1 Supplementary Material

1.1 Proof of Proposition 1

Fix a particular subgroup x. If the subgroup x is in the estimated positive set, then total expected error for subgroup x is given by $\lambda P_x(s)$. Similarly, if the subgroup x is in the estimated negative set, then total expected error for subgroup x is given by $(1-\lambda)(1-P_x(s))$. When $P_x(s) \geq 1-\lambda$, it holds that $\lambda P_x(s) \geq (1-\lambda)(1-P_x(s))$. In that case, $x \in H^+_{est}$ minimizes the total expected posterior error.

1.2 Additional experiments

Type 1 and Type 2 errors Figure 1.2 illustrates the trade-off between type-I and type-II errors for the RCT-KG algorithm (100 patients in each of 10 cohorts.) Of course, as λ increases, the type-I error decreases and the type-II error increases.

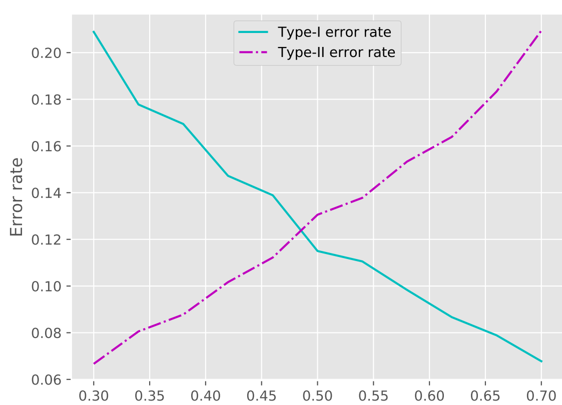

Figure 1: Tradeoffs between Type-I and Type-II errors

Informative Prior Finally, we present an experiment to illustrate the effect of having informative priors. We generated informative priors by sampling 50 patients from each subgroup and treatment group (but we did not count these patients as part of the total patient budget, which was either 500 or 1000). As seen from Table 1, having an informative prior increases the amount by which RCT-KG improves over UA. Having an informative prior is useful because it allows RCT-KG to make more informed decisions in the earlier stages and in particular to focus more on subgroups for which the difference between the true effectiveness of the treatment and control is smaller.

Budget	500	1000
Non-informative	0.1151	0.2813
Informative for sg 0,3	0.2000	0.3703
Informative for sg 1,2	0.1323	0.2878

Table 1: Improvement score for different budgets