We thank the reviewers for their detailed and insightful comments!

R3: In fact we have performed experiments using ϵ-DP in place of ACCU'. This procedure is somewhat more expensive but indeed still acceptable in terms of runtime. Our preliminary results showed that ACCU' led to almost identical results, so we adopted this simpler and cheaper alternative for our experiments. We will add a note to this effect. We also note that ACCU' can also be motivated from an intuitive perspective: if a coin i contributes p_i HEADs in expectation, how many coins do we need to flip to attain the desired number? This intuitively corresponds to the cost of active search. A more systematic theoretical and empirical study of ACCU' for NPB is an interesting topic on its own.

We have also investigated Monte Carlo (see `compute_negative_poisson_binomial_expectation_monte_carlo.cpp` in the supplementary material under `code/min_cost/`). Empirically (and unsurprisingly), it was much more expensive than ϵ-DP to achieve the same error level.

The specific k-nn probability model is described in Eq. (7) of Garnett et al. (2012) (line 327). Basically, the prior probability of each point being positive is the estimated marginal probability (e.g., 0.01). After each observation, these probabilities are updated by counting the proportion of positives in each point's k nearest neighbors, smoothed by the prior. This is a simple but effective model for active search. We will add more information to the appendix.

The only modification needed to adapt ENS to the cost-sensitive setting is appropriately specifying the "budget," as described lines 258–265 in the main text. The two settings are not directly related, as one is a maximum coverage problem and the other is a covering problem. However, the two problems are related in that one can be thought of as the dual of the other. This connection is perhaps why ENS is such a strong baseline for the CEAS setting.

In line 129, n is the number of candidate points to choose from.

R5: We strongly disagree that our "main contribution is a heuristic algorithm." We have (1) introduced a new optimization problem extending those considered in the literature, opening the potential for a new line of work and (2) established the optimal policy for this problem. This policy is computationally intractable; however, we (3) provide a fast and empirically strong approximate policy, guided by the optimal policy. Finally, we (4) provide interesting lower bounds on the approximation ratio for this problem. Although (3) is a heurisitc algorithm, our theoretical work (1, 2, 4) shows that heuristic algorithms are the best we can hope for due to the inherent hardness of the underlying problem (4).

"The authors have not compared with the large body of theoretical work on active search." Again, we disagree. We have compared with both the most relevant work on active search (Garnett, et al. (2012) and Jiang, et al. (2017, 2018a)), as well as with work from the stochastic submodular optimization literature. It is difficult to respond further as you have not identified any *particular* missing work. If the reviewer could cite the work they are referring to, we would be delighted to include it in our discussion.

"The improvement reported in Figure 1(b) is marginal." This is perhaps due to the visual contrast with the gap between one-step and two-step policies. Achieving a 5–10% improvement is exciting for applications where each experiment is costly. Further, our algorithm is consistently over 50% better than the popular greedy heuristic on drug discovery datasets, a massive reduction in cost.

The prior marginal distribution of points being positive is Bernoulli with a constant parameter p, the estimated ratio of positives in the pool (e.g., $p = 0.01$).

R6: $n^{0.16}$ vs $\sqrt{\log n}$: We absolutely agree with the reviewer that the two problems are different, analogous to set cover versus maximum coverage. It is not surprising that the two problems have different complexity. What we believe is interesting is that this paper formally establishes that cost effective active search is much harder to approximate than known bounds on active search. Beforehand, it was not obvious this problem should have a polynomial lower bound on the approximation ratio, and we suspected initially that it would be poly-logarithmic (such as in set cover).

Dependence on T: In the proof of Theorem 1, we used a construction where $T = n^{2\epsilon}$. In practice, the target number of positives usually grows as the total number of points. Besides, theoretically it is not very meaningful to assume constant T, since then even a random policy would have constant expected cost of T/p, where p is the ratio of positives.

R7: "how well the NPB expectation approximates the true expected cost": this is a great question. In short, it is provably infeasible to even approximate the true expected cost and we cannot expect *any* computationally fast algorithm to approximate the true expected cost (see Theorem 1 at line 135). We show this by proving the problem has a strong lower bound on the approximation ratio even if super polynomial time is allowed. Empirically, we cannot compute the true expected cost for even dozens of points (recall the $\mathcal{O}(n^t)$ complexity at line 129). This lower bound can perhaps be circumvented if some degree of conditional independence holds or some other useful structure in the probability model. Determining which probability distributions result in the ability to approximate the true expected cost is an exciting line of future research.