Taussig Cancer Institute
Cleveland Clinic
Jacob G. Scott, MD, DPhil (Oxon)

Dear Editor,

We are pleased to submit the manuscript, "Reinforcement Learning informs optimal treatment strategies to limit antibiotic resistance", for consideration as an article in *Proceedings of the National Academy of Sciences*. Our work highlights the development of a novel platform that leverages reinforcement learning to discover effective antibiotic cycling policies. In this paper, we present this potentially paradigm-shifting method, an associated python software package, and an application to a simulated population of *E. coli* evolving on empirically measured fitness landscapes.

Background and significance:

Drug resistant pathogens are a wide-spread and deadly phenomenon that were responsible for nearly 5 million deaths worldwide in 2019. The increasing prevalence of pan-drug resistance has prompted the CDC to declare that we have entered a "post-antibiotic era". Despite this evident public health crisis, development of novel antibiotics has all but ceased due to the poor return on investment currently associated with this class of drugs. Novel approaches to designing therapies that explicitly take into account the adaptive nature of microbial cell populations while leveraging existing treatment options are desperately needed.

Significant theoretical work has attempted to design optimal drug treatments for models of evolving systems. However, most methods for optimization of these models require a complete understanding of the underlying system dynamics. Such detailed knowledge is currently unobtainable in the clinical setting. In this study, we describe an approach that can approximate these optimal policies given only a fraction of the available information, filling a key unmet need in evolutionary medicine.

Major results: We present three main results from this manuscript.

- First, we develop a novel reinforcement learning framework for developing evolution-inspired antibiotic cycling regimes.
- Second, we show that reinforcement learning agents stably learn superior multi-drug policies in a simulated system of evolution given limited information about the system dynamics.
- Third, we show that policies learned by reinforcement learning agents display a capacity for long-term planning.

While we apply this method first to a model of *E. coli*, we believe that these principles can be applied to any evolving system with a measurable resistance phenotype (viruses, cancer, bacteria, fungi). For all these reasons, we believe our work to be a major advance with broad interest in the scientific community, well suited to the readership of *PNAS*.

Sincerely,

Jacob G. Scott, MD, DPhil (Oxon)