Supplementary Material for:
"An Optimal Stratified Simon Two-Stage Design"

Deepak Parashar*, Jack Bowden, Colin Starr, Lorenz Wernisch, Adrian Mander

Alternative methods of error rate control

As well as the error framework proposed in our paper, there are a variety of alternative error control methods. Here we describe three. It should be noted that no one method is 'correct'; instead, a framework should be chosen to suit the requirements of a particular trial.

Starting with the proposed framework, one could choose to additionally control the level of 'wrong positives' stemming from rejecting only H_0^+, i.e. concluding that the treatment is effective in the positive subgroup only when it is actually effective in an unselected population. This framework is summarised in the table below. (Note that the probability of a wrong positive occuring by concluding that the treatment is effective in the unselected population when it is only effective in the positive subpopulation is already controlled by the FWER, because $R_1()$ is independent of p^+ and hence has the same value in rows 1 and 3 of the table.)

	Reject both H_0^- and H_0^+	Reject H_0^+
(p_0^-, p_0^+)	$\{ \sum \leq \alpha \}$	
(p_1^-, p_1^-)	$\geq$ power	$\leq \alpha$
(p_0^-, p_1^+)	$\leq \alpha$	$\geq$ power

Weak IO control is defined as controlling the FWER in the weak sense, as well as forcing control of each false positive at the $\alpha/2$ level. Controlling each false positive in this manner automatically means that the FWER, i.e. $R_{123}(p_0^-, p_0^+)$, is controlled by at least α.

	Reject both H_0^- and H_0^+	Reject H_0^+
(p_0^-, p_0^+)	$\leq \alpha/2$	$\leq \alpha/2$
(p_1^-, p_1^-)	$\geq$ power	
(p_0^-, p_1^+)		$\geq$ power

In our design, the false positive control for rejecting both hypotheses is conditional on the false positive control for rejecting the positive-only null hypothesis and hence the trial will have a smaller FWER. A natural choice for individual false positive control would be to have equal weighting as indicated above, but this is in not a unique choice and appropriate variants could be explored.
Strong IO control is defined as weak IO control plus controlling the wrong positive errors as seen below. This option offers the strongest control of all the potential errors.

* *Corresponding author*, D.Parashar@warwick.ac.uk

	Reject both H_0^- and H_0^+	Reject H_0^+
(p_0^-, p_0^+)	$\leq \alpha/2$	$\leq \alpha/2$
(p_1^-, p_1^-)	$\geq$ power	$\leq \alpha$
(p_0^-, p_1^+)	$\leq \alpha$	$\geq$ power

(It is worth remembering that in both IO cases the probability of rejecting both hypotheses in the positive-only case is actually controlled at $\alpha/2$, because $R_1()$ is independent of p^+.)

Optimal designs for each of these error control variants are given below.

Finally, it is possible to set custom limits for any of the combinations of trial outcomes and realities, to suit the circumstances of a particular trial. For example, it may be desirable to place a stricter limit on incorrectly declaring positive-only, to avoid incorrectly excluding a portion of the population from the benefits of a promising treatment. As with all the designs presented here, this should be considered on a case-by-case basis.

Besides offering clinicians and trialists a choice of type I error controls, the methods presented here set the scene for studying different outcomes such as progression-free survival and overall survival, explore *minimax* designs (*i.e.* designs that minimise the total sample size of both stages), and seek a natural extention of the adaptive enrichment design to randomised Phase II trials.

A Selection of Optimal Trial Designs

We present a selection of optimal trial designs for different estimated values of p_0, p_1^- and p_1^+, and different error control schemes. First we present designs for the weak FWER error scheme described in the main paper, with two different values of the minimum power and four values of p_0. Then we pick one combination of power and p_0 and present tables for each of the three error schemes described above.

In every case, $p_0^- = p_0^+$, and $\alpha = 0.05$.

Weak FWER

p_1^-	p_1^+	ESS	Trial Design
0.10	0.10	110.212	(3 2)/(44 34) → (7/104) | (9 4)/(135 53)
0.10	0.15	77.923	(2 2)/(32 21) → (6/67) | (7 3)/(106 29)
0.10	0.20	66.355	(2 1)/(32 11) → (4/34) | (7 2)/(106 11)
0.10	0.25	59.952	(2 1)/(34 8) → (4/29) | (6 2)/(87 9)
0.10	0.30	56.581	(2 1)/(33 6) → (4/22) | (6 2)/(90 9)
0.10	0.40	52.975	(2 1)/(32 4) → (4/19) | (6 2)/(93 6)
0.15	0.15	46.935	(2 1)/(20 12) → (4/43) | (6 2)/(66 21)
0.15	0.20	37.251	(2 1)/(20 9) → (3/25) | (5 2)/(56 14)
0.15	0.25	32.501	(1 1)/(12 7) → (4/28) | (4 2)/(43 11)
0.15	0.30	29.477	(1 1)/(12 5) → (3/20) | (4 2)/(43 10)
0.15	0.40	26.630	(1 1)/(11 4) → (3/14) | (4 2)/(47 6)
0.20	0.20	26.900	(1 1)/(8 9) → (3/25) | (4 2)/(35 14)
0.20	0.25	22.552	(1 1)/(9 7) → (3/21) | (3 2)/(24 11)
0.20	0.30	19.920	(1 1)/(8 5) → (3/19) | (3 2)/(27 11)
0.20	0.40	17.071	(1 1)/(8 4) → (2/8) | (3 2)/(27 7)
0.20	0.50	15.494	(1 1)/(8 3) → (3/9) | (3 1)/(27 3)
0.30	0.30	13.309	(1 1)/(5 5) → (2/13) | (3 2)/(18 9)
0.30	0.40	10.717	(1 1)/(5 4) → (3/12) | (2 1)/(12 4)
0.30	0.50	9.353	(1 1)/(5 3) → (2/7) | (2 2)/(12 4)
0.40	0.40	8.826	(1 1)/(4 4) → (2/8) | (2 1)/(8 4)
0.40	0.50	7.663	(1 1)/(4 3) → (2/6) | (2 1)/(8 3)
0.40	0.60	6.644	(1 1)/(4 2) → (2/6) | (2 1)/(8 2)

$p_0 = 0.03$, power = 0.8

p_1^-	p_1^+	ESS	Trial Design
0.10	0.20	142.883	(6 2)/(82 17) → (9/67) | (17 5)/(225 32)
0.10	0.25	131.624	(5 2)/(72 14) → (9/53) | (16 4)/(208 20)
0.10	0.30	125.693	(5 2)/(72 11) → (6/33) | (16 4)/(208 16)
0.10	0.40	119.500	(5 1)/(72 4) → (5/20) | (16 3)/(208 10)
0.15	0.20	61.347	(3 2)/(30 16) → (7/56) | (8 3)/(76 19)
0.15	0.25	53.052	(3 2)/(30 13) → (6/38) | (7 3)/(67 15)
0.15	0.30	48.090	(2 2)/(22 11) → (5/28) | (7 3)/(67 12)
0.15	0.40	42.620	(2 1)/(21 5) → (4/15) | (7 3)/(70 9)
0.20	0.20	41.757	(2 2)/(15 16) → (5/40) | (7 3)/(54 18)
0.20	0.25	33.850	(2 1)/(15 7) → (4/25) | (7 2)/(54 11)
0.20	0.30	29.744	(2 1)/(15 6) → (4/21) | (5 3)/(40 13)
0.20	0.40	25.491	(2 1)/(15 4) → (3/13) | (5 3)/(40 9)
0.20	0.50	23.374	(2 1)/(15 3) → (3/11) | (5 2)/(40 4)
0.30	0.30	18.353	(1 1)/(5 5) → (4/23) | (4 2)/(23 10)
0.30	0.40	14.619	(1 1)/(6 5) → (3/11) | (3 2)/(15 6)
0.30	0.50	12.148	(1 1)/(5 3) → (3/10) | (3 2)/(18 5)
0.40	0.40	10.836	(1 1)/(4 4) → (3/13) | (3 2)/(11 5)
0.40	0.50	8.921	(1 1)/(4 3) → (2/7) | (3 2)/(11 4)
0.40	0.60	7.527	(1 1)/(4 2) → (2/5) | (3 1)/(11 2)

$p_0 = 0.05$, power = 0.8

p_1^-	p_1^+	ESS	Trial Design
0.15	0.40	192.564	(13 3)/(109 12) → (12/44) | (43 6)/(340 16)
0.20	0.20	121.556	(6 6)/(40 42) → (19/122) | (22 10)/(141 59)
0.20	0.25	89.562	(6 4)/(40 24) → (13/70) | (18 7)/(117 33)
0.20	0.30	76.142	(6 3)/(40 15) → (11/51) | (17 5)/(111 21)
0.20	0.40	64.554	(5 2)/(34 8) → (8/30) | (16 5)/(106 14)
0.20	0.50	59.503	(5 2)/(35 6) → (6/18) | (15 4)/(97 9)
0.30	0.30	35.920	(3 2)/(14 10) → (8/38) | (9 4)/(42 18)
0.30	0.40	26.672	(2 2)/(10 8) → (7/25) | (7 3)/(33 9)
0.30	0.50	22.583	(2 1)/(10 3) → (5/14) | (7 3)/(33 8)
0.40	0.40	17.666	(2 1)/(7 4) → (4/14) | (7 3)/(25 11)
0.40	0.50	13.937	(2 1)/(7 3) → (4/11) | (5 2)/(18 6)
0.40	0.60	11.913	(2 1)/(7 2) → (3/8) | (5 2)/(18 4)

$p_0 = 0.1$, power = 0.8

p_1^-	p_1^+	ESS	Trial Design
0.30	0.40	112.038	(13 6)/(53 20) → (22/69) | (43 11)/(168 32)
0.30	0.50	93.955	(13 4)/(53 11) → (16/41) | (39 7)/(153 16)
0.40	0.40	50.068	(6 4)/(19 14) → (15/48) | (19 9)/(60 25)
0.40	0.50	36.364	(5 3)/(16 8) → (12/32) | (15 7)/(49 17)
0.40	0.60	30.202	(4 2)/(14 5) → (9/20) | (13 4)/(40 7)

$p_0 = 0.2$, power = 0.8

p_1^-	p_1^+	ESS	Trial Design
0.10	0.10	161.341	(4 3)/(68 56) → (9/147) | (12 4)/(187 66)
0.10	0.15	114.629	(3 2)/(56 31) → (6/70) | (9 3)/(142 32)
0.10	0.20	98.707	(3 2)/(56 22) → (5/48) | (8 3)/(129 25)
0.10	0.25	90.217	(3 1)/(55 11) → (4/32) | (8 3)/(132 20)
0.10	0.30	85.340	(3 1)/(55 9) → (4/27) | (8 2)/(132 11)
0.10	0.40	79.698	(3 1)/(56 7) → (4/20) | (7 3)/(115 11)
0.15	0.15	69.317	(2 2)/(27 28) → (5/61) | (6 3)/(68 33)
0.15	0.20	55.224	(2 1)/(27 13) → (4/37) | (6 2)/(68 18)
0.15	0.25	48.078	(2 1)/(26 10) → (3/23) | (6 2)/(72 15)
0.15	0.30	43.974	(2 1)/(27 8) → (3/22) | (5 2)/(58 11)
0.15	0.40	40.012	(2 1)/(27 6) → (5/25) | (4 2)/(48 9)
0.20	0.20	40.946	(2 1)/(19 12) → (3/29) | (6 3)/(53 27)
0.20	0.25	34.817	(1 1)/(12 10) → (4/30) | (4 2)/(37 15)
0.20	0.30	30.658	(1 1)/(12 8) → (3/20) | (4 2)/(37 12)
0.20	0.40	26.686	(1 1)/(12 6) → (2/10) | (4 2)/(37 8)
0.20	0.50	23.600	(1 1)/(12 5) → (4/17) | (3 2)/(29 6)
0.30	0.30	20.449	(1 1)/(7 7) → (3/24) | (3 2)/(21 13)
0.30	0.40	16.043	(1 1)/(7 5) → (3/17) | (3 1)/(21 5)
0.30	0.50	14.324	(1 1)/(7 4) → (3/11) | (3 1)/(21 4)
0.40	0.40	12.376	(1 1)/(5 5) → (2/11) | (3 2)/(14 8)
0.40	0.50	10.451	(1 1)/(5 4) → (3/12) | (2 1)/(10 4)
0.40	0.60	8.955	(1 1)/(5 3) → (2/7) | (2 1)/(10 3)

$p_0 = 0.03$, power = 0.9

p_1^-	p_1^+	ESS	Trial Design
0.15	0.15	118.533	(4 3)/(45 37) → (11/115) | (13 5)/(130 52)
0.15	0.20	90.147	(4 2)/(45 22) → (7/57) | (11 4)/(113 32)
0.15	0.25	77.698	(3 2)/(36 17) → (6/42) | (10 4)/(105 26)
0.15	0.30	70.472	(3 2)/(37 15) → (6/37) | (9 3)/(93 15)
0.15	0.40	63.025	(3 1)/(37 6) → (4/18) | (9 3)/(93 13)
0.20	0.20	61.451	(3 2)/(26 20) → (6/52) | (9 4)/(74 31)
0.20	0.25	50.456	(2 2)/(19 16) → (5/35) | (8 3)/(67 20)
0.20	0.30	44.561	(2 2)/(20 13) → (6/39) | (6 3)/(50 17)
0.20	0.40	37.600	(2 1)/(19 6) → (4/20) | (6 3)/(53 12)
0.20	0.50	34.113	(2 1)/(19 5) → (4/14) | (6 2)/(53 6)
0.30	0.30	27.224	(2 1)/(12 8) → (4/24) | (5 2)/(30 11)
0.30	0.40	21.584	(2 1)/(12 5) → (3/15) | (5 2)/(30 9)
0.30	0.50	18.911	(1 1)/(8 4) → (3/11) | (4 2)/(23 7)
0.40	0.40	15.535	(1 1)/(5 5) → (3/16) | (4 2)/(18 8)
0.40	0.50	12.703	(1 1)/(5 4) → (3/11) | (3 2)/(14 7)
0.40	0.60	11.248	(1 1)/(5 3) → (3/10) | (3 2)/(14 5)

$p_0 = 0.05$, power = 0.9

p_1^-	p_1^+	ESS	Trial Design
0.20	0.20	174.609	(10 7)/(71 54) → (22/148) | (29 14)/(189 94)
0.20	0.25	130.040	(8 5)/(59 34) → (16/93) | (24 8)/(160 43)
0.20	0.30	110.942	(8 4)/(59 23) → (12/62) | (22 7)/(148 33)
0.20	0.40	94.411	(7 3)/(53 14) → (11/43) | (20 5)/(136 18)
0.20	0.50	86.975	(7 2)/(53 8) → (7/24) | (20 4)/(136 11)
0.30	0.30	52.085	(4 3)/(21 18) → (9/45) | (13 4)/(65 20)
0.30	0.40	39.197	(3 2)/(17 9) → (7/29) | (10 4)/(51 16)
0.30	0.50	32.980	(3 2)/(17 7) → (6/21) | (9 3)/(46 9)
0.40	0.40	25.969	(2 2)/(9 9) → (6/25) | (8 3)/(30 11)
0.40	0.50	20.801	(2 1)/(9 4) → (4/13) | (7 3)/(27 10)
0.40	0.60	18.108	(2 1)/(9 3) → (5/15) | (5 3)/(20 8)

$p_0 = 0.1$, power = 0.9

p_1^-	p_1^+	ESS	Trial Design
0.30	0.50	136.533	(18 5)/(78 16) → (19/53) | (50 9)/(200 23)
0.40	0.40	71.389	(8 7)/(28 26) → (18/60) | (24 11)/(77 34)
0.40	0.50	52.331	(8 4)/(28 13) → (11/30) | (21 6)/(68 16)
0.40	0.60	44.086	(6 3)/(22 8) → (12/29) | (18 5)/(60 11)

$p_0 = 0.2$, power = 0.9

Strong FWER

p_1^-	p_1^+	ESS	Trial Design
0.10	0.10	151.230	(4 2)/(75 34) → (7/106) | (12 4)/(198 51)
0.10	0.15	90.118	(2 3)/(31 31) → (25/214) | (7 3)/(111 31)
0.10	0.25	60.375	(2 1)/(34 7) → (5/35) | (6 2)/(87 11)
0.15	0.15	65.666	(2 1)/(30 12) → (4/43) | (7 3)/(85 29)
0.15	0.25	36.582	(1 2)/(11 12) → (19/108) | (4 2)/(47 12)
0.15	0.35	28.408	(1 1)/(11 5) → (4/18) | (4 2)/(47 8)
0.25	0.25	26.618	(1 1)/(12 7) → (4/28) | (3 2)/(24 11)
0.25	0.40	16.388	(1 1)/(6 4) → (11/36) | (3 1)/(25 5)

$p_0 = 0.03$, power = 0.8

Weak IO

p_1^-	p_1^+	ESS	Trial Design
0.10	0.10	113.195	(3 2)/(43 36) → (8/117) | (8 4)/(129 49)
0.10	0.15	78.493	(2 1)/(34 13) → (5/54) | (7 3)/(100 27)
0.10	0.25	61.963	(2 1)/(34 7) → (3/21) | (7 2)/(100 11)
0.15	0.15	49.126	(2 1)/(20 13) → (5/52) | (5 2)/(56 18)
0.15	0.25	32.796	(1 1)/(14 7) → (3/21) | (4 2)/(39 11)
0.15	0.35	28.790	(1 1)/(14 5) → (2/10) | (4 2)/(39 7)
0.25	0.25	18.552	(1 1)/(6 7) → (3/22) | (3 2)/(23 9)
0.25	0.40	13.454	(1 1)/(6 4) → (2/9) | (3 2)/(23 5)

$p_0 = 0.03$, power $= 0.8$

Strong IO

p_1^-	p_1^+	ESS	Trial Design
0.10	0.10	153.387	(4 2)/(76 35) → (8/117) | (10 5)/(167 66)
0.10	0.15	92.085	(2 3)/(31 30) → (25/215) | (8 3)/(125 31)
0.10	0.25	62.810	(2 1)/(34 7) → (4/27) | (7 2)/(100 11)
0.15	0.15	66.495	(2 2)/(30 21) → (5/58) | (6 3)/(76 27)
0.15	0.25	38.566	(1 2)/(11 12) → (15/83) | (5 2)/(57 12)
0.15	0.35	30.016	(1 1)/(14 5) → (5/23) | (4 2)/(39 7)
0.25	0.25	26.878	(1 1)/(11 7) → (3/22) | (4 2)/(32 10)
0.25	0.40	16.388	(1 1)/(6 4) → (11/36) | (3 1)/(25 5)

$p_0 = 0.03$, power $= 0.8$