stat in Med,* 2012. Available at.

KJC Statistics Ltd
Sovereign House
Bramhall Village
Cheshire
SK7 1AW
UK
Email: kevin.carroll@kjcstatistics.com
Phone: +44 (0) 161 478 4949
Mobile: +44 (0) 7578 324 824