Position Statements on Responsible Uses of Technology and Health Data During Times of Crisis: Making Real-World Data Useful within the New Drug Application Process

The coronavirus pandemic has placed renewed focus on expanded access (EA) programs that provide compassionate use exceptions to waves of patients seeking innovative pharmacological solutions.[1] While commendable, justifiable, and empathetic, The US, as well as most if not all EA programs[2-4] are not designed to collect clinical data for the FDA New Drug Application (NDA) process, or similar processes worldwide.[5]

Currently, EA programs lack the necessary rigor of meticulously crafted and controlled randomized controlled trials (RCT), [6] the gold standard in FDA clinical trials,[7] which, among other requirements, ensure (i) that each patient is closely monitored for adverse effects, (ii) that the data is properly documented, (iii) that the results are reliable and not biased by externalities, and (iv) that the results are casual not simply correlated. Further, EAs impede the approval process through diverting limited supplies of the drug, putative patients, and healthcare professionals. A resulting delay of regulatory approval harms the general public.

We propose that the negative effects of EAs be mitigated through a practical merging of EA programs into clinical trials. EA data can be valuable in an FDA framework that incorporates Real-World Data (RWD) as suitable substitutes for RCT data (which is often biased against underrepresented minorities, among other concerns[8]). To its credit, the UK, under the Early Access to Medicines Scheme, was the first to allow RWD from a compassionate use program to be officially considered as part of regulatory submission.[9]

In the US, the consideration of RWD within the clinical trial process is already mandated by the 21st Century Cures Act, but progress has been slow.[10] However, desperate times have allowed for the development and implementation of technological and regulatory wallflowers.[11,12] Perhaps the time is ripe for RWD incorporation as well. The FDA has an opportunity now to set standards for RWD such that it is statistical meaningful and actionable with minimal bias, transparent and verifiable.[13] This framework could allow for cohorts of EA patients to be included in regulatory trials, impelling rather than impeding drug development.

Expanded data collection is only part of the solution. Digital twins, i.e., *in silico* representations of real-world objects can be an alternative to morally problematic placebos in EA datasets. Originally conceived for NASA space vehicle lifecycle management,[14] they have been used primarily in engineering fields wherein devices can be stress-tested virtually.[15 16] Emerging research and patent

applications suggest promising developments in creating digital twins of living organisms that could be used in EA clinical trials.[17-23]

Patient digital twins would use novel data sources[24] like demographic data, family history electronic health records, laboratory results, physiologic measurements and insurance data, imaging and signal data and other unstructured health data, as well as substantial 'omic information arising from a patient's genome, microbiome, transcriptome, proteome and metabalome.[25]

Similarly, RWD can be extracted from underused data sources, including: electronic health records (EHR), billing, claims, and insurance data, self-identifying information provided by patient registries, groups, social media pages, and information collected by professional and recreational internet of things (IoT) devices ranging from insulin pumps to Apple Watches.[26]

While promising, this proposal has technical limitations as RWD and the health data necessary for digital twins typically lack structure, standardization and clarity, and often includes confounding factors. Data for both RWD and Digital Twins also create ethical, legal and social limitations, particularly privacy concerns relating to the various data sources,[27] as well as confounding biases such as: (i) information biases stemming from errors in data capture, lack of standardization, or incomplete data retrieval; (ii) attrition biases given the lack of structured surveillance; (iii) compliance or performance bias resulting from EA patients failing to adhere to regimented treatment; (iv) confounding biases as a result of the heterogeneity of the patients from varied demographics, clinical environments, and comorbidities; (v) immortal time biases; and, (vi) selection bias. Furthermore, although it might look like that the privacy and identifiability issues relating to these data can be resolved with suggested anonymization techniques within the current legal framework (e.g. HIPPA), recent evidence suggests that there are hidden identifying information and privacy leakages when multiple sources of health data are combined.[28,29] Hence, any proposed framework needs to incorporate solutions to anonymity and privacy beyond the traditional approaches that are currently in place.

With RWD efforts already underway for various aspects of COVID-19 treatment, this area of research is expanding rapidly.[30] There are already some efforts to incorporate RWD into the clinical trial process,[31] and with the FDA promising to update its current guidelines[32] and tools[33] for the use of RWD by the end of 2021, this is an opportune moment to work toward a better understanding of the ethical, social and legal limitations relating to areas such as patient consent, anonymity, patient data de-identification, security, privacy, liability, transparency and standardization of RWD such that future RWD collection efforts can move forward safely and effectively.

[1] *Abigail Alliance for Better Access v. Eschenbach*, 469 F.3d 129 (D.C. Cir. 2006).
[2] 21 CFR 312 section & 21 CFR part 812
[3] http://righttotry.org/in-your-state/
[4] Trickett Wendler, Frank Mongiello, Jordan McLinn, and Matthew Bellina Right to Try Act of 2017 (S. 204, Pub.L. 115–176)
[5] Mussa Rahbari a & Nuh N Rahbari, Compassionate use of medicinal products in Europe: current status and perspectives *Bulletin of the World Health Organization* 2011;89:163-163.
[6] Chapman, C.R., Moch, K.I., McFadyen, A., Kearns, L., Watson, T., Furlong, P. and Bateman-House, A., 2019. What compassionate use means for gene therapies. *Nature biotechnology*, *37*(4), pp.352-355.
[7] https://www.fda.gov/media/120060/download
[8] Britton, A., McKee, M., Black, N., McPherson, K., Sanderson, C. and Bain, C., 1999. Threats to applicability of randomised trials: exclusions and selective participation. *Journal of health services research & policy*, *4*(2), pp.112-121.
[9] Stein, D. and Soni, M., 2018. Early Access Programs Opportunities and Challenges for Real-World Data Collection. In *The Evidence Forum* (pp. 4-9).
[10] Klonoff DC. The new FDA real-world evidence program to support development of drugs and biologics. Journal of diabetes science and technology. 2020 Mar;14(2):345-9.
[11] Dov Greenbaum, The Coronavirus Crisis Can Push Technology Forward, Calcalist CTECH March 20, 2020
[12] Dov Greenbaum, After decades of being unappreciated, the pandemic could give distance learning a new lease on life, Calcalist CTECH June 9, 2020
[13] Schneeweiss S, Glynn RJ. Real-world data analytics fit for regulatory decision-making. American journal of law & medicine. 2018 May;44(2-3):197-217.
[14] Glaessgen E, Stargel D. The digital twin paradigm for future NASA and US Air Force vehicles. In53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA 2012 Apr 16 (p. 1818).
[15] Grieves MW. Product lifecycle management: the new paradigm for enterprises. International Journal of Product Development. 2005 Jan 1;2(1-2):71-84.
[16] Grieves M, Vickers J. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In Transdisciplinary perspectives on complex systems 2017 (pp. 85-113). Springer, Cham.
[17] Georges-Filteau J, Cirillo E. Generative Adversarial Networks Applied to Observational Health Data. CoRR. 2020 May 11.
[18] Guan J, Li R, Yu S, Zhang X. A Method for Generating Synthetic Electronic Medical Record Text. IEEE/ACM Transactions on Computational Biology and Bioinformatics. 2019 Oct 23.
[19] Yahi A, Vanguri R, Elhadad N, Tatonetti NP. Generative adversarial networks for electronic health records: A framework for exploring and evaluating methods for predicting drug-induced laboratory test trajectories. arXiv preprint arXiv:1712.00164. 2017 Dec 1.
[20] Fisher CK, Smith AM, Walsh JR. Machine learning for comprehensive forecasting of Alzheimer's Disease progression. Scientific reports. 2019 Sep 20;9(1):1-4.
[21] Walsh JR, Smith AM, Pouliot Y, Li-Bland D, Loukianov A, Fisher CK. Generating Digital Twins with Multiple Sclerosis Using Probabilistic Neural Networks. arXiv preprint arXiv:2002.02779. 2020 Feb 4.
[22]Farrell S, Mitnitski A, Rockwood K, Rutenberg A. Generating individual aging trajectories with a network model using cross-sectional data. bioRxiv. 2020 Jan 1.
[23]EP3646331A1, Jonathan Zimmerman Methods and systems for generating a patient digital twin General Electric Company
[24] Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, Liu PJ, Liu X, Marcus J, Sun M, Sundberg P. Scalable and accurate deep learning with electronic health records. NPJ Digital Medicine. 2018 May 8;1(1):18.
[25] Greenbaum D, Luscombe NM, Jansen R, Qian J, Gerstein M. Interrelating different types of genomic data, from proteome to secretome:'oming in on function. Genome research. 2001 Sep 1;11(9):1463-8.
[26] Sherman, Maya, Ziv Idan, and Dov Greenbaum. "Who Watches the Step-Watchers: The Ups and Downs of Turning Anecdotal Citizen Science into Actionable Clinical Data." *The American Journal of Bioethics* 19.8 (2019): 44-46; Dov Greenbaum "Avoiding Overregulation in the Medical Internet of Things" in Cohen, I. G., Lynch, H. F., Vayena, E., & Gasser, U. (Eds.). (2018). Big Data, Health Law, and Bioethics. Cambridge University Press.
[27] https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-records-and-electronic-signatures-clinical-investigations-under-21-cfr-part-11
[28] Harmanci A and Gerstein M. "Quantification of private information leakage from phenotype-genotype data: Linking attacks" (2016) Nature Methods 13:251-6

[29] Gursoy G and Gerstein M. "As coronavirus testing expands, new personal privacy issues arise." (2020) Hartford Courant

[30] COVID-19 Evidence Accelerator: https://evidenceaccelerator.org/
[31] Gliklich, R.E. and Leavy, M.B., 2020. Assessing Real-World Data Quality: The Application of Patient Registry Quality Criteria to Real-World Data and Real-World Evidence. *Therapeutic Innovation & Regulatory Science*, pp.1-5.
[32] FDA, Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drugs and Biologics Guidance for Industry Draft Guidance, May 2019 https://www.fda.gov/media/124795/download
[33] Baumfeld Andre, E., Reynolds, R., Caubel, P., Azoulay, L. and Dreyer, N.A., 2019. Trial designs using real-world data: The changing landscape of the regulatory approval process. *Pharmacoepidemiology and Drug Safety*.