Supplementary Material for "Bayesian Interim Analysis in Basket Trials"

Cheng Huang, Chenghao Chu, Yimeng Lu, Bingming Yi, and Ming-Hui Chen

S.1 Supplementary Results for Models without covariates

We consider the case that the outcomes are not influenced by any covariates. Following notation in the main paper, the outcomes are modeled as

$$\begin{aligned} Y_{cki} &= \beta_{k0} + \epsilon_{cki}, \\ \text{with} \quad & i = 1, \ldots, n_{ck}, \\ Y_{tki'} &= \beta_{k0} + \tau_k + \epsilon_{tki'}, \\ \text{with} \quad & i' = 1, \ldots, n_{tk}, \end{aligned} \tag{S.1}$$

where $k = 1, \ldots, 4$. The number of subjects in control and treatment arm are assumed to be the same ($n_{ck} = n_{tk} \doteq n_k$). For baskets 1 to 4, they are assumed to be $n_1 = 30$, $n_2 = 30$, $n_3 = 20$, and $n_4 = 20$, and the same sample sizes are assumed for each stage (if the basket is available for analysis in that stage). Let β_{k0} denotes the intercept in the control arm for the k-th basket. τ_k denotes the difference in the intercept between treatment and control for the k-th basket. Each of the error terms $\epsilon_{cki}(\epsilon_{tki'}) \sim N(0, 1)$. To evaluate the performance of these various methods, we generate data assuming the control effects to be $(0, 0, 0, 0)$. We consider 6 scenarios of the true treatment effects, $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_4)$, which are summarized in Table S.1. Cut-off values are summarized in Table S.2. Note that the first scenario corresponds to the global null (GN).

Simulation results for the one-stage design are reported in Table S.3; simulation results for the two-stage design with efficacy and futility stopping are reported in Table S.4; simulation results

Table S.1: True effects in the treatment arm

Scenarios	τ_1	τ_2	τ_3	τ_4
1 GN	0	0	0	0
2	0.4	0	0	0
3	0.4	0.4	0	0
4	0.4	0.4	0.4	0
5	0.4	0.4	0.4	0.4
6	0	0.2	0.4	0.6

Table S.2: Value of q_1

Method	One-Stage	Two-Stage, both stopping ($\eta = \frac{1}{2}$)	Two-Stage, futility stopping only ($\eta = 0$)
Models without covariates			
SEP	0.9762	0.9629	0.9400
BHM	0.9520	0.9320	0.9180
CBHM	0.9760	0.9641	0.9400
MFM	0.9080	0.8800	0.8700

for the two-stage design with futility stopping only are reported in Table S.5; The graphs for comparing the performance of these methods are shown in Figure S.1.

Table S.3: Performance of SEP, BHM, CBHM and MFM without interim analysis (models without covariates)

Scn	Method	FWER	P1	P2	P3	% Reject				100 x RMSE			
1						1	2	3	4	1	2	3	4
	SEP	10.0	0.0	0.0	0.0	2.7	2.6	2.8	2.4	25.8	25.9	31.5	30.8
	BHM	10.0	0.0	0.0	0.0	4.1	3.6	3.2	2.0	18.4	18.3	19.7	19.2
	CBHM	10.1	0.0	0.0	0.0	2.8	2.8	3.2	3.5	19.9	20.2	21.7	22.2
	MFM	10.0	0.0	0.0	0.0	3.8	3.6	2.5	2.1	13.3	13.3	13.0	12.7
2						[1]	2	3	4	[1]	2	3	4
	SEP	7.4	34.1	31.3	31.3	34.1	2.3	2.5	2.8	25.9	25.7	31.3	31.9
	BHM	13.5	34.2	26.0	26.0	34.2	6.5	6.1	4.2	22.9	19.6	21.4	21.4
	CBHM	11.9	30.4	24.8	24.8	30.4	4.4	5.2	5.7	23.2	20.9	23.0	23.4
	MFM	12.6	43.0	34.3	34.3	43.0	5.6	4.3	4.5	22.6	14.4	14.3	14.5
3						[1]	[2]	3	4	[1]	[2]	3	4
	SEP	5.0	56.3	53.5	11.0	33.2	34.8	2.4	2.7	25.8	25.8	31.6	31.9
	BHM	15.4	63.4	50.2	18.6	43.1	45.0	10.4	7.0	20.4	20.0	24.3	23.9
	CBHM	14.1	54.2	43.6	9.2	35.2	33.0	8.5	9.0	21.5	22.1	24.7	25.0
	MFM	12.2	72.6	61.5	25.5	51.8	52.6	6.6	6.7	20.3	20.0	15.8	15.9
4						[1]	[2]	[3]	4	[1]	[2]	[3]	4
	SEP	2.6	68.0	66.3	3.0	35.3	34.7	24.6	2.6	25.9	25.9	31.6	31.8
	BHM	11.2	79.8	69.4	12.8	49.3	50.3	44.8	11.2	19.1	18.9	20.5	27.0
	CBHM	13.4	71.2	58.9	5.7	38.3	37.7	40.7	13.4	20.6	21.0	22.5	26.9
	MFM	9.2	85.5	76.7	17.5	59.3	59.1	45.4	9.2	18.6	18.5	20.4	17.3
5						[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]
	SEP	0.0	76.1	76.1	0.7	34.3	34.4	25.0	24.8	25.7	25.6	31.9	31.6
	BHM	0.0	89.1	89.1	10.2	53.8	54.6	52.3	43.2	18.2	18.2	19.9	19.7
	CBHM	0.0	80.8	80.8	10.0	41.2	42.9	46.5	39.0	19.9	20.2	21.7	22.2
	MFM	0.0	92.4	92.4	18.1	64.1	64.5	52.7	52.0	17.1	17.1	18.8	18.7
6						1	[2]	[3]	[4]	1	[2]	[3]	[4]
	SEP	2.6	65.3	63.6	1.6	2.6	12.9	25.2	47.3	25.6	25.9	31.0	31.9
	BHM	13.2	70.0	58.0	7.1	13.2	29.9	42.2	44.8	22.9	19.5	20.9	26.8
	CBHM	8.2	66.8	59.6	3.8	8.2	20.0	38.4	42.8	22.5	20.8	23.1	26.9
	MFM	9.7	77.9	68.9	9.5	9.7	28.0	42.1	63.4	16.0	15.2	21.0	31.7

Table S.4: Performance of SEP, BHM, CBHM and MFM, two-stage design and $\eta = \frac{1}{2}$ (models without covariates)

Scn	Method	FWER	P1	P2	P3	% Reject				100 x RMSE				Average Enrollment			
1						1	2	3	4	1	2	3	4	1	2	3	4
	SEP	10.0	0.0	0.0	0.0	2.6	2.7	2.8	2.3	25.2	25.3	30.7	29.9	30.6	30.6	20.4	20.4
	BHM	10.0	0.0	0.0	0.0	3.5	3.6	3.0	2.1	17.9	17.9	19.4	18.9	30.9	31.1	20.6	20.5
	CBHM	10.0	0.0	0.0	0.0	2.7	2.6	3.1	3.4	19.4	19.8	21.4	21.9	30.9	30.8	20.4	20.5
	MFM	9.9	0.0	0.0	0.0	3.4	3.5	2.5	2.1	12.9	13.0	12.9	12.5	31.1	31.2	20.6	20.6
2						[1]	2	3	4	[1]	2	3	4	[1]	2	3	4
	SEP	7.1	39.1	36.2	36.2	39.1	2.1	2.3	2.8	26.2	25.0	30.5	31.2	33.3	30.6	20.4	20.4
	BHM	13.3	42.2	32.5	32.5	42.2	6.2	5.4	4.3	23.3	18.8	20.5	20.8	33.7	31.8	21.1	20.8
	CBHM	11.3	40.1	33.0	33.0	40.1	4.1	4.8	5.3	23.8	20.3	22.5	22.9	34.3	31.1	20.6	20.6
	MFM	11.4	50.0	41.6	41.6	50.0	4.8	3.7	4.1	22.3	13.8	13.7	14.1	35.9	31.7	21.0	20.9
3						[1]	[2]	3	4	[1]	[2]	3	4	[1]	[2]	3	4
	SEP	5.1	64.0	60.8	15.2	39.2	40.8	2.5	2.7	26.0	26.0	30.9	31.2	33.4	33.3	20.4	20.4
	BHM	13.5	73.1	60.6	27.6	53.4	55.6	8.3	6.7	20.8	20.5	22.8	22.8	34.0	35.0	21.8	21.2
	CBHM	13.8	66.5	54.4	17.6	46.6	44.6	8.4	9.0	22.1	22.7	24.0	24.4	34.5	34.4	20.8	20.7
	MFM	11.0	78.9	68.7	32.4	58.2	59.5	5.9	5.8	20.1	19.7	15.1	15.1	36.1	36.0	21.4	21.4
4						[1]	[2]	[3]	4	[1]	[2]	[3]	4	[1]	[2]	[3]	4
	SEP	2.7	74.4	72.4	4.4	40.4	40.3	28.6	2.7	26.1	26.2	31.5	31.1	33.3	33.4	21.9	20.4
	BHM	10.6	86.1	75.8	23.1	62.8	63.4	51.7	10.6	19.5	19.4	21.6	25.5	34.0	34.7	23.7	21.6
	CBHM	14.0	81.1	67.5	12.2	53.8	51.3	46.5	14.0	21.1	21.6	23.0	26.3	34.7	34.5	22.0	20.8
	MFM	7.8	89.6	82.0	22.6	65.0	64.6	50.3	7.8	18.4	18.4	20.7	16.2	35.9	35.7	24.3	21.6
5						[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]
	SEP	0.0	82.1	82.1	1.3	39.6	40.2	29.3	28.6	25.9	25.8	31.8	31.5	33.2	33.5	22.0	21.9
	BHM	0.0	93.1	93.1	26.4	69.0	70.7	59.6	53.6	18.7	18.6	21.0	20.5	33.7	34.1	23.4	23.2
	CBHM	0.0	90.2	90.2	25.4	61.8	59.9	56.4	55.3	20.4	20.7	22.1	22.7	34.5	34.3	21.7	21.9
	MFM	0.0	94.9	94.9	21.1	68.9	69.2	56.0	55.9	17.2	17.1	19.2	19.1	35.7	35.4	24.2	24.3
6						1	[2]	[3]	[4]	1	[2]	[3]	[4]	1	[2]	[3]	[4]
	SEP	2.6	71.3	69.4	2.2	2.6	14.4	29.0	53.4	25.0	25.3	30.9	32.3	30.6	31.8	21.9	22.3
	BHM	11.2	78.4	67.7	12.8	11.2	32.1	47.8	63.5	21.7	19.6	21.8	26.4	32.2	34.1	23.4	23.7
	CBHM	7.4	76.6	69.6	7.9	7.4	22.2	43.3	62.2	21.3	20.7	23.6	26.8	31.9	33.3	22.1	22.2
	MFM	8.2	83.5	75.7	11.9	8.2	29.1	46.8	71.2	15.1	15.4	21.3	30.1	32.5	34.6	24.1	24.2

Table S.5: Performance of SEP, BHM, CBHM and MFM, two-stage design and $\eta = 0$ (models without covariates)

Scn	Method	FWER	P1	P2	P3	% Reject				100 x RMSE				Average Enrollment			
1						1	2	3	4	1	2	3	4	1	2	3	4
	SEP	10.0	0.0	0.0	0.0	2.5	2.8	2.8	2.4	23.3	23.7	28.5	28.1	31.9	31.9	21.3	21.2
	BHM	9.9	0.0	0.0	0.0	3.0	3.5	3.0	2.2	17.0	17.1	18.6	18.4	32.1	32.3	21.3	20.9
	CBHM	10.0	0.0	0.0	0.0	2.9	2.9	2.9	3.0	18.7	19.2	20.4	20.9	31.8	31.7	21.2	21.3
	MFM	10.0	0.0	0.0	0.0	3.2	3.6	2.8	2.3	12.5	12.7	12.7	12.4	32.0	31.9	21.0	20.9
2						[1]	2	3	4	[1]	2	3	4	[1]	2	3	4
	SEP	7.6	46.9	43.3	43.3	46.9	2.4	2.5	2.9	22.7	23.6	28.5	28.9	45.2	31.8	21.2	21.4
	BHM	14.0	37.6	31.4	31.4	37.6	6.3	5.6	4.7	24.4	17.4	19.1	19.7	43.5	33.8	22.3	21.8
	CBHM	10.5	39.7	34.4	34.4	39.7	3.8	4.2	4.1	24.6	19.3	20.9	21.4	43.1	32.5	21.9	21.9
	MFM	12.4	50.0	41.7	41.7	50.0	5.1	4.3	4.5	22.5	13.4	13.5	13.8	46.1	32.8	21.6	21.6
3						[1]	[2]	3	4	[1]	[2]	3	4	[1]	[2]	3	4
	SEP	5.1	72.2	68.7	21.6	46.5	48.3	2.7	2.5	22.5	22.2	28.7	29.1	45.1	45.5	21.3	21.3
	BHM	14.6	64.0	53.6	10.3	36.5	41.2	8.8	7.5	21.7	21.2	20.7	20.9	46.2	47.8	23.6	22.9
	CBHM	10.3	61.1	53.6	7.5	39.8	30.8	5.7	5.6	22.8	23.3	21.5	22.0	44.7	43.6	22.8	22.7
	MFM	12.1	78.5	67.6	29.1	56.4	57.4	6.5	6.5	20.2	19.8	14.8	14.9	48.7	49.0	22.3	22.2
4						[1]	[2]	[3]	4	[1]	[2]	[3]	4	[1]	[2]	[3]	4
	SEP	2.8	82.3	80.1	7.4	47.5	47.9	34.6	2.8	22.4	22.4	26.9	29.0	45.3	45.5	28.0	21.3
	BHM	11.1	78.7	69.0	8.5	39.4	41.7	44.4	11.1	20.1	20.0	21.3	22.4	47.9	48.9	31.9	24.2
	CBHM	7.2	71.9	65.4	4.1	37.8	31.8	33.7	7.2	21.7	22.2	22.9	23.0	45.6	44.9	30.3	23.7
	MFM	9.0	89.1	80.4	22.1	62.4	62.2	52.4	9.0	18.4	18.4	20.5	15.9	50.8	50.7	31.5	23.0
5						[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]
	SEP	0.0	88.4	88.4	2.9	46.8	47.8	35.0	34.8	22.3	22.2	27.3	26.9	45.0	45.3	28.0	27.9
	BHM	0.0	88.0	88.0	4.4	42.4	41.6	49.7	45.1	19.0	18.9	20.3	19.9	49.0	49.4	33.0	31.8
	CBHM	0.0	76.3	76.3	2.5	35.6	32.6	36.4	28.9	20.8	21.3	21.7	23.3	46.2	46.0	31.3	29.8
	MFM	0.0	94.7	94.7	22.8	66.9	67.1	58.0	58.3	17.0	17.0	18.9	18.8	52.1	52.1	32.8	32.7
6						1	[2]	[3]	[4]	1	[2]	[3]	[4]	1	[2]	[3]	[4]
	SEP	2.7	78.3	76.2	3.7	2.7	17.2	34.4	60.9	23.3	21.8	26.3	28.4	32.0	36.9	27.9	32.6
	BHM	10.0	70.9	62.5	6.0	10.0	28.2	40.5	37.6	19.0	17.9	22.1	27.2	36.0	42.4	31.0	32.0
	CBHM	5.6	68.0	63.6	2.7	5.6	20.7	36.6	33.8	19.3	19.5	23.6	27.8	34.3	38.9	29.9	30.6
	MFM	8.4	83.5	75.7	11.6	8.4	30.4	48.5	65.5	14.3	15.5	21.4	28.3	34.6	41.2	30.7	34.9

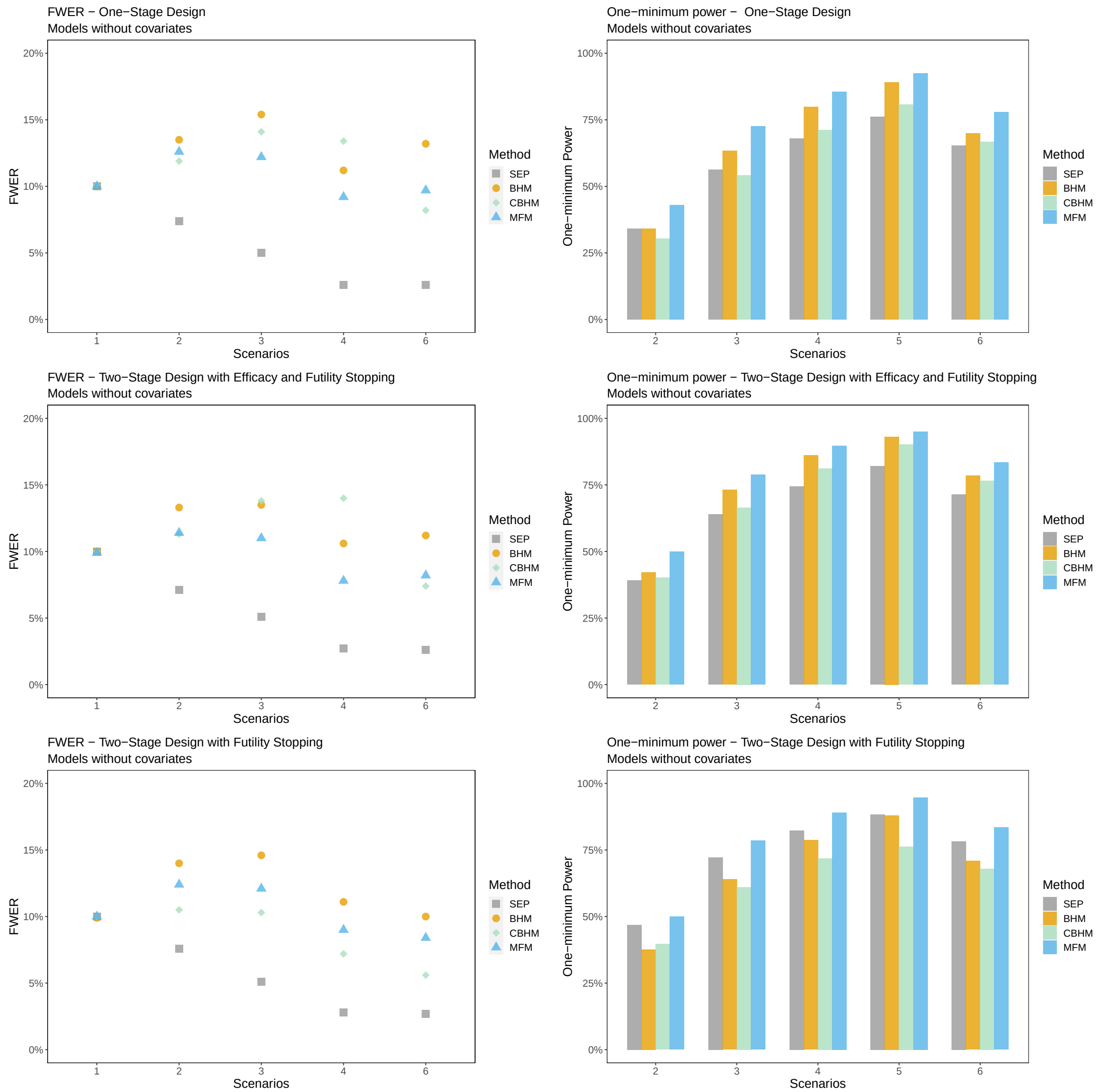

Figure S.1: One-minimum Power and FWER for Bayesian Methods, models without covariates

S.2 Sensitivity Analysis of MFM

The following tables display the sensitivity analyses of the MFM performance if various hyper-parameters ξ and ϕ are given under different scenarios and in the three study designs. As shown in these tables, the hyper-parameters that result in higher powers also tend to increase the FWER inflation. Well-chosen ξ and ϕ will balance power and FWER optimally. After comparing the results, we pick $\phi = 0.1$ and $\xi = 0.16$. MFM performs consistently in models with or without covariates.

Table S.6: Sensitivity Analysis, one-stage design (models without covariates)

		Scn 2		Scn 3		Scn 4		Scn 5	Scn 6	
$\phi(\tau)$	$\xi(\theta)$	FWER	P1	FWER	P1	FWER	P1	P1	FWER	P1
0.01	0.064	15.2	47.4	19.1	80.2	17.9	91.2	96.2	17.9	84.0
0.01	0.16	14.1	43.6	17.2	74.6	15.8	86.8	93.6	16.3	77.5
0.01	0.4	13.9	41.4	17.0	72.0	14.8	84.2	91.8	15.6	73.5
0.1	0.064	12.9	48.1	13.2	78.3	10.0	90.0	95.4	10.4	84.1
0.1	0.16	12.6	43.0	12.2	72.6	9.2	85.5	92.4	9.7	77.9
0.1	0.4	11.8	39.6	11.7	69.0	8.6	82.0	90.0	9.1	73.5
1	0.064	15.1	45.4	16.5	74.9	13.8	89.0	95.3	11.9	83.2
1	0.16	13.6	39.6	14.4	69.2	11.6	83.9	92.3	10.1	76.5
1	0.4	12.9	36.8	13.7	65.7	10.6	80.2	89.4	9.4	72.7

Table S.7: Sensitivity Analysis, two-stage design and $\eta = 1/2$ (models without covariates)

		Scn 2		Scn 3		Scn 4		Scn 5	Scn 6	
$\phi(\tau)$	$\xi(\theta)$	FWER	P1	FWER	P1	FWER	P1	P1	FWER	P1
0.01	0.064	15.0	56.9	16.4	85.7	13.5	94.1	97.8	14.4	89.4
0.01	0.16	14.4	50.7	15.4	80.7	12.6	90.8	95.8	13.4	83.3
0.01	0.4	14.4	48.5	15.4	78.8	12.8	89.1	94.4	13.2	80.2
0.1	0.064	12.0	55.5	11.4	84.2	8.3	93.4	97.1	8.7	89.2
0.1	0.16	11.4	50.0	11.0	78.9	7.8	89.6	94.9	8.2	83.5
0.1	0.4	11.2	47.1	10.3	75.8	7.5	87.1	93.2	7.9	80.1
1	0.064	14.3	53.1	14.6	82.5	11.7	93.1	97.3	10.1	89.7
1	0.16	13.0	47.6	13.3	77.1	10.5	89.3	95.1	9.2	84.1
1	0.4	12.5	45.1	12.8	74.3	9.9	86.9	93.4	8.4	81.1

Table S.8: Sensitivity Analysis, two-stage design and $\eta = 0$ (models without covariates)

		Scn 2		Scn 3		Scn 4		Scn 5	Scn 6	
$\phi(\tau)$	$\xi(\theta)$	FWER	P1	FWER	P1	FWER	P1	P1	FWER	P1
0.01	0.064	16.7	56.9	20.4	85.6	18.9	94.1	97.8	17.1	89.5
0.01	0.16	15.8	50.9	19.1	80.9	17.0	90.9	95.9	15.8	83.5
0.01	0.4	15.6	48.4	18.8	78.8	16.5	89.1	94.4	15.1	80.2
0.1	0.064	13.0	55.2	12.4	82.5	9.6	92.1	96.5	8.6	88.2
0.1	0.16	12.4	50.0	12.1	78.5	9.0	89.1	94.7	8.4	83.5
0.1	0.4	12.2	47.8	11.5	75.5	8.6	87.0	93.0	8.0	80.2
1	0.064	14.6	51.5	15.8	76.5	13.4	88.6	94.7	10.5	84.2
1	0.16	13.4	46.0	14.3	72.2	11.7	85.0	92.5	8.9	79.0
1	0.4	13.3	43.6	13.6	69.7	10.9	82.8	90.6	8.5	76.0

Table S.9: Sensitivity Analysis, one-stage design (models adjusting for covariates)

		Scn 2		Scn 3		Scn 4		Scn 5	Scn 6	
$\phi(\tau)$	$\xi(\theta)$	FWER	P1	FWER	P1	FWER	P1	P1	FWER	P1
0.01	0.064	15.9	45.4	18.4	72.9	15.0	83.1	91.4	15.6	72.6
0.01	0.16	14.1	41.2	16.1	70.4	13.4	81.6	88.5	13.5	68.3
0.01	0.4	14.6	40.5	16.8	70.5	13.4	79.8	86.8	13.7	66.5
0.1	0.064	13.3	43.1	12.8	71.1	9.2	81.2	89.8	9.2	73.3
0.1	0.16	11.5	40.7	11.7	68.9	8.3	78.7	87.6	8.2	69.3
0.1	0.4	11.5	40.2	11.8	67.4	8.4	78.3	86.7	8.2	66.7
1	0.064	13.6	37.1	13.5	65.8	10.1	77.7	88.6	8.4	70.8
1	0.16	13.5	37.8	13.7	66.1	9.7	77.3	86.8	8.2	68.5
1	0.4	13.0	34.2	12.9	62.8	9.6	75.8	84.8	7.7	66.5

Table S.10: Sensitivity Analysis, two-stage design and $\eta = 1/2$ (models adjusting for covariates)

		Scn 2		Scn 3		Scn 4		Scn 5	Scn 6	
$\phi(\tau)$	$\xi(\theta)$	FWER	P1	FWER	P1	FWER	P1	P1	FWER	P1
0.01	0.064	14.7	50.5	15.9	78.0	11.5	87.3	93.4	11.9	78.8
0.01	0.16	14.0	48.5	15.9	76.0	9.8	85.9	92.1	11.7	76.5
0.01	0.4	14.3	49.7	15.9	77.1	10.7	84.6	92.1	12.8	76.3
0.1	0.064	11.5	49.6	10.8	76.8	7.7	86.6	92.2	7.4	79.5
0.1	0.16	10.9	46.4	9.6	74.4	6.3	83.7	91.0	6.3	75.5
0.1	0.4	11.7	46.1	9.9	73.4	7.0	82.8	90.3	7.1	74.2
1	0.064	13.4	47.9	13.2	75.4	9.7	86.4	92.3	7.6	80.4
1	0.16	12.4	43.0	11.9	72.5	8.7	83.4	90.4	6.6	76.7
1	0.4	12.9	44.0	12.3	72.1	9.1	82.2	90.2	6.8	75.3

Table S.11: Sensitivity Analysis, two-stage design and $\eta = 0$ (models adjusting for covariates)

		Scn 2		Scn 3		Scn 4		Scn 5	Scn 6	
$\phi(\tau)$	$\xi(\theta)$	FWER	P1	FWER	P1	FWER	P1	P1	FWER	P1
0.01	0.064	16.0	51.0	18.8	78.3	15.5	87.6	93.4	14.6	79.0
0.01	0.16	14.6	48.5	19.3	75.9	14.0	85.9	92.1	13.9	76.5
0.01	0.4	15.2	49.7	19.3	77.1	14.7	84.6	92.1	14.5	76.3
0.1	0.064	12.2	49.7	12.5	75.3	7.5	86.0	91.5	6.9	78.0
0.1	0.16	11.9	47.8	11.5	74.1	7.1	85.1	90.9	6.6	77.2
0.1	0.4	12.5	46.2	11.9	72.7	7.5	83.4	91.2	6.6	75.5
1	0.064	13.5	46.5	14.2	69.9	9.8	81.1	90.0	7.2	73.9
1	0.16	12.8	43.5	13.8	67.7	9.1	81.5	88.4	7.0	72.2
1	0.4	12.0	40.5	13.5	66.6	8.5	78.6	86.4	5.9	69.4

S.3 Derivation of MFM sampling algorithm formula

The derivation for $m(\mathcal{D}|\sigma^2, \theta_k)$ in the MFM sampling algorithm for models without covariates is given as follows:

$$
\begin{aligned}
m(\mathcal{D}|\sigma^2, \theta_k) = &\int [\mathcal{L}(\bar{y}_{ck}|\theta_k, \sigma^2)\mathcal{L}(\bar{y}_{tk}, z_k|\theta_k, \sigma^2, \tau_k)\pi(\tau_{z_k}|\sigma^2)]d\tau_{z_k} \\
=&\mathcal{L}(\bar{y}_{ck}|\theta_k, \sigma^2)\frac{1}{(\sqrt{2\pi\sigma^2})^{n_{tk}}}\frac{1}{\sqrt{2\pi\phi\sigma^2}} \\
&\int \exp\Big[-\frac{(n_{tk}-1)S_{tk}^2 + n_{tk}(\bar{y}_{tk} - \theta_k - \tau_{z_k})^2}{2\sigma^2}\Big]\exp\Big(-\frac{\tau_{z_k}^2}{2\phi\sigma^2}\Big)d\tau_{z_k} \\
=&\mathcal{L}(\bar{y}_{ck}|\theta_k, \sigma^2)\frac{1}{(\sqrt{2\pi\sigma^2})^{n_{tk}}}\frac{1}{\sqrt{\phi n_{tk}+1}}\exp\Big[-\frac{(n_{tk}-1)S_{tk}^2}{2\sigma^2} - \frac{n_{tk}(\bar{y}_{tk}-\theta_k)^2}{2\sigma^2(n_{tk}\phi+1)}\Big].
\end{aligned}
\tag{S.2}
$$

The derivation for $m(\mathcal{D}|\boldsymbol{\beta}_k, \sigma^2)$ in the MFM sampling algorithm for models adjusting for covariates is given as follows:

$$
\begin{aligned}
m(\mathcal{D}|\boldsymbol{\beta}_k, \sigma^2) = &\int \mathcal{L}(\boldsymbol{X}_k, \boldsymbol{y}_k|\boldsymbol{\beta}_k, \tau_{z_k}, \sigma^2)\pi(\tau_{z_k}|\sigma^2)d\tau_{z_k} \\
=&\mathcal{L}(\boldsymbol{X}_{ck}, \boldsymbol{y}_{ck}|\boldsymbol{\beta}_k, \sigma^2)\frac{1}{(\sqrt{2\pi\sigma^2})^{n_{tk}}}\frac{1}{\sqrt{2\pi\phi\sigma^2}} \\
&\int \exp\Big[-\frac{1}{2\sigma^2}||\boldsymbol{y}_{tk} - \tau_{z_k}\mathbf{1}_{n_{tk}} - \boldsymbol{X}_{tk}\boldsymbol{\beta}_k||^2 - \frac{\tau_{z_k}^2}{2\phi\sigma^2}\Big]d\tau_{z_k} \\
=&\mathcal{L}(\boldsymbol{X}_{ck}, \boldsymbol{y}_{ck}|\boldsymbol{\beta}_k, \sigma^2)\frac{1}{(\sqrt{2\pi\sigma^2})^{n_{tk}}}\frac{1}{\sqrt{2\pi\phi\sigma^2}}\exp\Big(-\frac{1}{2\sigma^2}||\boldsymbol{y}_{tk} - \boldsymbol{X}_{tk}\boldsymbol{\beta}_k||^2\Big) \\
&\int \exp\Bigg[-\frac{1}{2\phi\sigma^2}[(\phi\mathbf{1}_{n_{tk}}'\mathbf{1}_{n_{tk}}+1)\tau_{z_k}^2 - 2\phi(\boldsymbol{y}_{tk} - \boldsymbol{X}_{tk}\boldsymbol{\beta}_k)'\mathbf{1}_{n_{tk}}\tau_{z_k}]\Bigg]d\tau_{z_k} \\
=&\mathcal{L}(\boldsymbol{X}_{ck}, \boldsymbol{y}_{ck}|\boldsymbol{\beta}_k, \sigma^2)\frac{1}{(\sqrt{2\pi\sigma^2})^{n_{tk}}}\frac{1}{\sqrt{\phi\mathbf{1}_{n_{tk}}'\mathbf{1}_{n_{tk}}+1}} \\
&\exp\Bigg\{\frac{\phi[(\boldsymbol{y}_{tk} - \boldsymbol{X}_{tk}\boldsymbol{\beta}_k)'\mathbf{1}_{n_{tk}}]^2}{2\sigma^2(\phi\mathbf{1}_{n_{tk}}'\mathbf{1}_{n_{tk}}+1)} - \frac{1}{2\sigma^2}||\boldsymbol{y}_{tk} - \boldsymbol{X}_{tk}\boldsymbol{\beta}_k||^2\Bigg\}.
\end{aligned}
\tag{S.3}
$$

S.4 Testing treatment difference among baskets for CBHM

Assuming $K = 4$ available baskets at the stage of analysis, we construct the following dummy variables

$$\begin{aligned}(Z_1, Z_2, Z_3) =&(1, 0, 0) \text{ for basket 1,}\\ =&(0, 1, 0) \text{ for basket 2,}\\ =&(0, 0, 1) \text{ for basket 3,}\\ =&(0, 0, 0) \text{ for basket 4.}\end{aligned} \tag{S.4}$$

Define $T = 1$ for treatment arm, $T = 0$ for control arm.

Under models without covariates, fit the model on all available data:

$$\text{Full model: } Y =\theta_4 + \tau_4 T + \alpha_1 Z_1 + \alpha_2 Z_2 + \alpha_3 Z_3 + \gamma_1 Z_1 T + \gamma_2 Z_2 T + \gamma_3 Z_3 T + \epsilon. \tag{S.5}$$

Test it against the reduced model to obtain the F test statistics:

$$\text{Reduced model: } Y =\theta_4 + \tau_4 T + \alpha_1 Z_1 + \alpha_2 Z_2 + \alpha_3 Z_3 + \epsilon. \tag{S.6}$$

Under two covariates (X_1, X_2) adjusted modeling, fit the model on all available data:

$$\begin{aligned}\text{Full model: } Y =&\theta_4 + \tau_4 T + \beta_1 X_1 + \beta_2 X_2 +\\ &\alpha_1 Z_1 + \alpha_2 Z_2 + \alpha_3 Z_3 +\\ &\gamma_1 Z_1 T + \gamma_2 Z_2 T + \gamma_3 Z_3 T +\\ &\delta_1 Z_1 X_1 + \delta_2 Z_2 X_1 + \delta_3 Z_3 X_1 +\\ &\lambda_1 Z_1 X_2 + \lambda_2 Z_2 X_2 + \lambda_3 Z_3 X_2 + \epsilon.\end{aligned} \tag{S.7}$$

Test it against the reduced model to obtain the F test statistics:

$$\begin{aligned}\text{Reduced model: } Y =&\theta_4 + \tau_4 T + \beta_1 X_1 + \beta_2 X_2 +\\ &\alpha_1 Z_1 + \alpha_2 Z_2 + \alpha_3 Z_3 +\\ &\delta_1 Z_1 X_1 + \delta_2 Z_2 X_1 + \delta_3 Z_3 X_1 +\\ &\lambda_1 Z_1 X_2 + \lambda_2 Z_2 X_2 + \lambda_3 Z_3 X_2 + \epsilon.\end{aligned} \tag{S.8}$$