

Commentary

Artificial Intelligence in Pharmacovigilance: Scoping Points to Consider



Manfred Hauben, MD, MPH^{1,2}; and Craig G. Hartford, MBBCh, PhD³

¹Safety Sciences Research, Pfizer Inc, New York, NY, USA; ²Department of Medicine, NYU Langone Health, New York, NY, USA; and ³Pharmaceutical Physician, Kent, United Kingdom

ABSTRACT

Artificial intelligence (AI), a highly interdisciplinary science, is an increasing presence in pharmacovigilance (PV). A better understanding of the scope of artificial intelligence in pharmacovigilance (AIPV) may be advantageous to more sharply defining, for example, which terms, methods, tasks, and data sets are suitably subsumed under the application of AIPV. Accordingly, this article explores relevant points to consider regarding defining the scope of AIPV and offers a potential working definition of the scope of AIPV. (*Clin Ther.* 2021;43:372–379) © 2020 Elsevier Inc.

Key words: AI, artificial intelligence, machine learning, pharmacovigilance, technology.

INTRODUCTION

Artificial intelligence (AI) is being increasingly used in pharmacovigilance (PV). On the basis of a MEDLINE search for the terms *artificial intelligence* and *pharmacovigilance*, the field of artificial intelligence in pharmacovigilance (AIPV) is rapidly growing (Figure 1). Although the recent increase shown in this figure is only a crude signal of the magnitude of increasing interest, it aligns with observations of scientific meeting agendas and initiatives in this area. For example, the newly formed Drug Safety Research Unit International Working Group on Signal Detection and Evaluation has a subgroup devoted to AI.¹

A dictionary definition of AI is different from a working definition of AI. An insufficient understanding of the scope of AIPV, for example,

which terms, methods, tasks, and data sets are ordinarily considered to be included in the application of AIPV, will likely hamper efficient and consistent execution of various activities for which decisions about what is and is not AIPV must be made. Such activities potentially adversely affected would include, for example, conducting systematic reviews or planning scientific meetings or working groups. AIPV scoping is even more essential given the nested and overlapping fields of AI, machine learning (ML), deep learning (DL), data mining, and cognitive computing. In addition, AIPV stakeholders are a diverse range of scientific and policymaking disciplines with variable baseline knowledge.

Another key use of a scoping exercise is to create a well-defined mapping of what is and is not known about a topic.² Systematic literature reviews can often return many false-positive results, even with sensitive and specific search strategies,³ which then necessitates arduous and potentially inconsistent relevance adjudication via title, abstract, and possibly full-text review. A case in point is a recent systematic review of one circumscribed subset of AIPV, natural language processing for electronic health records–based PV, which returned 1422 citations of which all but 48 were excluded.⁴ An upfront sharper topic scope may ease the burden following such systematic literature reviews. A well-defined scope of AIPV can also provide orientation for new entrants into the field. In this article, we explore relevant points to consider (PTCs) regarding defining the

Accepted for publication December 19, 2020

<https://doi.org/10.1016/j.clinthera.2020.12.014>

0149-2918/\$ - see front matter

© 2020 Elsevier Inc.

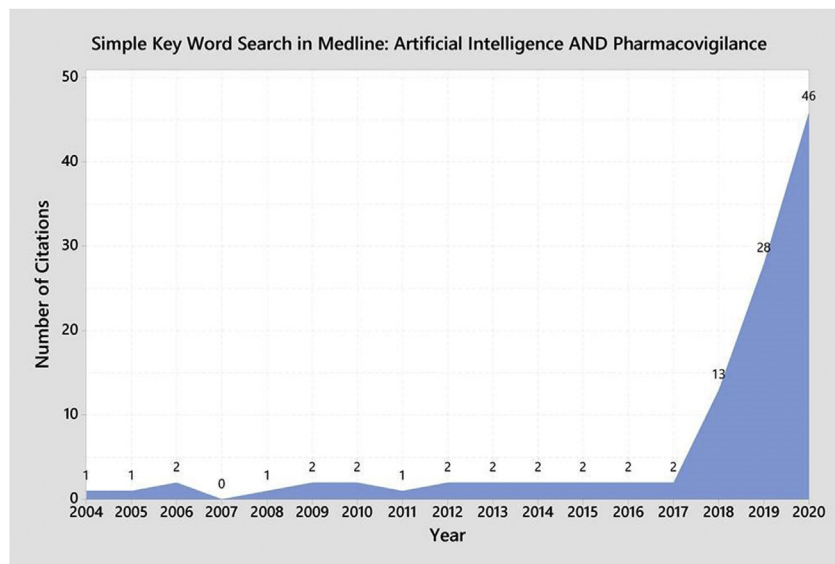


Figure 1. Number of unique citations returned by a simple keyword search for *artificial intelligence* and *pharmacovigilance* in MEDLINE on January 10, 2021, without application of automatic term explosion or mapping or an adjacency operator, from 2004 through 2020.

scope of AIPV and offer a potential working definition of the scope of AIPV to, in conjunction with the PTCs, facilitate further deliberation on this topic.

PTCs: AI TERMS

Artificial is not particularly difficult to define, but *intelligence* is harder to define. Legg and Hutter⁵ reviewed 70 definitions, including 18 from AI researchers. The psychologist Gardner⁶ proposed that individuals possess ≥ 8 relatively autonomous intelligences, including visual-spatial, verbal-linguistic, logical-mathematical, bodily-kinesthetic, and interpersonal intelligence. These intelligences may correspond to narrow AI tasks, such as machine vision, automated driving, image captioning, self-driving cars, reading radiographs or interpreting biopsy results, and intelligent agents. These are tasks that people perform. Mapping intelligence definitions to the AI tasks that people perform corresponds to much of AI, but it is not a perfect correspondence because many AI applications involve extraction of information from data sets that is not and cannot be performed by humans without machines, even given unlimited time. AIPV is currently narrow AI, applied to tasks such as natural language processing of

information from medical text (eg, adverse event extraction).

Definitions of the term *AI* abound. Wang⁷ presents one thorough analysis of defining AI and notes that a *dictionary* definition of AI is different from a *working* definition of AI. Three elements of Wang's analysis are particularly pertinent to embarking on developing a working definition of AIPV. First, defining AI is important to prevent conflicting implicit assumptions and misunderstandings that could degrade research, discussion, and debates. Second, there are 5 typical categories of AI definitions: structure, behavior, capability, function, and principle. Definition by capability is based on specific application domains. Third, a working definition guides a research agenda or objective and facilitates its efficient and consistent execution. It is based on 4 guiding principles: similarity to common use, sharpness, leading to fruitful research, and simple as possible.

PTCs: METHODS

AI is a broad area of computer science, inclusive of ML, which in turn includes DL and other forms of ML, sometimes known as traditional ML, as well as

non-ML forms of AI, sometimes referred to as “good old-fashioned AI (GOFAI)”. ML can be discussed from the perspective of types of learning (eg, supervised, unsupervised, and reinforcement learning), general ML tasks, and specific methods.

A historic, widely cited, and authoritative intentional definition of ML is “a computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance T, as measured by P, improves with experience E.”⁸ This definition leaves residual fuzziness in the boundary between ML, especially some traditional ML, and basic statistical methods. After all, ordinary least-squares regression, which is frequently listed in the gray literature as a form of traditional ML, might not come to mind when we think of cutting-edge AI.

However, improved accuracy with increased experience, the essence of McCarthy's original definition of ML, is compatible with the increased power and precision of ordinary least-squares regression with increasing sample size. This inevitably leads to the still much debated question of what is the difference, if any, between ML and other statistical methods or modeling?

Table I summarizes several typical characteristics of ML and non-ML data analyses. Table I is not intended to be fully comprehensive, and although the characteristics listed in Table I facilitate some understanding of methods, current applications, and practices, they do not establish a bright and immutable line between ML and non-ML. Not every element in Table I is decisive in all cases; therefore, scoping decisions require a holistic consideration of

Table I. Typical characteristics of ML and non-ML approaches.

Machine Learning	Non-ML (eg, Statistical Analyses, Computations, and Models)
<i>Relative</i> emphasis on classification, potential for utility in prediction, with data split for training and testing; model interpretability, albeit potentially desirable, is less emphasized	<i>Relative</i> emphasis on inference, estimation, <i>P</i> values, CIs, goodness of fit, and relationships between variables and their individual contributions and model interpretability; typically analyzes the entire data set at once; observations from changes in model fits can provide insights about behavior of variables as individual predictors
ML architecture and learning defined by external hyperparameters initialized before training and optimization, not estimated from the data (eg, number of hidden units, layers and epochs, learning rate, dropout rate in neural networks)	Model hyperparameters of prior parametric distribution and direct model parameters often estimated from the data when model is run, although can be assigned based on prior beliefs in Bayesian models
Accommodates big data with hundreds of millions of parameters, high or poorly defined dimensionality, many nonlinearities and ambiguous data structure	Usually accommodate small to moderate sized, relatively well-behaved data sets with limited number of parameters
Representational learning and automated feature extraction	Extensive manual feature engineering
More ways to improve performance (eg, hyperparameter tuning, increased data, run time, or ensemble learning)	Performance improvement with increased sample size
More nonparametric, fewer assumptions about data and model theory	More parametric and requires assumptions to be made about structure of data and model theory

Abbreviation: ML, machine learning.

all the elements listed in Table I as well as the other PTCs.

Representational learning (Table I) is a key concept for ML, automatically extracting predictive features from data. This procedure reduces the need for the extensive manual feature engineering more often, but not exclusively, used in non-ML methods. Some techniques, such as regression, can be enhanced for representational learning (eg, by implementation with regularization).

Many basic methods are clearly ML (and thus AI), fitting multiple machine learning characteristics listed in Table I. These methods include neural networks, individual decision trees, random forests, support vector machines, and naive Bayes theorems. However, some methods are both ML and non-ML. For example, logistic regression is often used for classification and prediction but also routinely provides parameter estimates, *P* values, CIs, goodness of fit, and model interpretability.

Using the criteria listed in Table I, the commonly performed disproportionality analysis of spontaneous reporting system data, namely data mining,⁹ is not in scope for AIPV. For example, calculation of a proportional reporting ratio is executed with code that fully defines a single, static arithmetical calculation, although many calculations are performed simultaneously and in a low-dimensional space.

Bayesian logic is frequently applied in ML but does Bayesian logic therefore elevate Bayesian disproportionality analysis to a form of AIPV? Contemporary forms of Bayesian disproportionality analysis may be viewed as learning from the data, updating hyperparameters with increasing experience, to converge toward a final model, thus reminiscent of intelligence. However, it is still a prespecified (parametrically via conjugate distributions or using simple arithmetical adjustments roughly akin to plus four CIs) density estimation of observed-to-expected reporting frequencies. It focuses on dampening the sample variance in low-dimensional projection of large sparse data (ie, 2×2 contingency tables). It is therefore arguably not in scope when also considering other parts of Table I. Future enhanced versions, accommodating more variable dimensions or ensembling with other methods or information, could well more readily fall into scope, so an open mind is desirable.

Some ML existed long before other types of ML and big data did (eg, principle components analysis). Principle components analysis is a dimensionality reduction technique developed by Karl Pearson in 1901. Its inclusion as ML is justified because (1) it is adaptive, performing automated feature extraction (ie, a new basis defined by weighted combinations of the input variables defining a lower dimensional space); (2) more contemporary formulations, such as robust and randomized principle components analysis, are more specialized for big and/or sparse data sets that are well described by nonlinear models; (3) it is frequently used for ML tasks and optimizations; and (4) minimum assumptions are required for exploratory purposes.¹⁰ Note that dimensionality reduction includes more advanced methods specialized for highly nonlinear big data, such as various forms of embedding.

How about in real-world practice? Python (Python Software Foundation, Wilmington, Delaware) is the most commonly used programming language for ML. On the GitHub repository, it is the number 1 language for ML. In a Kaggle survey, 87% of data scientists reported using Python, more than any other language. The Python ML libraries, frameworks, and associated documentation may therefore provide a reality check on our method PTCs.

One basic Python ML library is Scikit-learn. Scikit-learn, version 0.22.2 documentation includes a graphical guidance flowchart that provides an aerial view of basic ML options for numerical data in Python (Figure 2).^{11,12} Note the 4 basic ML tasks are classification, regression, clustering, and dimensionality reduction, but classic statistical methods such as regression, are included when implemented with regularization or optimization, consistent with our PTCs.

Non-ML AI (including rules engines, expert systems, knowledge graphs, or symbolic AI) explicitly code domain knowledge and symbolic representations of human reasoning. These analyses originally used static rules without self-learning or self-correction and therefore can be considered as out-of-scope for AIPV. Examples of such non-ML approaches that are therefore out of scope for AIPV might be rules-based duplicate report detection and automation of established causality assessment algorithms, without corresponding automated case report information extraction and entry.

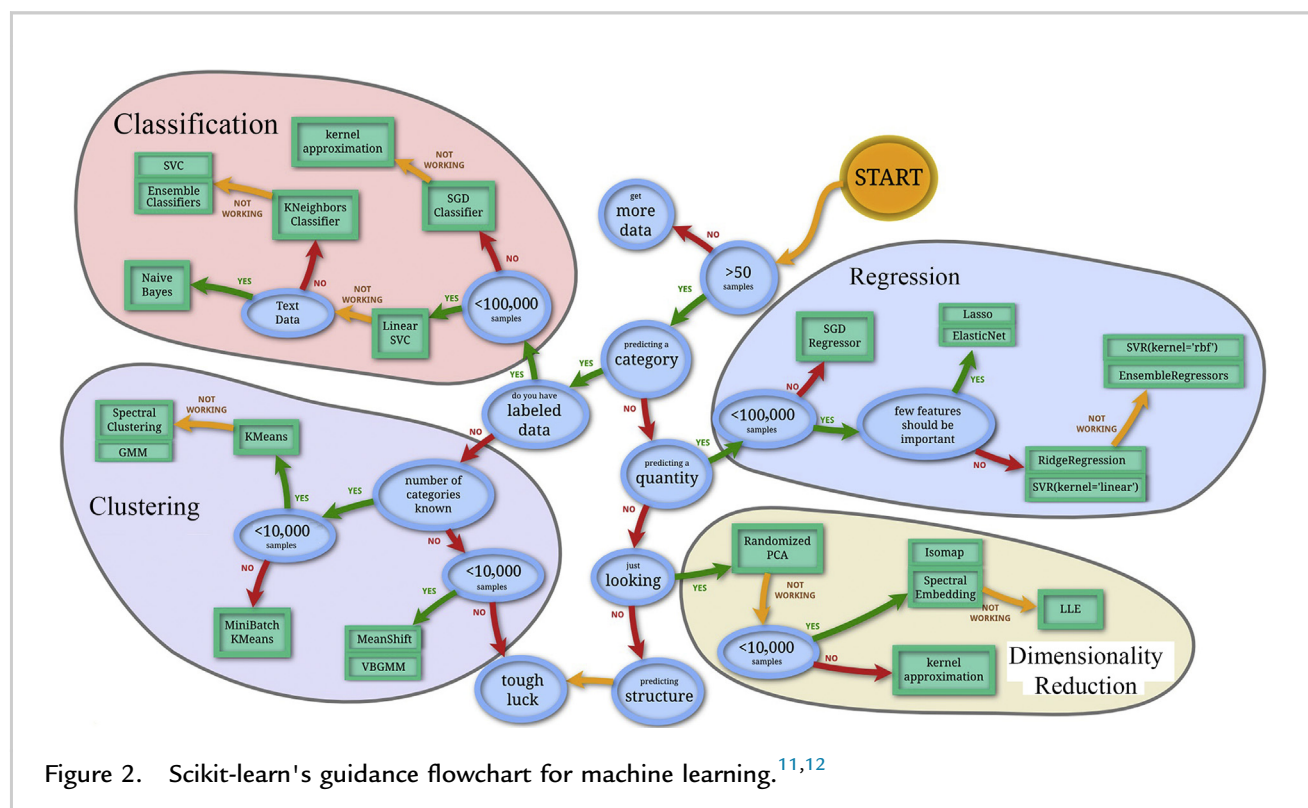


Figure 2. Scikit-learn's guidance flowchart for machine learning.^{11,12}

However, AIPV, like AI in general, continues to evolve, unfortunately sometimes along hype cycles.¹³ Non-ML symbolic AI is seeing something of a 21st-century resurgence and should not be dismissed out of hand. Two examples are illustrative. First, description logic is a web-based knowledge representation language that allows the construction of machine-readable and analyzable data. This language facilitates enhanced adverse event case definitions and queries, integration of data sets, and inferencing. These factors are obviously important in multiple areas of PV.¹⁴ Second, hybrid approaches, such as neurosymbolic AI, combine DL with symbolic AI to complement each other. These approaches are already being applied in areas of relevance to PV.¹⁵

In addition, some other approaches customized to enable the analysis methods to be deployed in PV are of obvious interest and relevance and therefore in scope. These approaches include ancillary computational, data management, and workflow strategies and technologies, such as platforms,

frameworks, and pipelines, within the knowledge discovery in data bases framework.^{16,17}

Generally, computer algorithms and ancillary technologies that should be regarded as in scope for AIPV are those that make it more feasible to define, represent, combine, learn, reason on, and/or extract information from numerical and nonnumerical data sets that otherwise pose difficulties for human processing and standard statistical models. Difficulties posed by data sets may be attributable to size, dimensionality, heterogeneity, nonlinearity, and/or otherwise ambiguous, incomplete, and/or erroneous structure.

PTCs: TASKS

A widely cited definition of PV is “the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug-related problem.”¹⁸ We often envision PV as dedicated professionals toiling over increasingly large individual case safety report (ICSR) data sets, periodic summary reports, signal detection

and evaluation, risk management, and product labeling. These tasks are all domains of good PV practice for organizations with statutory or legal requirements to monitor drug safety. The high-throughput nature of some of these tasks are enticing targets for AI. Mokute et al¹⁹ derived a set of 51 decision points corresponding to 25 essential cognitive services in the domain of ICSR processing. A broader array of actual and potential applications of AIPV exist,²⁰ including aggregate analysis, periodic reporting, signal detection, benefit/risk assessment, and risk management.^{21,22} Signal detection and evaluation, being a dominant PV activity, makes AI-supported evidence synthesis an appealing goal.²³

PV, as defined above, is also practiced in daily life by health care practitioners, scientists performing research, and patients reporting adverse event data, the last potentially through wearable health devices. AI has realized potential applications here as well, such as using neural networks to analyze multifocal electroretinograms for early detection of hydroxychloroquine toxicity.²⁴ The diagnosis of drug-induced illness from skin and fundus photographs, biopsies, and imaging studies are additional examples. An increasingly common use of AIPV is automating the processing tasks applicable to ICSRs, in which natural language processing and ML are already being used to extract ICSR information.

PTCs: DATA SETS AND INFORMATION SOURCES

Contemporary PV is a holistic process, especially signal evaluation with converging lines of evidence from multiple data streams, for which both clinical and preclinical data are sought. Numerous data from the molecular level to the human level are potentially in scope, with well-established sources of data, such as preclinical toxicology data, spontaneous reports, electronic claims and medical records, registries, and clinical trials. Other information sources include user-generated content (eg, social media, internet search logs, wearable health devices, and adverse event reporting apps), chemoinformatics, and systems biology and Omics databases. ML can discover features and relationships within these large biological networks.^{25–28} Many of these modern data sets have one or more characteristics and challenges of big data, such as the increasing number of V's (volume, velocity, variety, value, validity, variability,

veracity, viability, viscosity, and volatility).²⁹ These big data characteristics, also alluded to in the PTCs Methods section, present challenges for capture, integration, storage, and knowledge extraction.

Collaborative initiatives frequently overlap disciplines of systems and network biology, chemical biology, computational toxicology, and bioinformatics. These initiatives use numerous publicly available big data repositories that contain millions of compounds and corresponding data points on full molecular structures, chemical descriptors, toxicity, and multiple Omics data. These repositories are potential treasure troves from which AI can identify complex pathways that lead to patient adverse outcomes, such as potentially identifying toxicity during drug development. AI might also support postapproval signal evaluation by identifying signature patterns of network connections and activations, bolstering the Bradford Hill criteria of biological plausibility in signal evaluation, which was implemented to some degree as signal substantiation in the European Union Adverse Drug Reaction initiative.²⁷ Complex analyses might even tease apart drug-specific signals within a class.²⁸ Note that some of these data sets include drugs as well as agricultural, environmental, and occupational toxins, requiring careful review to avoid missing relevant data sets. Although smaller, traditional, preclinical animal toxicology data sets are not conducive to AI, it is reasonable not to limit searches to human studies to prevent missing articles about AIPV for transspecies toxicity prediction.³⁰

AIPV SCOPE DEFINITION

The above PTCs provide an orientation to the AI and PV field and support defining the scope of AIPV. Inspired by these definitional PTCs and considering the criteria of Wang,⁷ we consider that a definition for the scope of AIPV would likely be positioned around the following elements: development, study, and use of computerized algorithms and ancillary technologies; supporting and/or performing operational and scientific PV tasks; improving PV knowledge discovery, process accuracy, and/or efficiency; representing, managing, integrating, reasoning on, and/or learning from data sets; and application to data sets that pose difficulties for human processing and standard statistical models because of their size, dimensionality, heterogeneity,

nonlinearity, and/or otherwise ambiguous, incomplete and/or erroneous structure.

With these elements in mind, but taking into account that a working definition should also always be as simple as possible,⁷ we offer a potential working definition for the scope of AIPV as follows: *AIPV is the development and use of computerized algorithms and supporting technologies that perform operational and scientific tasks to improve PV knowledge and process performance by transacting on or learning from data sets that pose difficulties for human processing and standard statistical analysis.* This potential working definition and the PTCs are complementary and must be viewed together. For example, the working definition contains the term *PV* without elaboration, whereas the operations and information PTCs elaborate on specific PV activities and data sets that might otherwise be overlooked or inappropriately excluded. The PTCs and a working definition should, together, better equip interested parties to understand, discuss, and make scoping decisions regarding AIPV. As with any PTCs and AI scope definitions, they should be considered guidance rather than an immutable prescription regarding AIPV; the potential working definition for the scope of AIPV is offered, in conjunction with the PTCs, to facilitate further deliberation on this topic.

CONCLUSION

This Commentary raises and explores relevant PTCs regarding the scope of AIPV. On the basis of these definitional PTCs, we also offer a potential working definition of the scope of AIPV to, in conjunction with the PTCs, facilitate further deliberation on this topic.

DISCLOSURES

The authors have indicated that they have no conflicts of interest regarding the content of this article.

ACKNOWLEDGEMENT

Many thanks to 2 anonymous reviewers for thought-provoking comments, which improved the quality of the manuscript. Manfred Hauben was responsible for conceptualization, writing (original draft preparation, specifically writing the initial draft), and supervision. Craig Hartford was responsible for writing

(reviewing and editing, specifically critical review), commentary, and revision.

REFERENCES

1. <https://www.dsru.org/international-working-group-on-signal-detection-and-management-in-pharmacovigilance/>. Accessed December 13, 2020.
2. Sucharew H, Macaluso M. Methods for research evidence synthesis: the scoping review approach. *J Hosp Med*. 2019;14:416–418.
3. Shariff SZ, Cuerden MS, Haynes RB, et al. Evaluating the impact of MEDLINE filters on evidence retrieval: study protocol. *Implement Sci*. 2010;5:58.
4. Luo Y, Thompson WK, Herr TM. Natural language processing for EHR-based pharmacovigilance: a systematic review. *Drug Saf*. 2017;40:1075–1079.
5. Legg S, Hutter M. *A Collection of Definitions of Intelligence*. Technical Report; 15-July-2007. IDSAI 07-07 <https://arxiv.org/pdf/0706.3639.pdf>. Accessed December 13, 2020.
6. Gardner H. Frames of mind. In: *The Theory of Multiple Intelligences*. 3rd ed. New York: Basic Books; 29 Mar. 2011. ISBN-13 : 978-0465024339.
7. Wang P. On defining artificial intelligence. *J Artif Gen Intelligence*. 2019;10:1–37.
8. Mitchell T. Introduction. In: *Machine Learning*. Boston: McGraw-Hill; 1997:2, 1997.
9. Montastruc J-L, Sommet A, Bagheri H, Lapeyre-Mestre M. Benefits and strengths of the disproportionality analysis for identification of adverse drug reactions in a pharmacovigilance database. *Br J Clin Pharmacol*. 2011;72:905–908.
10. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Phil Trans R Soc A*. 2016;374:20150202.
11. https://scikit-learn.org/stable/tutorial/machine_learning_map/index.html. Accessed December 13, 2020.
12. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Machine Learn Res*. 2011;12:2825–2830.
13. Hauben M, Reynolds R, Caubel P. Deconstructing the pharmacovigilance hype cycle. *Clin Ther*. 2018;40:1981–1990.
14. Souvignet J, DeClerck G, Asfari H, et al. OntoADR a semantic resource describing adverse drug reactions to support searching, coding, and information retrieval. *J Biomed Inform*. 2016;63:100–107.
15. Alshahrani M, Khan MA, Maddouri O, et al. Neuro-symbolic representation learning on biological knowledge graphs. *Bioinformatics*. 2017;33:2723–2730.

16. Ma C, Zhang HH, Wang X. Machine learning for big data analytics in plants. *Trends Plant Sci.* December 2014;19:796–806.
17. Tsai C-W, Lai C-F, Chao H-C, et al. Big data analytics: a survey. *J Big Data.* 2015;2.
18. https://www.who.int/medicines/areas/quality_safety/safety_efficacy/pharmvigi/en/. Accessed December 13, 2020.
19. Mockute R, Desai S, Perera S, et al. Artificial intelligence within pharmacovigilance: a means to identify cognitive services and the framework for their validation. *Pharm Med.* 2019;33:109–120.
20. Hauben M. The potential of artificial intelligence in pharmacovigilance. *J Fac Pharm Med.* 10 November 2020.
21. Lewis DL, McCallum JF. Utilizing advanced technologies to augment pharmacovigilance systems: challenges and opportunities. *Therap Innov Regul Sci.* 2019.
22. Ghosh R, Kempf D, Pufko A, et al. Automation opportunities in pharmacovigilance: an industry survey. *Pharm Med.* 2020;34:7–18.
23. Marshall IJ, Byron BC. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Syst Rev.* 2019;8:163.
24. Wright T, Peng Y, Easterbrook M. Machine learning to identify multifocal ERG deficits in patients taking hydroxychloroquine. *Invest Ophthalmol Vis Sci.* July 2019;60:5959.
25. Ciallella H, Zhu H. Advancing computational toxicology in the big data era by artificial intelligence: data-driven and mechanism-driven modeling for chemical toxicity. *Chem Res Toxicol.* 2019;32:536–547.
26. Wu Y, Wang G. Machine learning based toxicity prediction: from chemical structural description to transcriptome analysis. *Int J Mol Sci.* 2018;19:2358.
27. Bauer-Mehren, van Mullingen EM, Avilach P, et al. Automatic filtering and substantiation of drug safety signals. *Plos Comput Biol.* 2012;8, e1002457.
28. Xie L, Li J, Xie L, et al. Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. *Plos Comput Biol.* 2009;5, e1000387.
29. Zhaohao S, Strang KD, Rongping L. Big data with ten big characteristics. October 27-29. In: *Proceedings of 2018 the 2nd International Conference on Big Data Research (ICBDR 2018)*. Weihai, China: ACM; 2018:56–61. ISBN 978-1-4503-6476-8. ACM.
30. Chen L, Cai C, Chen V, et al. Trans-species learning of cellular signaling systems with bimodal deep belief networks. *Bioinformatics.* 2015;31: 3008–3015, 2015.

Address correspondence to: Craig Hartford. E-mail: craighartford.1@outlook.com