LEADING ARTICLE

Combining the Power of Artificial Intelligence with the Richness of Healthcare Claims Data: Opportunities and Challenges

David Thesmar[1] · David Sraer[2] · Lisa Pinheiro[3] · Nick Dadson[3] · Razvan Veliche[4] · Paul Greenberg[4]

Abstract

Combinations of healthcare claims data with additional datasets provide large and rich sources of information. The dimensionality and complexity of these combined datasets can be challenging to handle with standard statistical analyses. However, recent developments in artificial intelligence (AI) have led to algorithms and systems that are able to learn and extract complex patterns from such data. AI has already been applied successfully to such combined datasets, with applications such as improving the insurance claim processing pipeline and reducing estimation biases in retrospective studies. Nevertheless, there is still the potential to do much more. The identification of complex patterns within high dimensional datasets may find new predictors for early onset of diseases or lead to a more proactive offering of personalized preventive services. While there are potential risks and challenges associated with the use of AI, these are not insurmountable. As with the introduction of any innovation, it will be necessary to be thoughtful and responsible as we increasingly apply AI methods in healthcare.

Key Points for Decision Makers
The use of artificial intelligence (AI) with claims data has the ability to identify and reduce common biases in healthcare, such as doctor and omitted variable biases.
The ability of AI to detect intricate patterns in claims data pooled with other sources can lead to better care (e.g., detection of underdiagnosed diseases) and insurance coverage processes.
Patient confidentiality, methodological transparency, and potential discrimination are important issues to consider when using AI with claims data.

✉ Lisa Pinheiro
lisa.pinheiro@analysisgroup.com

1 MIT Sloan School of Management, MIT, Cambridge, MA, USA

2 Department of Economics and Haas School of Business, UC Berkeley, Berkeley, CA, USA

3 Analysis Group, Inc., 1190 avenue des Canadiens-de-Montréal, Montreal, QC, Canada

4 Analysis Group, Inc., Boston, MA, USA

△ Adis

1 Introduction

Healthcare claims data provide structured information on patient interactions with the healthcare system, such as treatments given, providers used, billed amounts, and prescriptions filled. They consist of the billing codes that healthcare providers, such as physicians and hospitals, submit for payment by commercial and government health plans. Since they are collected regularly for administrative purposes, healthcare claims provide a relatively inexpensive source of information, over long time periods, for large numbers of patients. Moreover, as our payment structures change and patients ask more of their providers, the volume of available claims data is likely to increase rapidly.

The large sample sizes can be useful for studying rare conditions and the longitudinal dimension of the data allows researchers to examine treatment adherence and effects over time. However, healthcare claims typically provide only limited information on clinical severity, patients' health status, and other variables of interest [1]. This can be ameliorated by supplementing claims data via linkages with other data sources, such as census data to account for neighborhood income or electronic medical records (EMRs) to obtain more clinical information [2].[1] These linkages open the door for a

---

[1] Another, more recent, source of information that could possibly be combined with claims data is 'telehealth' technology. This technology has made it possible for individuals to use their computers and

number of research possibilities, as relevant healthcare data can come in a variety of unstructured formats, such as text, pictures, audio, or video files.[2]

While linked datasets permit a more comprehensive comparative analysis of alternative health technologies and interventions [4], incorporating the additional dimensionality and complexity of the data into the analysis can present challenges for standard statistical analysis [5]. In particular, explicit functional relationships in novel combinations of data may be unknown ex ante. Recent developments in artificial intelligence (AI) research have led to computer algorithms and systems with the ability to learn and extract complex patterns from raw data [6]. For example, machine learning algorithms, such as random forests and deep convolutional neural networks, often detect intricate relationships between input and output variables to improve out-of-sample predictions. Natural language processing algorithms can be used to parse and interpret written text. AI algorithms have been applied successfully in many domains of business, government, and science [7].

There are important differences in emphasis between standard statistics and AI algorithms. Standard statistics emphasize our understanding of underlying mechanisms (i.e., fitting a specific model and hypothesis testing). While AI algorithms can often detect unforeseen relationships and complicated nonlinear interactions within the data, their emphasis is typically on prediction accuracy (e.g., identifying the best treatment or course of action) [5].

Multiple approaches have been developed to palliate concerns about the interpretability of AI algorithms. For example, recent research has developed measures of the importance of specific covariates based on how much each contributes to the prediction accuracy of the model(s) [8, 9].[3] Like standard statistical methods, these importance measures can be biased based on the data used or the variables included in the model. However, they do allow researchers to tie predictive AI models back to intuitive relationships and existing knowledge.[4] Another approach is to develop

methods that ally the flexibility of AI and the interpretability of simpler models, such as optimal classification tree methods.[5] These approaches are particularly valuable in healthcare settings, in which it can be particularly important to understand the decision process behind a decision, such as a diagnosis or treatment.

In this paper we discuss some recent and potential applications of AI to healthcare claims data.[6] AI has the potential to improve the detection of diseases and treatment adverse effects, identify and reduce diagnostic errors or personal biases, and improve the monitoring of insurance costs and fraud. However, this potential is associated with risks and challenges. Patient confidentiality, methodological transparency, and potential discrimination are important issues to consider when using AI with claims data.

2 Making the Most of Healthcare Claims Data

'Artificial intelligence,' broadly defined, is the ability for machines to perform tasks characteristic of human intelligence. In this paper we focus on machine learning algorithms, which are a subset of AI. Rather than researchers assuming explicit functional relationships between variables, machine learning algorithms are designed to learn relationships from the data. These learned relationships can be exploited to improve data analytics relying on claims data in numerous ways. Some examples are presented in this section.

Footnote 1 (continued)

phones to access healthcare services remotely. For example, patients can potentially use their phones to upload their food logs, medications, or dosing for review by a doctor or nurse [3].

[2] Examples include doctors' notes, computed tomography (CT) scans, and video recordings of clinical encounters.

[3] These measures permute the values of the covariate, breaking the relationship between the covariate and outcome, and calculate the corresponding increase in the prediction error. A covariate is considered more important the more its permutation increases prediction errors, since this implies the model relied on the covariate for prediction.

[4] Questions relating to statistical significance are also often approached differently in the AI world. AI algorithms are typically

---

Footnote 4 (continued)

used with very large datasets, in which it is often possible to find statistical significance despite a variable being associated with very small differences in the outcome variable. In certain applications it may therefore be more useful to determine variables' economic or clinical significance rather than focusing solely on their statistical significance.

[5] Researchers continue to improve the predictive ability of machine learning algorithms. Among these algorithms, decision trees are often preferred due to their interpretability. Optimal classification trees are an example of recent research that improves predictions while maintaining interpretability [10].

[6] Our discussion in this paper is not intended to be exhaustive as the set of potential applications is likely quite large. For example, while we do not discuss it in depth here, AI can possibly help alleviate shortages of clinical staff by taking on some of their responsibilities (e.g., the diagnostic responsibilities of neurologists). In this way, AI may increase the access to healthcare, particularly in remote regions where it may be difficult to attract practitioners.

2.1 Pooling Knowledge to Provide Better Care

A significant advantage of AI over standard statistical analysis is its ability to examine large multidimensional data with many variables. A statistician may easily be overwhelmed by a large number of potentially usable variables. AI algorithms, in contrast, are often able to identify which variables are important and detect the optimal combination of these variables for the task at hand.

By combining claims data with other information sources, such as patients' laboratory results and EMRs, the knowledge and experience of a wide range of medical practitioners can be pooled together to detect complex patterns indicative of certain illnesses. The identification of these patterns can improve the early detection of diseases and the detection of underdiagnosed or rare diseases, provide more accurate diagnoses, or lead to a more proactive offering of personalized preventive services.

In the case of conditions that are underdiagnosed, for example, a sample of patient claims data (combined with laboratory values, specialist visits, and physician notes) can be analyzed by physicians to determine which of the patients should have been diagnosed and at what point in time. This process can then be used to train algorithms to detect other instances that should be flagged and identify early markers and indicators of future disease onset. For example, recent research has used AI to predict type 2 diabetes mellitus using claims data [11]. More generally, there have been a number of applications using AI with other types of data (e.g., CT scans, or genome sequencing in precision medicine) to support diagnoses, either through the prediction of disease onset or the identification of drug-resistant strains of disease (e.g., drug-resistant tuberculosis [12]).

Another promising application of AI is the detection of treatment adverse effects. Adverse effects are typically hard to discover in standard human trials because of relatively small sample sizes. Large claims data may help discover these adverse effects, in particular when combined with the pattern recognition power of AI. Recent research has used machine learning algorithms to detect adverse drug reactions using EMR data [13]. Healthcare administrative data can be linked with EMR data to add more clinical information [2], and may improve detection. More generally, as discussed by Dadson et al. [14], AI has been used to detect drug–drug interactions based on the text in government and scientific reporting systems and may be used with genomic data to determine genetic factors associated with particular adverse effects.

AI has also been used broadly to systematically learn from past experience. Examining patterns in pooled data can improve the monitoring of services, such as hospital readmissions or diseases contracted on site. For example, machine learning analysis on patients' electronic health records, demographics, medical histories, admission and hospitalization histories, and likely exposure to the *Clostridium difficile* bacterium was shown to increase the accuracy in predicting the risk of hospital infections [15].

2.2 Identifying and Reducing the Effects of Biases

AI also has the potential to identify and reduce the effects of common biases in healthcare, such as doctor biases and omitted variable bias.

2.2.1 Doctor Bias

Doctors may exhibit biases in their propensities to prescribe particular drugs. These biases may result from the limited time doctors have to make a diagnosis, the information doctors are exposed to (e.g., limited samples, medical press, or patient history), doctors' own, potentially self-reinforcing, experience (e.g., repeated diagnoses), or they may result from specific cognitive biases or personality traits (e.g., aversion to risk or ambiguity) or forecasting biases (e.g., over-confidence or belief in the law of small numbers) [16].

As described earlier, AI used on claims data pooled with information from other sources can provide support for better care. Taking advantage of more knowledge and varied experience can reduce doctors' biases in decision-making. Doctors could be provided with a statistical diagnosis tool using models calibrated with data. For example, AI algorithms could scan patient histories, data, and relevant literature as a doctor is entering observations and notify the doctor of potentially useful information. In some settings, the diagnostic tool could also incorporate prior expert knowledge and decision-making into a predictive AI algorithm [17]. Although this tool would be limited to the extent that the relevant data can be encoded and made machine-readable, it could complement the doctor's judgement with unbiased information [18].

2.2.2 Omitted Variable Bias

Claims data often contain information about specific treatments not available via randomized trials [19]. Using this information, algorithms can identify covariates correlated to both outcomes and treatment that previous researchers did not realize were important. This can reduce bias by ensuring no pertinent variables are omitted from the statistical analysis. In addition, AI prediction tools also feature a variety of methods to avoid overfitting and exclude superfluous variables.

Despite its richness, there is some potentially relevant information that may be missing or limited in claims data, even after pooling it with other sources, which can lead to omitted variable bias. The lack of lifestyle characteristics

information is often cited as a main limitation of claims data [20]. Information on how the recommended treatment was followed by the patient or the 'quality of care' may also be missing. The ability of AI to find complex patterns in the data can potentially approximate this missing information via combinations of the variables that are available. This can improve the matching of treated and untreated patients, which in turn helps correct for treatment selection biases in retrospective studies of treatment efficacy or safety [21].

For example, it is notoriously difficult to compare the effect of different treatments based on retrospective studies. The decision to prescribe one treatment over another is generally informed by a doctor's evaluation, and the factors that affect that evaluation, such as the patient's disease severity, other co-morbidities, and history of compliance, may be unobservable to the researcher. If, say, one treatment tends to be used for more severe cases, comparing its efficacy (or safety) against another treatment without controlling for this tendency will bias the results in favor of the treatment that is typically used for easier cases.

Propensity score matching is one statistical technique used to limit this bias. The method essentially relies on a two-step process in which the probability that the doctor will prescribe a given treatment is estimated in the first step and then used in second step to account for differences in patient characteristics that affected the prescribing decision. An intuitive way to think about the first step is that it is an attempt to mimic (or predict) the doctor's prescribing decision based on patient information. The better the prediction in the first step of the procedure, the smaller the bias in the overall comparison of treatments. Some researchers have advocated the use of a 'chain of proxies' often found in claims data to improve prediction (e.g., old age may serve as a proxy for co-morbidity or cognitive decline) [22]. Through the detection of intricate patterns amongst these proxies, AI methods can improve predictions even further. Recent research demonstrates that significant bias can remain in many claims data applications after using conventional methods and that the use of AI can significantly reduce, and essentially eliminate, such biases [21, 23, 24].

2.3 Improving Insurance Coverage Processes and Fairness

AI has had an influence on each step of the processing pipeline for insurance coverage: claim submission, claim adjudication, and fraud monitoring. Its influence is likely to continue well into the future.

An efficient and accurate intake process for claims submissions can reduce burdens on patients and the entire medical system. For example, AI image recognition can be used to facilitate the process of submitting and coding claims. One insurance company has trialed a procedure to automate claim approvals. Patients electronically submit photos of their hospital bills and, within moments, receive notification of receipt, approval, and credit to their account [25]. The cost savings associated with this automation process may allow insurance companies to scale up and provide more coverage. Some insurance companies have developed algorithms that recommend and incentivize healthy habits and behaviors to policyholders, such as exercise and nutrition strategies [26]. If successful, these may avoid claims submissions for preventable illnesses altogether.

Claim adjudication involves checking coverage, limits, contracts with providers, pharmacies, appropriate diagnosis, and procedure coding. Complex and suspicious claims trigger secondary, usually manual, processing steps. AI can save time via the efficient processing of normal claims (e.g., through increased automation in the settlement of claims based on their complexity and known patient history) [27]. In addition, AI can be used for early detection of abnormal price patterns among healthcare providers. This may help insurance companies monitor their costs and understand whether price increases reflect actual quality improvements.

Increasing the accuracy of triaging claims through AI can reduce losses due to fraud (e.g., a provider billing for services that were not actually provided or the provision of unnecessary medical services to generate insurance payments). While the share of actual fraudulent claims is generally very low [28], these can add up to very significant dollar amounts. Just counting savings to Medicare, the Health Care Fraud Unit of the US Department of Justice's Criminal Division recovered and transferred billions of dollars in the 2017 fiscal year [29]. Several data mining efforts have been undertaken with increasing levels of sophistication to improve fraud detection [30]. A Google Patent search for "healthcare machine learning fraud detection" reveals over 7700 results [31]. The general challenge is to limit false positives, which can lead to losses in reputation of targeted firms as well as enforcement efforts that are channeled inappropriately, while increasing the proportion of fraud detected. The rules to detect individual fraud are now relatively simple to implement. Recently, more complex network fraud pattern recognition algorithms have been developed with the potential to significantly reduce the costs for insurers [32].

3 Risks and Challenges

While there is great potential for AI applications with claims data, this potential is associated with risks and challenges. Issues related to confidentiality, clarity of scope, appropriate methodology, and transparency of results need to be addressed.

Table 1 Examples of benefits and challenges

A. Benefits

Pooling knowledge
  Early detection of disease and readmission
  Identification of underdiagnosed/rare diseases
  Personalized preventive services
  Monitoring of adverse effects

Reducing biases
  Doctor biases
  Omitted variable biases

Improving the insurance pipeline
  Claim submissions
  Claim adjudication
  Fraud monitoring

B. Challenges

Data provenance
  Confidentiality
  Need for exhaustive, unbiased data

Interpretation of methodological approaches
  Hasty generalization
  Lasting effects of real-life decisions

Discrimination
  Premiums based on personal traits
  Reproduction of biases

3.1 Data Provenance: Confidentiality and Quality

Most ethical problems related to confidentiality can be dealt with by anonymizing the data. In the past, removing social security numbers, phone numbers, names, and most date of birth information was sufficient to anonymize data. However, the recent emergence of third-party datasets (e.g., social media and cell phone tower access records) has provided additional avenues to triangulate datapoints and identify individuals within the data [33]. This creates additional challenges and a potential need for better de-identification techniques that ensure that shared information cannot be used to approximate personal information even when combined with pattern detection techniques.

3.2 Data Use: Interpretation of Methodological Approaches

As is the case with standard statistical methods, AI algorithms built on biased data will lead to biased predictions. Supervised machine learning algorithms, for instance, learn relationships between outcome and predictor variables after being trained on a set of examples for which the true outcome is known. Algorithms will have trouble classifying new examples that are very different to what they have seen

before. In machine learning this problem is referred to as "hasty generalization" [34]. While this problem is not unique to AI, the rapid speed at which individuals are now able to analyze large datasets has raised concerns that some believe the need for evidence of causal relationships is unnecessary, large databases by themselves are enough, and that the numbers will speak for themselves [35]. While complex correlation patterns are useful for making predictions, they may be of limited use for explaining *why* certain events occur [36].[7] The importance measures [8, 9] and highly interpretable AI models [10] discussed earlier can be useful in this regard.

Using AI or any statistical tool requires careful consideration of both the underlying data and the results. There are inherent biases in most claims databases due to sample selection (e.g., the characteristics of the insured population and varying standards of care), such that results may not be generalizable to other population groups. For instance, evaluating the effect of a drug that requires a very demanding treatment on a subpopulation with a high adherence rate may overestimate the effectiveness of the drug on the general population. In addition, the decisions made in real-life applications based on an algorithm's results can have lasting effects (e.g., if an individual is incorrectly declined insurance because their profile shares some characteristics with at-risk groups) [39].

Supplementing claims data with information from other sources may limit the potential bias from unobserved variables. As discussed earlier, particularly rich data may also be used to triangulate and proxy for information that is still missing. However, rather than relying on technology to solve potential issues ex post, recent research has investigated whether technology can help gather more exhaustive, rich, and less biased datasets ex ante. For example, the blockchain technology associated with bitcoin is being considered as a potential solution to securely store healthcare data. Individuals would have ownership over their own personal data and would potentially receive financial incentives to make it widely available. Blockchain technology could also help document the provenance, veracity, and selection of the data used in studies. This could greatly reduce the widespread difficulty in replicating published medical research. Relatedly, this documentation could reduce the ability to cherry-pick results [40].

---

[7] In addition, for a new machine learning algorithm called estimating the maximum (EMX), researchers have found situations in which it is impossible to prove mathematically whether or not the algorithm could actually solve a particular problem. These findings suggest that some level of caution should be exercised when using new AI methods [37, 38].

3.3 Usage: Prevention of an Adverse Welfare Impact from Discrimination

Just as AI algorithms can help detect a disease in its earliest stage, they can also help predict future health costs and therefore be used, in theory, to adjust insurance premiums based on combinations of personal characteristics. If an insurer can predict illness reliably, competition could lead to lower premiums for predictably healthier individuals and higher premiums for predictably sicker ones, even if the characteristics used for prediction are beyond an individual's control.

Even if such a process could be appropriately implemented, it may be economically inefficient. An individual who is uncertain about their future sickness would want to smooth their future income across possible states of nature (e.g., sick vs. not sick). The individual would want insurance, but if the insurance company has reliable forecasting technology, competition could undermine that person's ability to obtain insurance, leaving the individual worse off.[8]

The extent to which such insurer strategies are already used and how much should be prohibited by law is often the subject of debate, particularly since similar issues appear in other industries. For example, research on credit access has examined the use of variables correlated with race and gender when evaluating loan requests and the potential bias against certain groups [42, 43].

At issue in these cases is not the methodology itself but rather its use. Let us assume, for example, that loans are evaluated by a group including individuals with biases (e.g., against immigrant or older loan applicants). If an algorithm is trained to mimic the decisions of these individuals (i.e., predict the loan approval decision) it will implicitly reproduce these biases. However, if the algorithm is trained to predict the future return of the investments made (i.e., the loans), the algorithm may show that unbiased decisions lead to better outcomes and may therefore provide better recommendations. In that case the AI algorithms might, for example, better align loan decisions with firms' long-term profit interests while reducing biases [44].

AI methodologies are therefore not inherently good or bad, biased or clean. They are what we make of them. Interestingly, AI and claims data can be used to analyze the scope of the potential insurance discrimination problem, by examining the insurance premium distribution and identifying the groups that benefit for different scenarios of harmonization versus tailoring of premium rates.

4 Conclusion

Linkages of claims data with additional datasets provide rich sources of healthcare information. Given the complexity and scale of these datasets, it is unsurprising to observe that healthcare data have frequently been used in tandem with AI methods (Table 1). These methods can detect intricate patterns within complex data, improving their own performance through experience. For example, using claims data, laboratory results, and EMRs, the experiences of a large set of medical practitioners can be consolidated and used to detect patterns indicative of certain illnesses.

While these disease patterns can be used to improve the quality of care, this knowledge can also be used by doctors as a relatively unbiased complement to their decision-making. AI has the potential to correct for biases due to omitted variables as well, using the complex patterns detected in the data to approximate the missing information with the information that is available. This has significant potential for use in numerous applications, such as retrospective studies comparing the effectiveness of different treatments.

If insurers, providers, and regulators can work together to increase the linkages between datasets, in a way that respects privacy, the capabilities driven by AI can lead to cost savings and quality improvements throughout the healthcare industry.

There are risks and challenges associated with the increased application of AI to healthcare claims and related datasets (Table 1). For example, AI has had a strong influence in the health insurance coverage process. Claim submission, claim adjudication, and fraudulent claim detection have either already improved or have the potential to improve due to the use of AI. However, the predictive ability of AI can also be used to adjust insurance premiums in undesirable ways based on combinations of personal characteristics. More generally, other potential issues involve the confidentiality of data and appropriate use of AI tools.

These risks and challenges are not insurmountable, however. They only imply that, as we increasingly apply more tools from AI, we need to be careful and responsible.

Compliance with Ethical Standards

---

[8] This is called the Hirshleifer effect [41].

References

1. Birnbaum HG, Cremieux PY, Greenberg PE, LeLorier J, Ostrander J, Venditti L. Using healthcare claims data for outcomes research and pharmacoeconomic analysis. Pharmacoeconomics. 1999;16(4):1–8.

2. Cadarette SM, Wong L. An introduction to health care administrative data. Can J Hosp Pharm. 2015;68(3):232–7.

3. Mayo Clinic Staff. Telehealth: technology meets health care. Mayo Clinic; 2017 Aug 16. https://www.mayoclinic.org/healthy-lifestyle/consumer-health/in-depth/telehealth/art-20044878. Accessed 30 Jan 2019.

4. Onukwugha E, Jain R, Albarmawi H. Evidence generation using big data: challenges and opportunities. In: Birnbaum HG, Greenberg PE, editors. Decision making in a world of comparative effectiveness research: a practical guide. Singapore: Springer Nature Singapore Pte Ltd.; 2017. p. 253–63.

5. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods. 2015;15(4):233–4.

6. Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge: MIT Press; 2016.

7. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436–44.

8. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.

9. Fisher A, Rudin C, Dominici F. All models are wrong but many are useful: variable importance for black-box, proprietary, or mis-specified prediction models, using model class reliance; 2018 Nov. https://arxiv.org/pdf/1801.01489.pdf. Accessed 22 Jan 2019.

10. Bertsimas D, Dunn J. Optimal classification trees. Mach Learn. 2017;106(7):1039–82.

11. Razavian N, Blecker S, Schmidt AM, Smith-McLallen A, Nigam S, Sontag D. Population-level prediction of type 2 diabetes from claims data and analysis of risk factors. Big Data. 2015;3(4):277–87.

12. Chen ML, Doddi A, Royer J, Freschi L, Schito M, Ezewudo M, et al. Deep learning predicts tuberculosis drug resistance status from genome sequencing data. bioRxiv 275628 (preprint); 2018 Jun. https://www.biorxiv.org/content/10.1101/275628v2.

13. Jeong E, Park N, Choi Y, Park RW, Yoon D. Machine learning model combining features from algorithms with different analytical methodologies to detect laboratory-event-related adverse drug reaction signals. PLoS One. 2018;13(11):1–15.

14. Dadson N, Pinheiro L, Royer J. Decision making with machine learning in our modern, data-rich health care industry. In: Birnbaum HG, Greenberg PE, editors. Decision making in a world of comparative effectiveness research: a practical guide. Singapore: Springer Nature Singapore Pte Ltd.; 2017. p. 277–89.

15. Slabodkin G. Machine learning, HER data helping to combat hospital infections. Health Data Management; 2018 Apr 3. https://www.healthdatamanagement.com/news/machine-learning-ehr-data-helping-to-combat-hospital-infections. Accessed 19 Oct 2018.

16. Saposnik G, Redelmeier D, Ruff CC, Tobler PN. Cognitive biases associated with medical decisions: a systematic review. BMC Med Inform Decis Mak. 2016;16(138):1–14.

17. Gibert K, Garcia-Alonso C, Salvador-Carulla L. Integrating clinicians, knowledge and data: expert-based cooperative analysis in healthcare decision support. Health Res Policy Syst. 2010;8(28):1–16.

18. Seidman AD, Pilewskie ML, Robson ME, Kelvin JF, Zauderer MG, Epstein AE, et al. Integration of multi-modality treatment planning for early stage breast cancer (BC) into Watson for oncology, a decision support system: seeing the forest and the trees. J Clin Oncol. 2015; 33(S15). http://ascopubs.org/doi/abs/10.1200/jco.2015.33.15_suppl.e12042.

19. Franklin JM, Schneeweiss S, Polinski JM, Rassen JA. Plasmode simulation for the evaluation of pharmacoepidemiologic methods in complex healthcare databases. Comput Stat Data Anal. 2014;72:219–26.

20. Stein JD, Lum F, Lee PP, Rich WL, Coleman AL. Use of health care claims data to study patients with ophthalmologic conditions. Ophthalmology. 2014;121(5):1134–41.

21. Royer J, Merrigan P, Brown K. Estimating average treatment effects with propensity scores estimated with four machine learning procedures: simulation results in HD settings and with time to event outcomes. SSRN; 2018 Sep. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3272396. Accessed 23 Jan 2019.

22. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology. 2009;20:512–22.

23. Karim ME, Pang M, Platt RW. Can we train machine learning methods to outperform the high-dimensional propensity score algorithm? Epidemiology. 2018;29(2):191–8.

24. Franklin JM, Eddings W, Austin PC, Stuart EA, Schneeweiss S. Comparing the performance of propensity score methods in healthcare database studies with rare outcomes. Stat Med. 2017;36(12):1946–63.

25. GovInsider. AI is changing healthcare – and insurers are taking notice; 2018. https://govinsider.asia/inclusive-gov/ai-changing-healthcare-insurers-taking-notice/. Accessed 19 Oct 2018.

26. Institute of International Finance. Innovation in insurance: how technology is changing the industry; 2016 Sep. https://www.iif.com/portals/0/Files/private/32370132_insurance_innovation_report_2016.pdf. Accessed 19 Jan 2019.

27. Hehner S, Kors B, Martin M, Uhrmann-Klingen E, Waldron J. Artificial intelligence in health insurance: smart claims management with self-learning software. McKinsey & Company; 2017 Sep. https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/artificial-intelligence-in-health-insurance-smart-claims-management-with-self-learning-software. Accessed 19 Oct 2018.

28. LexisNexis. Bending the cost curve: analytics-driven enterprise fraud control; 2011 Apr. http://lexisnexis.com/risk/downloads/idm/bending-the-cost-curve-analytic-driven-enterprise-fraud-control.pdf. Accessed 19 Oct 2018.

29. Office of Inspector General, U.S. Department of Health and Human Services. Health care fraud and abuse control program annual report for fiscal year 2017; 2018 Apr. https://oig.hhs.gov/publications/docs/hcfac/FY2017-hcfac.pdf. Accessed 23 Jan 2019.

30. Kose I, Gokturk M, Kilic K. An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance. Appl Soft Comput. 2015;36:283–99.

31. Google. Google Patents search: "healthcare machine learning fraud detection"; 2018. https://patents.google.com/?q=healthcare&q=machine+learning&q=fraud+detection&oq=healthcare+machine+learning+fraud+detection. Accessed 14 Sept 2018.

32. Liu J, Bier E, Wilson A, Guerra-Gomez JA, Honda T, Kumar Sricharan, et al. Graph analysis for detecting fraud, waste, and abuse in health-care data. AI Mag. 2016;Summer:33–46.

33. Berinato S. There's no such thing as anonymous data. Harvard Business Review; 2015 Feb 9. https://hbr.org/2015/02/theres-no-such-thing-as-anonymous-data. Accessed 19 Oct 2018.

34. Wallace BC, Trikalinos TA, Lau J, Brodley C, Schmid CH. Semi-automated screening of biomedical citations for systematic reviews. BMC Inform. 2010;11(55):1–11.

35. Anderson C. The end of theory: the data deluge makes the scientific method obsolete. Wired; 2008 Jun 23. https://www.wired.com/2008/06/pb-theory/. Accessed 19 Oct 2018.

36. Williams BA, Brooks CF, Shmargad Y. How algorithms discriminate based on data they lack: challenges, solutions, and policy implications. J Inf Policy. 2018;8:78–115.

37. Ben-David S, Hrubes P, Moran S, Shpilka A, Yehudayoff A. Learnability can be undecidable. Nat Mach Intell. 2019;1:44–8.

38. Reyzin L. Unprovability comes to machine learning. Nature. 2019;565:166–7.

39. O'Neil C. Weapons of math destruction: how big data increases inequality and threatens democracy. New York: Broadway Books; 2016.

40. Sbeglia C. Microsoft delves into using blockchain as part of a process flow. Research and Development. 2018 Aug 1. https://www.rdmag.com/article/2018/08/microsoft-delves-using-blockchain-part-process-flow. Accessed 19 Oct 2018.

41. Hirshleifer J. The private and social value of information and the reward to inventive activity. Am Econ Rev. 1971;61(4):561–74.

42. Ongena S, Popov A. Gender bias and credit access. J Money Credit Bank. 2016;48(8):1691–724.

43. Waddell K. How algorithms can bring down minorities' credit scores. The Atlantic; 2016 Dec 2. https://www.theatlantic.com/technology/archive/2016/12/how-algorithms-can-bring-down-minorities-credit-scores/509333/. Accessed 19 Oct 2018.

44. Doobie W, Liberman A, Paravisini D, Pathania V. Measuring bias in consumer lending. NBER Working Paper Series; 2018 Aug. https://www.nber.org/papers/w24953. Accessed 15 Oct 2018.