A teachable moment for dual-use

To the Editor — Dual-use research of concern (DURC) can be defined as research, mainly in the life sciences, that has the potential to be misapplied for harmful purposes[1]. Key examples are the synthesis of mousepox[2], the synthesis of poliovirus[3], the generation of the 1918 influenza virus[4], gain-of-function studies with H5N1 in ferrets[5] and the synthesis of horsepox, the viral cousin of smallpox[6]. These seminal events involved the physical synthesis of a biological agent. However, the time may have come to also consider dual-use risk of the development of toxic agents in silico, in the light of the alarming results of a computational experiment we recently performed for a biennial arms control conference[7]; we used a generative artificial intelligence (AI) approach previously developed for drug discovery applications, and found it could easily design a range of nerve agents including VX. The experiment demonstrated the alarming speed and ease with which such software — based on open-source tools and datasets from the public domain — could be used for bad purposes. Our experiment was subsequently covered widely in the media, reaching a network of scientists, experts and lay people alike[8], and its implications were recognized at the highest levels of governments within a matter of days. The level of interest was probably amplified owing to the enfolding war in Ukraine, with Russia's invasion and the threat of biochemical weapons use.

Although we are a small team, the perspectives we provide are nationally diverse, span the private, academic and government sectors, and draw on expertise from the natural and social sciences as well as more technical fields such as computing and drug discovery. We believe our experiences from reporting and discussing our computational results so widely provided several important lessons that we wish to share with the scientific, ethics and security communities. First, the experiment is a powerful example of a concrete dual-use risk concern arising from converging technologies and this could be used to raise awareness of the security dimension of life sciences research. Second, our experience as a whole — obtained from reviewer and editorial feedback on our paper[7] to interactions with many groups after publication including with several interviewers — taught us the importance of increasing awareness in a responsible, non-alarmist way. Third, we need to consider what these dual-use findings mean for responsible science in drug discovery, and what action the community should be taking.

Responses to our article[7] have been varied widely. Some academics and government employees have requested the compound structures (this was denied); some suggested we should only use the technology for good (yes). Others asked whether the software could help to identify treatments for diseases of interest to them (yes, potentially). Some felt our thought experiment was obvious; whereas several experts on chemical weapons accepted that they had not considered it and saw novelty. Many were concerned about the security of the data generated. There have also been questions on why we published and whether the details of the experiment should have even been published at all — in line with responses to previous biological dual-use examples. In reply to this point, we believe that this new example highlights an important message, that dual-use risk potential in the life sciences goes beyond the synthesis of biological agents. For governments, our thought experiment highlights the challenge of how and when to limit access to generative and machine-learning software, including through export controls. For the drug-design community, it will now be necessary to agree on ways to share data and models securely.

Our thought experiment has already become a 'teachable moment for dual-use' — a positive unintended consequence of our study. It can be drawn on as a test case for considering the risks of research that involves converging technologies, in contrast to previous dual-use examples that focus on physical biological agents. It can also be used to provide dual-use risk training for those applying AI in drug discovery in the context of nerve agents and chemical weapons. Dual-use concerns in AI is already an urgent topic on the agenda for policymakers, but our results point to the need for further action in the development of regulation. Our pre-emptive publication may lead to increased diligence around AI technologies, datasets, models and related software for designing new molecules and the subsequent consideration of ethics and societal consequences[9]. Dual-use potential of AI is of concern to all scientists, not just those in the field of drug discovery. We hope that our thought experiment puts dual-use risk on the radar for a wider area without raising undue alarm and that it stimulates the search for potential solutions[8]. ❐

Fabio Urbina[1], Filippa Lentzos[2], Cédric Invernizzi[3] and Sean Ekins[1]✉

[1]*Collaborations Pharmaceuticals, Raleigh, NC, USA.* [2]*Department of War Studies and Department of Global Health & Social Medicine, King's College London, London, UK.* [3]*Spiez Laboratory, Federal Department of Defence, Civil Protection and Sports, Austrasse, Spiez, Switzerland.*

✉*e-mail:* sean@collaborationspharma.com

Published online: 12 July 2022
https://doi.org/10.1038/s42256-022-00511-6

References

1. NIH. https://osp.od.nih.gov/biotechnology/dual-use-research-of-concern/ (accessed 15 June 2022).
2. Jackson, R. J. et al. *J. Virol.* **75**, 1205–1210 (2001).
3. Cello, J., Paul, A. V. & Wimmer, E. *Science* **297**, 1016–1018 (2002).
4. Tumpey, T. M. et al. *Science* **310**, 77–80 (2005).
5. Herfst, S. et al. *Science* **336**, 1534–1541 (2012).
6. Noyce, R. S., Lederman, S. & Evans, D. H. *PLoS ONE* **13**, e0188453 (2018).
7. Urbina, F., Lentzos, F., Invernizzi, C. & Ekins, S. *Nat. Mach. Intell.* **4**, 189–191 (2022).
8. Editorial. *Nat. Mach. Intell.* **4**, 313 (2022).
9. Shankar, S. & Zare, R. N. *Nat. Mach. Intell.* **4**, 314–315 (2022).

Acknowledgements

C.I. contributed to this article in his personal capacity. The views expressed in this article are those of the authors only and do not necessarily represent the position or opinion of the Spiez Laboratory or the Swiss Government. We acknowledge funding from the National Institutes of Health (NIH) grant number R44GM122196-02A1, the National Institute of General Medical Sciences (NIGMS) grant numbers 1R43ES031038-01 and 1R43ES033855-01, and the National Institute of Environmental Health Sciences (NIEHS) for our machine-learning software development and applications. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Competing interests

F.U. and S.E. work for Collaborations Pharmaceuticals. F.L. and C.I. have no conflicts of interest.