



The role of big-data in clinical studies in laboratory medicine

Zhongheng Zhang

Department of Emergency Medicine, Sir Run-Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310016, China

Correspondence to: Zhongheng Zhang. Department of Emergency Medicine, Sir Run-Run Shaw Hospital, Zhejiang University School of Medicine, No. 3, East Qingchun Road, Hangzhou 310016, China. Email: zh_zhang1984@hotmail.com.

Abstract: The primary aim of laboratory medicine is to make an accurate diagnosis or risk stratification. A diagnostic tool is said to be useful if its diagnostic performance is good, but it cannot be routinely recommended for clinical use unless it also provides clinical benefits in patient-important outcomes. The latter condition is the effectiveness of the diagnostic tool. While the gold standard of effectiveness relies on randomized controlled trials (RCTs), big-data clinical study employing electronic healthcare records (EHR) provides an alternative to provide evidence to guide decision-making. RCTs are limited by their strict inclusion/exclusion criteria, high cost and ethical constraint. EHR contains data on patient-level granularity and can help to disentangle complex research questions. The Achilles' heel of observational studies (e.g., big-data study is a kind of observational study in nature) is uncontrolled confounding. Although it is still impossible to control for unmeasured confounding factors, the re-use of EHR helps to control as much confounding as possible. As compared with registry and administrative databases, EHR contains more detailed information on demographics, drugs, procedures and operations. Recent decades have witnessed exponential increase in the application of machine learning techniques in laboratory medicine and these novel techniques continue to provide advanced tool for better prediction of diseases.

Keywords: Big data; laboratory medicine; machine learning

Received: 12 March 2017; Accepted: 05 June 2017; Published: 22 June 2017.

doi: 10.21037/jlpm.2017.06.07

View this article at: <http://dx.doi.org/10.21037/jlpm.2017.06.07>

Evidence pyramid in laboratory medicine

Laboratory medicine aims to provide tests to guide clinical decision making. Generally speaking, it doesn't provide direct treatment and/or interventions and its effectiveness can only be reflected from its ability to classify patients into those who will benefit from treatment and those whose will not. Some biomarkers may be useful to identify patients with certain disorders, but there may not exist an effective treatment for such disorders, and these biomarkers will be considered to be clinically effective (1). The usefulness is not equivalent to effectiveness in the paradigm of evidence-based medicine. For instance, the pulse-indicated continuous cardiac output (PiCCO) device is a powerful tool in bedside monitoring of hemodynamics for critically ill patients (2). It provides dozens of hemodynamic parameters that helps doctors to gain a global view on circulatory status including cardiac function, vascular resistance, pulmonary edema

and volume status (3). However, these parameters have not successfully translated into solid clinical benefits for those critically ill patients (4). In other words, PiCCO device is useful to obtain a global assessment of hemodynamic status, but clinicians lack treatment strategies to translate the knowledge into clinical benefits. One successful example in laboratory medicine is the use of procalcitonin (PCT) to guide antibiotic therapy (5). It has long been recognized that PCT is a specific biomarker of bacterial infection (6,7). Although there are many confounding factors that may influence the linkage between PCT and infection, PCT is much better than C-reactive protein. Thus, it is considered to be useful in identifying infection. More recently, there are randomized controlled trials (RCT) showing that discontinuation of antibiotics according to the decrease in PCT can help to reduce the use of antibiotics but not increase the risk of infection relapse (8-10). The lesson we can learn from the PCT case is that the high-grade evidence

for the effectiveness of laboratory biomarkers should be based on RCTs employing patient outcomes as the study end-points such as mortality, morbidity, quality-of-life or cost of treatment. Without high quality RCTs, these biomarkers can only be useful in identifying certain type of disorders, but they can never be recommended strongly for clinical use. In the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework, the diagnostic accuracy as represented by sensitivity and specificity is considered to provide lower level of evidence than other patient-important outcomes (11,12). Brain natriuretic peptide (BNP) is another example in laboratory medicine that it can help to distinguish patients with dyspnea from respiratory or cardiac origin (13-16). A BNP level of greater than 500 ng/L should direct clinicians to start treatment for heart failure, which in turn translates to improved patient outcomes (16,17). However, BNP is only recommended for use in emergency conditions and its utility has not been validated in in-hospital or critical care settings. The diagnostic accuracy of a test can be influenced by the target population. In critical care setting, the patients typically suffer from multiple organ failure, which increases the burden of confounding.

Predicament in current evidence based medicine

In the paradigm of current evidence based medicine, high-grade evidence is based on well-designed RCTs which are costly, labor-intensive and sometimes unethical to perform (18,19). As a result, the generation of high-grade evidence cannot meet the need of daily clinical practice. In other words, most clinical practices involving ordering blood tests and other laboratory measurements are not based on empirical evidence. For instance, point-of-care ultrasound is an area of intensive research at present because the ultrasound can provide a large amount of information on organs and hemodynamic status of critically ill patients (20,21). However, there is a lack of strong evidence that such measurement can be translated into benefits in patient-important outcomes such as mortality. Nevertheless, clinicians are still keen on performing point-of-care ultrasound, advocating that ultrasound is a stethoscope in the 21 century (22-24). We need to pay attention to the fact that there is lack of evidence for such a diagnostic tool, although the ultrasound imposes no direct harm to patient it bears an opportunity cost.

Cardiology is a specialty with advanced evidence-based

medicine, but scientific evidence underlying cardiology practice is mostly based on low level of evidence. Tricoci and colleagues have examined the recommendations issued by the American College of Cardiology (ACC) and the American Heart Association (AHA) and found that only 314 of the 2,711 examined recommendations (11%) are classified as level of evidence A, whereas 1,246 (48%) are at the level of evidence C (25). The situation is similar in oncology practice (26). There is no data on laboratory medicine, but the situation cannot be better. There is a large body of evidence showing that clinicians frequently raise questions in their daily practice, but roughly half of them are never pursued (27-30). The primary reasons that the doctors did not continue to pursue answers to the questions were the doubt that a useful answer exists (31). There is a large gap between the needs of high-level evidence in real-world practice and the lack of such evidence generated by high quality RCTs.

Potential application of clinical database in laboratory medicine

While RCTs are gold standard of clinical practice, they cannot cover all decisions that clinicians made on a daily basis. Reasons prohibiting the performance of RCTs includes, but not limited to, ethical issues, constrained financial and human resources (32,33). An alternative to RCTs is the employment of electronic databases generated by other purposes such as insurance, administration and registry. With the development of information technology and big-data processing techniques, more and more electronic big-data are available to investigators (34-36). These clinical databases are typically large in sample sizes, providing an important source of evidence to guide clinical practice.

Electronic healthcare records (EHR) have the patient level granularity of complete data to address complex research questions, which are not realizable by using traditional discharge records or registry databases. Some databases are created for specific purposes and lack sufficient variables to control for confounders in complex research questions. For example, the Dartmouth Atlas utilize Medicare collected data to monitor the medical costs and patient outcomes (<http://www.dartmouthatlas.org/>), which cannot be employed to investigate complex interaction between laboratory variables and treatments. The Laboratory Information Management System (LIS) has been widely used in Chinese hospitals for several

decades, and the system has stored millions of laboratory values that can be linked to EHR system (18). The latter system contains medical orders, discharge outcomes, operations and treatments. As a result, laboratory values can be associated with interventions and outcomes, which provided opportunity to answer complex clinical questions. Our previous studies have utilized such information systems to explore the association between platelet indices and mortality risk (37), the causal relationship between serum chloride levels and development of acute kidney injury (AKI) (38). These studies included thousands participants, which is infeasible by designing a prospective cohort study and data being collected by hand. Although such studies cannot provide high-level evidence for clinical practice according to GRADE guideline, they provide insights into interactions between laboratory values and clinical outcomes. The next step will be to investigate for instance, whether the administration of balanced fluid (less chloride) compared with chloride-rich fluid is able to reduce the incidence of AKI.

Traditional techniques to control for confounding

The confounding issues are the Achilles heel in observational studies. Studies employing existing databases are observational in nature, in which the control of unmeasured confounding factors is impossible. However, all conventional techniques to control confounding are applicable to big-data research. These techniques include stratification, matching, multivariable regression and propensity score analysis (39,40). In conventional cohort studies, because the sample size is limited and thus variables employed to control for confounding cannot exceed a certain number. In multivariable logistic regression, the number of covariates are typically less than one tenth of the event of interest (41,42). Also, the variables collected by hand are predefined for the primary research purpose, which markedly limit the secondary analysis for another research question. In a study investigating the effectiveness corticosteroids in patients with acute respiratory distress syndrome (ARDS), we employed data from a RCT evaluating the effectiveness of statins. Thus, many confounding factors that may influence the use of corticosteroids are lacking (43). By employing EHR, there will be no such restriction because the EHR contains nearly all information recorded during clinical practice (44). The large sample size of clinical databases enables the inclusion of more covariates in multivariable regression.

Machine learning in laboratory medicine

Machine learning methods are probably the most popular in the area of big data. A large proportion of clinicians' endeavor in medical decision making is classification. All these questions are does my patient have the disease? In what stage does it belong to? Which group of patients can benefit from the treatment? All these questions related to classification or diagnosis. In many fields of diagnostics, machine learning algorithms appear to be superior or equivalent to the most sophisticated doctors in making correct diagnosis. For example, scientists have succeeded in building a deep learning algorithm which is able to detect diabetic retinopathy with the sensitivity and specificity of 96.1% and 93.9%, respectively (45). Similarly, a machine learning method called convolutional neural networks was able to identify skin cancer with a level of competence comparable to sophisticated dermatologists (46).

Since the major task of laboratory medicine is to make diagnosis or identify patients who will benefit the most from interventions, the machine learning methods have extensive applications. Actually, the application of machine learning methods in laboratory medicine has witnessed an exponential increase in recent years. PubMed was searched with the key words "laboratory medicine" and "machine learning" and found that the use of machine learning is increasing (*Figure 1*). Although the search strategy is not systematic and comprehensive, this result indicates the enthusiasm on the application of machine learning in recent years.

Typical examples of machine learning techniques in laboratory medicine include to enhance understanding of the relationship between gamma glutamyl transferase and other components of liver function, improve the prediction of hepatitis C and B, and the association of bilirubin and white blood cell count (47). Pattanapairoj and colleagues developed a C4.5 decision tree classification model to distinguish cholangiocarcinoma from other benign disorders, and the model showed a good diagnostic performance (48). The inclusion of test for anemia by using random forests model can help to improve the diagnostic accuracy of breast cancer (49).

Summary

The primary aim of laboratory medicine is to make an accurate diagnosis and risk stratification. The rule of thumb to judge whether a biomarker could be recommended

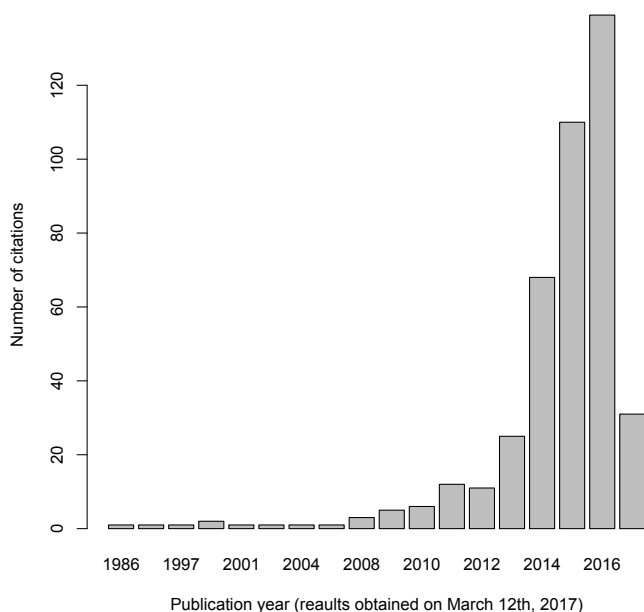


Figure 1 Exponential increase in the application of machine learning techniques in laboratory medicine.

for clinical use lies in its effectiveness in improving patient important outcomes such as mortality, morbidity, hospital length of stay and cost. While the gold standard of effectiveness relies on RCTs, big-data clinical study employing EHR provides an alternative to guide clinical practice. RCTs are limited by their strict inclusion/exclusion criteria, high cost and ethical constraint. EHR contains data on patient-level granularity and can help to disentangle complex research questions. The Achilles' heel of observational results (e.g., big-data study is a kind of observational study in nature) is uncontrolled confounding. Although it is still impossible to control for unmeasured confounding factors, the re-use of EHR helps to control as much confounding as possible. Recent decades have witnessed exponential increase in the application of machine learning techniques in laboratory medicine and these novel techniques continue to provide cutting edge tool for better prediction and stratification of diseases.

Acknowledgments

Funding: The study was funded by Zhejiang Engineering Research Center of Intelligent Medicine (2016E10011) from the First Affiliated Hospital of Wenzhou Medical University.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Tony Badrick and Tze Ping Loh) for the series “Clinical Database in Laboratory Medicine Research Column” published in *Journal of Laboratory and Precision Medicine*. The article has undergone external peer review.

Conflicts of Interest: The author has completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/jlpm.2017.06.07>). The series “Clinical Database in Laboratory Medicine Research Column” was commissioned by the editorial office without any funding or sponsorship. The author has no other conflicts of interest to declare.

Ethical Statement: The author is accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Lippi G, Mattiuzzi C. The biomarker paradigm: between diagnostic efficiency and clinical efficacy. *Pol Arch Med Wewn* 2015;125:282-8.
2. Zhang Z, Xu X, Yao M, et al. Use of the PiCCO system in critically ill patients with septic shock and acute respiratory distress syndrome: a study protocol for a randomized controlled trial. *Trials* 2013;14:32.
3. Zhang Z, Lu B, Sheng X, et al. Accuracy of stroke volume variation in predicting fluid responsiveness: a systematic review and meta-analysis. *J Anesth* 2011;25:904-16.
4. Zhang Z, Ni H, Qian Z. Effectiveness of treatment based on PiCCO parameters in critically ill patients with

- septic shock and/or acute respiratory distress syndrome: a randomized controlled trial. *Intensive Care Med* 2015;41:444-51.
5. Hohn A, Heising B, Schütte JK, et al. Procalcitonin-guided antibiotic treatment in critically ill patients. *Langenbecks Arch Surg* 2017;402:1-13.
 6. Cabral L, Afreixo V, Almeida L, et al. The Use of Procalcitonin (PCT) for Diagnosis of Sepsis in Burn Patients: A Meta-Analysis. *PLoS One* 2016;11:e0168475.
 7. Zhang Z, Smischney NJ, Zhang H, et al. AME evidence series 001-The Society for Translational Medicine: clinical practice guidelines for diagnosis and early identification of sepsis in the hospital. *J Thorac Dis*. 2016;8:2654-65.
 8. de Jong E, van Oers JA, Beishuizen A, et al. Efficacy and safety of procalcitonin guidance in reducing the duration of antibiotic treatment in critically ill patients: a randomised, controlled, open-label trial. *Lancet Infect Dis* 2016;16:819-27.
 9. Verduri A, Luppi F, D'Amico R, et al. Antibiotic treatment of severe exacerbations of chronic obstructive pulmonary disease with procalcitonin: a randomized noninferiority trial. *PLoS One* 2015;10:e0118241.
 10. Huang TS, Huang SS, Shyu YC, et al. A procalcitonin-based algorithm to guide antibiotic therapy in secondary peritonitis following emergency surgery: a prospective study with propensity score matching analysis. *PLoS One* 2014;9:e90539.
 11. Chen Y, Yao L, Du L, et al. Rationales, methods, challenges and development tendency of using GRADE in systematic reviews of diagnostic accuracy tests. *Chinese Journal of Evidence-Based Medicine* 2014;14:1402-6.
 12. Gopalakrishna G, Mustafa RA, Davenport C, et al. Applying Grading of Recommendations Assessment, Development and Evaluation (GRADE) to diagnostic tests was challenging but doable. *J Clin Epidemiol* 2014;67:760-8.
 13. Saligari E, Pagani L, Jund J, et al. Rationalising BNP prescription in the Emergency Department. *Am J Emerg Med* 2017. [Epub ahead of print].
 14. Islam MA, Bari MS, Islam MN, et al. B-type Natriuretic Peptide Assay in Differentiating Congestive Heart Failure from Lung Disease in Patients Presenting with Dyspnea. *Mymensingh Med J* 2016;25:470-6.
 15. Burri E, Hochholzer K, Arenja N, et al. B-type natriuretic peptide in the evaluation and management of dyspnoea in primary care. *J Intern Med* 2012;272:504-13.
 16. Lam LL, Cameron PA, Schneider HG, et al. Meta-analysis: effect of B-type natriuretic peptide testing on clinical outcomes in patients with acute dyspnea in the emergency setting. *Ann Intern Med* 2010;153:728-35.
 17. Schneider HG, Lam L, Lokuge A, et al. B-type natriuretic peptide testing, clinical outcomes, and health services use in emergency department patients with dyspnea: a randomized trial. *Ann Intern Med* 2009;150:365-71.
 18. Zhang Z. Big data and clinical research: focusing on the area of critical care medicine in mainland China. *Quant Imaging Med Surg* 2014;4:426-9.
 19. Zhang Z. Big data and clinical research: perspective from a clinician. *J Thorac Dis* 2014;6:1659-64.
 20. Zennaro F, Neri E, Nappi F, et al. Real-Time Tele-Mentored Low Cost "Point-of-Care US" in the Hands of Paediatricians in the Emergency Department: Diagnostic Accuracy Compared to Expert Radiologists. *PLoS One* 2016;11:e0164539.
 21. Acar Y, Tezel O, Salman N, et al. 12th WINFOCUS world congress on ultrasound in emergency and critical care. *Crit Ultrasound J* 2016;8:12.
 22. Fischer LM, Woo MY, Lee AC, et al. Emergency medicine point-of-care ultrasonography: a national needs assessment of competencies for general and expert practice. *CJEM* 2015;17:74-88.
 23. Copetti R. Is lung ultrasound the stethoscope of the new millennium? Definitely yes! *Acta Med Acad* 2016;45:80-1.
 24. Sekiguchi H. Tools of the Trade: Point-of-Care Ultrasonography as a Stethoscope. *Semin Respir Crit Care Med* 2016;37:68-87.
 25. Tricoci P, Allen JM, Kramer JM, et al. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA* 2009;301:831-41.
 26. Poonacha TK, Go RS. Level of scientific evidence underlying recommendations arising from the National Comprehensive Cancer Network clinical practice guidelines. *J Clin Oncol* 2011;29:186-91.
 27. Ely JW, Osheroff JA, Ebell MH, et al. Analysis of questions asked by family doctors regarding patient care. *BMJ* 1999;319:358-61.
 28. Cogdill KW. Information needs and information seeking in primary care: a study of nurse practitioners. *J Med Libr Assoc* 2003;91:203-15.
 29. Ely JW, Osheroff JA, Chambliss ML, et al. Answering physicians' clinical questions: obstacles and potential solutions. *J Am Med Inform Assoc* 2005;12:217-24.
 30. Gorman PN, Helfand M. Information seeking in primary care: how physicians choose which clinical questions to pursue and which to leave unanswered. *Med Decis Making* 1995;15:113-9.

31. Del Fiol G, Workman TE, Gorman PN. Clinical questions raised by clinicians at the point of care: a systematic review. *JAMA Intern Med* 2014;174:710-8.
32. Weijer C. Placebo-controlled trials in schizophrenia: are they ethical? Are they necessary? *Schizophr Res* 1999;35:211-8; discussion 227-36.
33. Colli A, Pagliaro L, Duca P. The ethical problem of randomization. *Intern Emerg Med* 2014;9:799-804.
34. Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035.
35. Cook SF, Visscher WA, Hobbs CL, et al. Project IMPACT: results from a pilot validity study of a new observational database. *Crit Care Med* 2002;30:2765-70.
36. Zhang Z. Accessing critical care big data: a step by step approach. *J Thorac Dis* 2015;7:238-42.
37. Zhang Z, Xu X, Ni H, et al. Platelet indices are novel predictors of hospital mortality in intensive care unit patients. *J Crit Care* 2014;29:885.e1-6.
38. Zhang Z, Xu X, Fan H, et al. Higher serum chloride concentrations are associated with acute kidney injury in unselected critically ill patients. *BMC Nephrol* 2013;14:235.
39. Zhang Z. Propensity score method: a non-parametric technique to reduce model dependence. *Ann Transl Med* 2017;5:7.
40. Zhang Z. Variable selection with stepwise and best subset approaches. *Ann Transl Med* 2016;4:136.
41. Zhang Z. Model building strategy for logistic regression: purposeful selection. *Ann Transl Med* 2016;4:111.
42. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol* 2007;165:710-8.
43. Zhang Z, Chen L, Ni H. The effectiveness of Corticosteroids on mortality in patients with acute respiratory distress syndrome or acute lung injury: a secondary analysis. *Sci Rep* 2015;5:17654.
44. Marco-Ruiz L, Moner D, Maldonado JA, et al. Archetype-based data warehouse environment to enable the reuse of electronic health record data. *Int J Med Inform* 2015;84:702-14.
45. Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016;316:2402-10.
46. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115-8.
47. Richardson A, Signor BM, Lidbury BA, et al. Clinical chemistry in higher dimensions: Machine-learning and enhanced prediction from routine clinical chemistry data. *Clin Biochem* 2016;49:1213-20.
48. Pattanapairoj S, Silsirivanit A, Muisuk K, et al. Improve discrimination power of serum markers for diagnosis of cholangiocarcinoma using data mining-based approach. *Clin Biochem* 2015;48:668-73.
49. Biljan M, Dmitrović B, Kristek J, et al. Statistical learning confirms the diagnostic significance of the anemia panel in breast cancer. *Clin Chem Lab Med* 2012;50:1671-8.

doi: 10.21037/jlpm.2017.06.07

Cite this article as: Zhang Z. The role of Big-data in clinical studies in laboratory medicine. *J Lab Precis Med* 2017;2:34.