Table 6 Sensitivity analysis for the proposed BMA-bCRM approach using three different sets of working models under scenario 2

Scenario (p_{ej}, p_{ij})	Design	Selection probabilities (%) for RD at the end of trial: Dose level 1	2	3	4	5	None	Average percentage of efficacy	Average percentage of toxicity	Average number of patients
2		(25, 15)	(30, 25)	(35, 45)	(40, 55)	(45, 65)				
					Set 1					
	WM_1	(10, 5)	(15, 15)	(20, 30)	(25, 45)	(30, 55)				
	WM_2	(40, 25)	(45, 45)	(50, 60)	(55, 75)	(60, 80)				
	WM_3	(40, 5)	(45, 15)	(50, 30)	(55, 45)	(60, 55)				
	WM_4	(10, 10)	(20, 15)	(35, 20)	(40, 30)	(45, 50)				
	BMA-bCRM	12	**53.9**	30.1	0.2	0.1	3.7	13.3	12.2	43.9
	Number of patients	12	17.4	18	5.6	4.3				
					Set 2					
	WM_1	(10, 10)	(15, 15)	(20, 35)	(22, 50)	(24, 60)				
	WM_2	(25, 25)	(30, 30)	(35, 40)	(38, 45)	(40, 50)				
	WM_3	(20, 5)	(30, 15)	(40, 30)	(50, 45)	(55, 55)				
	WM_4	(25, 20)	(30, 40)	(35, 45)	(40, 50)	(45, 55)				
	BMA-bCRM	3.9	**60.7**	30.5	1.3	0.2	3.4	13.5	12.9	44
	Number of patients	9	20.3	16.6	7.3	4.3				
					Set 3					
	WM_1	(10, 10)	(20, 15)	(25, 20)	(30, 30)	(35, 45)				
	WM_2	(20, 20)	(30, 30)	(32, 40)	(33, 45)	(34, 50)				
	WM_3	(30, 20)	(35, 25)	(40, 30)	(45, 45)	(50, 50)				
	WM_4	(30, 25)	(32, 35)	(34, 40)	(36, 42)	(38, 44)				
	BMA-bCRM	1.6	**41.9**	52.2	0.2	0.3	3.8	14	14.2	43.8
	Number of patients	5.4	17.1	20.5	5.7	4.8				

20%, between "Set 2" and "Set 3." Especially, "Set 3" could not select the true RD with the highest probability. These results indicated that it was essential for the proposed BMA-bCRM approach not to lay out misplaced sets of working models, although this was obvious from its strategic nature.

We examined the operating characteristics of the proposed BMA-bCRM approach using the logistic model with a fixed intercept of 3. In this sensitivity analysis, four sets of working models are assumed to have the same efficacy and toxicity probabilities as in Table 2. The prior distributions for each parameter are identical to those of the power model because they are also sufficiently noninformative in this setting. Table 7 shows the results under scenarios 3 and 5. According to Table 7, the proposed BMA-bCRM approach using the logistic model is comparable to that using the power model with respect to the correct RD selection probability.

Table 7 Sensitivity analysis for the proposed BMA-bCRM approach using the logistic model with a fixed intercept of 3 under scenario 3 and 5

Scenario (p_{ej}, p_{tj})	Design	Selection probabilities (%) for RD at the end of trial: Dose level 1	2	3	4	5	None	Average percentage of efficacy	Average percentage of toxicity	Average number of patients
3		(15, 5)	(20, 10)	(25, 15)	(40, 20)	(45, 25)				
	WM_1	0	0	2	14.3	**77.5**	6.2	16	8.7	43
	Number of patients	3.7	3.2	5	8.5	28.2				
	WM_2	0	1.6	24	26.1	**40.8**	7.5	13.6	7.4	42.7
	Number of patients	4.2	6	12	10.7	23.2				
	WM_3	0	0	0	2	**91.8**	6.2	16.5	9.1	43.0
	Number of patients	3.7	3.3	3.2	4.4	30.9				
	WM_4	0	0	87.1	0	**0**	12.9	9.8	5.7	41.4
	Number of patients	3.7	3.2	35.9	0	0				
	BMA-bCRM	0	0	1.1	18.8	**73.9**	6.2	15.9	8.6	43
	Number of patients	3.7	3.3	4.5	10.7	26.4				
5		(20, 20)	(30, 30)	(40, 50)	(50, 55)	(60, 60)				
	WM_1	9	**39.2**	23.4	0.3	0	28.1	10.7	14.4	38.4
	Number of patients	7.7	13.9	16.5	8.5	5.9				
	WM_2	14	**59.7**	1.9	0	0	24.4	9	11	38.7
	Number of patients	14.2	22.7	8.7	4	5				
	WM_3	10.8	**41.9**	21.4	2	0.2	23.7	11.7	14.8	38.9
	Number of patients	8.7	13.7	10.9	7.5	8.5				
	WM_4	8.6	**39.1**	13.6	0	0	38.7	9.8	13.1	36.7
	Number of patients	7.8	15.2	16.7	0	0				
	BMA-bCRM	9	**47.9**	18	0.9	0.3	23.9	10.8	14.3	39.1
	Number of patients	7.9	16.4	14.6	8.5	5.8				

5. DISCUSSION

We have proposed the BMA-bCRM approach for jointly evaluating the efficacy and toxicity in Phase I clinical trials to determine the RD. In practice, it is also true that investigators have multiple ideas with regard to the true dose-response relations. The idea of BMA is more natural and acceptable in such a situation. Furthermore, the operating characteristics of the bCRM approach strongly depended on the assumed skeletons for efficacy and toxicity according to our simulation studies. Some skeletons often showed disastrous operating characteristics under specific scenarios. These facts also motivate our use of multiple combinations of skeletons and the idea of BMA. The proposed BMA-bCRM approach could mitigate the risk of underperforming results in terms of the correct RD selection probability caused by an inappropriate choice of skeletons than the bCRM approach. According to our simulation studies, the correct RD selection probability of the proposed BMA-bCRM was relatively close to that of the best-fitting working model while mitigating the poor RD selection caused by poorly fitting working models.

Obviously, the proposed BMA-bCRM approach could be easily implemented under other dose-response models, such as the complementary log–log model and hyperbolic tangent model. In the case of targeted agents, monotonically non-decreasing efficacy with no toxicity or flat efficacy with monotonically non-decreasing toxicity assumptions within the dose range, what we investigated might be reasonable from a biological perspective. For such a situation, the idea of BMA-bCRM might provide a reasonable solution to investigators by incorporating their assumptions as one of the working models.

ACKNOWLEDGMENT

The authors thank the associate editor and two referees for very helpful suggestions that substantially improved the article.

REFERENCES

Braun, T. M. (2002). The bivariate continual reassessment method: Extending the CRM to phase I trials of two competing outcomes. *Controlled Clinical Trials*. 23(3):240–256.

Chevret, S. (1993). The continual reassessment method in cancer Phase I clinical trials: A simulation study. *Statistics in Medicine*. 12(12):1093–1108.

Faries, D. (1993). Practical modifications of the continual reassessment method for Phase I cancer clinical trials. *Journal of Biopharmaceutical Statistics*. 4(2):147–164.

Hoeting, J. A., Madigan, D., Raftery, A. E., Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. *Statistical Science*. 14(4):382–401.

Ivanova, A. (2003). A new dose-finding design for bivariate outcomes. *Biometrics*. 59(4):1001–1007.

Mandrekar, S. J., Qin, R., Sargent, D. J. (2010). Model-based Phase I designs Incorporating toxicity and efficacy for single and dual agent drug combinations: methods and challenges. *Statistics in Medicine*. 29(10):1077–1083.

O'Quigley, J., Chevret, S. (1991). Methods for dose finding studies in cancer clinical trials: A review and results of a Monte Carlo study. *Statistics in Medicine*. 10(11):1647–1664.

O'Quigley, J., Hughes, M. D., Fenton, T. (2001). Dose-finding designs for HIV studies. *Biometrics*. 57(4):1018–1129.

O'Quigley, J., Pepe, M., Fisher, L. (1990). Continual reassessment method: A practical design for Phase I clinical trials in cancer. *Biometrics*. 46(1):33–48.

Raftery, A. E., Madigan, D., Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*. 92(437):179–191.

Thall, P. F., Cook, J. D. (2004). Dose–finding based on efficacy-toxicity trade-offs. *Biometrics*. 60(3):684–693.

Thall, P. F., Russell, K. E. (1998). A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in Phase I/II clinical trials. *Biometrics*. 54(1):251–264.

Whitehead, J., Zhou, Y., Stevens, J., Blakey, G., Price, J., Leadbetter, J. (2006). Bayesian decision procedures for dose-escalation based on evidence of undesirable events and therapeutic benefit. *Statistics in Medicine*. 25(1):37–53.

Yin, G., Yuan, Y. (2009). Bayesian model averaging continual reassessment method in Phase I clinical trials. *Journal of the American Statistical Association*. 104(487):954–968.

Zhang, W., Sargent, D. J., Mandrekar, S. J. (2006). An adaptive dose-finding design incorporating both toxicity and efficacy. *Statistics in Medicine*. 25(14):2365–2383.