

Reviewing the relationship between machines and radiology: the application of artificial intelligence

Acta Radiologica Open
10(2) 1–7
© The Foundation Acta
Radiologica 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2058460121990296
journals.sagepub.com/home/arr



Rani Ahmad 

Abstract

Background: The scope and productivity of artificial intelligence applications in health science and medicine, particularly in medical imaging, are rapidly progressing, with relatively recent developments in big data and deep learning and increasingly powerful computer algorithms. Accordingly, there are a number of opportunities and challenges for the radiological community.

Purpose: To provide review on the challenges and barriers experienced in diagnostic radiology on the basis of the key clinical applications of machine learning techniques.

Material and Methods: Studies published in 2010–2019 were selected that report on the efficacy of machine learning models. A single contingency table was selected for each study to report the highest accuracy of radiology professionals and machine learning algorithms, and a meta-analysis of studies was conducted based on contingency tables.

Results: The specificity for all the deep learning models ranged from 39% to 100%, whereas sensitivity ranged from 85% to 100%. The pooled sensitivity and specificity were 89% and 85% for the deep learning algorithms for detecting abnormalities compared to 75% and 91% for radiology experts, respectively. The pooled specificity and sensitivity for comparison between radiology professionals and deep learning algorithms were 91% and 81% for deep learning models and 85% and 73% for radiology professionals ($p < 0.000$), respectively. The pooled sensitivity detection was 82% for health-care professionals and 83% for deep learning algorithms ($p < 0.005$).

Conclusion: Radiomic information extracted through machine learning programs from images that may not be discernible through visual examination, thus may improve the prognostic and diagnostic value of data sets.

Keywords

artificial intelligence, machine learning, radiology

Received 8 October 2020; accepted 7 January 2021

Introduction

Medical imaging is one of the first branches of health science to utilize machine learning and artificial intelligence (AI) to assist human medical practice.¹ Machine learning allows computers to learn in an analogous way to humans, extracting patterns or classes based on input experience or a data set. This parallelism in learning is becoming increasingly close with the continuous innovations in data science and the progress of machine learning and AI.² With the advancement in medical imaging technology and the incorporation of large data sets, machines can extract features that are arguably beyond the reach of human perception and cognition. In recent years, a number of machine

learning algorithms have been used in content-based image retrieval systems for improving efficiency and accuracy.^{3,4}

Several computational principles can be used to categorize machine learning algorithms, such as unsupervised learning, supervised learning, and semi-supervised learning.⁴ In supervised learning, a system is

King Abdulaziz University, King Abdulaziz University Hospital, Jeddah, Saudi Arabia

Corresponding author:

Rani Ahmad, King Abdulaziz University, King Abdulaziz University Hospital, Jeddah, Saudi Arabia.
Email: Rahmad@kau.edu.sa



provided with input and output features, and the emphasis is on understanding how these are mapped to each other. In unsupervised learning, inferences are drawn from data sets comprising input data without any labeled data. Cluster analysis is the most common unsupervised learning method, which is used for identifying trends or groups in data through exploratory data analysis.⁵ It functions by grouping sets of unlabeled data into clusters of similar features without differentiating between dependent and non-dependent variables. Semi-supervised learning is a compromise between supervised and unsupervised learning techniques, utilizing some labeled data to leverage the analysis of unlabeled data. Speech analysis is the most common application of semi-supervised learning models.⁶⁻⁸

Presently, a new era of AI in radiology is emerging with, focus on analyzing images which has been showing promising results for some time. Indeed, expectations the application of AI to radiological images have increased significantly. This suggests a need to review the existing literature on the application of machine learning and AI to radiological modalities, so that their potential effect can be understood. The purpose of this study is to provide a review on the challenges and barriers experienced in diagnostic radiology on the basis of the key clinical applications of machine learning techniques. The following hypothesis will be examined in the study:

$H_1 =$ *The radiomic information extracted through machine learning programs from images that improves the prognostic and diagnostic value of data sets.*

Methodology

Design and eligibility

This review has carried out a search for studies that explore the challenges and barriers in diagnostic radiology through the context of machine learning techniques. Only studies published in 2010–2019 and in English were included. The setting was hospital-based or clinical-based, and concerned reporting the

effectiveness of machine learning models or AI algorithms on the ability to detect and interpret radiological findings. Narrative reviews, letters, preprints, and scientific reports were also included in the review. Interventions and findings related to home-based settings were excluded, as were studies on non-human or animal samples or duplicate data were excluded (Table 1). The review assumes that expert opinion or consensus opinion and standard-of-care diagnoses are accurate.

Sources and search strategy

This review has searched EMBASE, Science Citation Index, Conference Proceedings Citation Index, and Ovid-MEDLINE for studies published from 1 January 2010 to 30 December 2019, in English. The following keywords were used: Machine learning AND imaging; Machine learning AND Radiology; Deep learning AND algorithms AND imaging; and AI AND Radiology

Patients and intervention

Studies that include patients diagnosed with any type of disease detected using machine learning algorithms were selected. Prospective assessments were undertaken for identifying the effect of these algorithms upon diagnostic yield and also on therapeutic yield. This review explores the implementation of machine learning algorithms and its “downstream effects” on the clinical pathway. The Consolidated Standards of Reporting Trials and Standard Protocol Items were reviewed for prospective trials.

Data management

A manual search of citations, related articles, and bibliographies of included studies was undertaken to identify any further relevant articles that might have been missed during the automated search process. The analysis for studies providing contingency tables for both machine learning algorithm performance and

Table 1. Description of inclusion and exclusion criteria.

	Inclusion criteria	Exclusion criteria
Publication	Between 2010 to 2019	Before 2010
Setting	Hospital or clinical based	Home based
Outcome	Effectiveness of AI on radiological findings	Effectiveness of AI non-radiological medical equipment
Study type	Experimental studies, observational studies, narrative reviews, letters, preprints, and scientific reports	Blogs, Newspapers, web-based guidelines
Population	Humans	Non-human or animal samples
Language	English	Non-English

health-care professional performance was done using the sample external validation data sets.

Risk of bias

The recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement were followed throughout. Methods of analysis and inclusion criteria were specified in advance. The research question was formulated based on previously published recommendations for systematic reviews of prediction models, the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modeling Studies.

Data synthesis

A single contingency table was selected for each study to report the highest accuracy of radiology professionals and machine learning algorithms. Binary diagnostic accuracy data were extracted. Contingency tables of true-negative, false-negative, true-positive, and false-positive were used for calculating sensitivity and specificity.

Secondary outcomes

A meta-analysis of studies was based on contingency tables to estimate the accuracy of machine learning algorithms. This review has assumed contingency tables to be independent from each other, whether a study provides various contingency tables for the same or different algorithms. A unified hierarchical model was used for the meta-analysis of diagnostic accuracy studies and the plotted summary receiver operating characteristic (ROC) curves.

Results

The search identified 10,758 records, out of which 5035 were screened. The study found and evaluated 102 full text articles for eligibility; 23 studies⁹⁻³² were included in this systematic review. Sixteen studies collected retrospective data and seven studies collected data prospectively. A pre-specified sample size calculation was not reported in any of the studies. The condition that health-care professionals were provided with additional clinical information alongside the image was examined in four studies.

An algorithm-plus-clinician condition was undertaken with diagnostic performance in three studies. One study depended on single expert consensus, five studies used histopathology for confirmation, four studies used clinical follow-up, five studies used various models of expert consensus, two studies used clinical trial data; one study used surgical confirmation, and two studies used clinical care notes or labels.

The present review has also pooled performances of radiology professionals and deep learning models obtained from matched internally validated samples as an exploratory analysis. In addition, a single contingency table was selected to report each study with the highest accuracy.

The specificity for all the deep learning models ranged from 39% to 100%, whereas sensitivity ranged from 85% to 100%. The pooled sensitivity was 89% for the deep learning algorithms for detecting abnormalities compared to 75% for radiology experts, when averaging across studies using hierarchical summary ROC curves (Fig. 1). Similarly, the pooled specificity was 85% for health-care professionals and 91% for deep learning algorithms ($p = 0.000$).

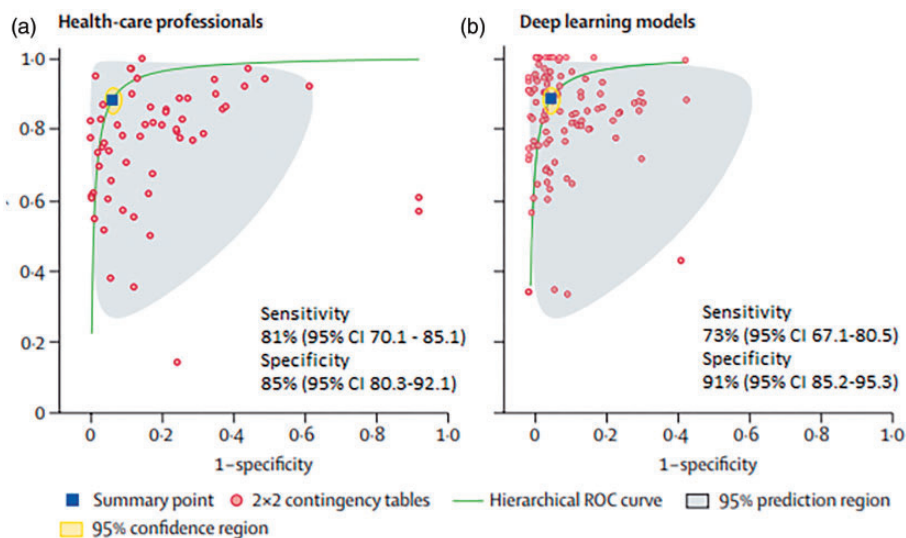


Fig. 1. ROC curves of all studies. CI: confidence interval.

A comparison between radiology professionals and deep learning algorithms was made in 12 studies. Of these 12 studies, the pooled specificity for detection of abnormalities was 91% for deep learning models and 85% for radiology professionals ($p < 0.000$). The pooled sensitivity was 73% for radiology professionals and 81% for deep learning algorithms. The pooled sensitivity detection was 82% for health-care professionals and 83% for deep learning algorithms ($p < 0.005$).

Discussion

A number of machine learning algorithms, in particular, deep neural networks, have been implemented in content-based image retrieval systems for improving query efficiency and accuracy. The processing of radiology text reports is another application of machine learning in radiology. Large text databases comprise the accumulated reports of daily radiology practice.¹⁰ Modern information processing technologies are used to exploit these radiology report databases to enhance retrieval, report search, and assist radiologists in making accurate diagnoses. Natural language understanding and natural language processing offer a more effective approach for managing and retrieving appropriate information hidden in the radiology reports.¹¹ They can extract meaningful information as well as manage large-scale data in a more efficient way that is not possible for human readers.³³ The importance of machine learning has grown in recent years owing to these attributes, being above all a practical way to carry out a text analysis of radiology report databases.

The advantage of rapid technological change in radiology, and the rapidly evolving field of radiomics, is similar to other fields that have benefited from transitioning to digital systems, while issues continue to present themselves, such as those surrounding the perception that machines and computers take jobs away from humans, often considered to be a cultural barrier in the implementation of AI in radiology.¹⁵ It has been predicted that much of the work of anatomic pathologists and radiologists will be possible through machine learning in the future, and thus, human occupations will become threatened. In addition, it is likely that machine learning techniques will become even more sophisticated in the next 5–10 years, which may threaten radiology as a thriving human discipline.¹⁸

Medical images are highly heterogeneous at both a population and an individual level, and so to train AI systems for a given application is a complex task if the number of available labeled images is restricted.²¹ In this regard, there may often be a risk of over-fitting the data, and there will be a loss of generalizability. Therefore, the practice of radiology may beneficially integrate AI

methods, rather than replace radiologists, and improve the efficiency of digital imaging methods.²²

A robust source of ground truth for each detection is needed to validate with AI programs trained on proven or known cases, whether the learning is supervised or unsupervised.¹⁶ Patient outcomes, gold standard testing results, and imaging methods can provide the source of the ground truth, but this should be comprehensively elucidated for each AI program that is established and used clinically.¹⁸ Fast computing systems are as yet not generally available in medical institutions for supplying results in a clinically relevant time frame for urgent or emergency diagnoses. However, this may not be a practical concern owing to the easy access to “cloud” computing solutions and the rapid development of graphic processing units at lower costs.¹⁹

The endeavor for generalizability of results, beyond the patient population in which the research was performed, is a broad challenge in clinical research. Data mining can be a complicated and costly process, despite the potential of big data to show the efficiency of technology with respect to patient outcomes in health care.²⁴ The lack of structured reporting can be a strong hindrance behind this technology. Therefore, approved nomenclature and standards could be created by the radiology industry that provides definition and structure, to locate the same types of data across reports, irrespective of the format that each facility utilizes. Information generated on images will not be dictated quantifiably but will be directed from the image viewing solutions in prospective events.²⁶ The improvement in accuracy and efficiency of machine learning emerges when location, anatomical findings, and measurements are developed as a result of the radiological viewing workflow.

The automatic generalizability of health-care knowledge from training data to future test data is one of the most significant contributions of machine learning. For instance, a computer can explain and make decisions on masses or microcalcifications of the human breast in a mammography Computer-Aided Detection (CAD) system.⁶ In this respect, then, the knowledge of radiologists of mammography diagnosis may be said to be transferred to the computer. For data, the input should be original, the associated problems should be anatomical structure of the patient and previous knowledge of the object of interest, and the objective should be segmentation of an object of interest in the image for medical image segmentation.¹

The extraction of useful features and their identification, and designing an adequate objective function, is the second step in machine learning. Various problems can be addressed toward the task of fitting the data to the anatomical structures of the target. Training the algorithm and finding the best parameters for the

graph cut model are the last step in the process. An improved scanning capability is produced through the trained machine learning segmentation algorithm for deep learning in radiology.³¹

Probabilistic models solve segmentation and image content analysis in radiological applications.⁵ Various processes are included therein, such as integration or marginalization of a complete probability model, independent and dependent variable identification, and probability density function, for making sure distributions meet the target variable or the objective function.⁷ Previous studies have addressed the segmentation problem of brain magnetic resonance images using a hidden Markov random field model¹¹ which is a stochastic process generated by Markov chain. Similarly, a hidden Markov random field model was used to capture the association between unidentified cluster labels and observations under spatial barriers.

Diagnostic imaging is one of the first medical disciplines that has optimally applied machine learning algorithms toward the automation of health care, while other medical fields have marked potential in this regard, particularly but not exclusively, cardiology, dermatology, gastroenterology, and pathology.¹⁰ Machine learning approaches may also further personalize health care to include a wider spectrum of data, such as genetic, laboratory, imaging, clinical, and laboratory information.

Linear regression (prediction of the dependent variable of the output by fitting a linear function to correlate the input/output pairs, which have a continuous range of values),³⁴ logistic regression (where the prediction is carried out on dependent categorical variables),³⁵ artificial neural networks (nonlinear connection of the input to the output, emulating the biological neurons found in the brain),³⁶ and decision trees (in which the entry “nodes” are labeled with features that are arranged to form multiple element “classes,” where a “leaf” of a decision “tree” can reach a finite discrete target in each pathway) are the algorithms used for supervised learning.³⁷ AI algorithms have facilitated significant progress in image-recognition tasks, specifically through the use of deep learning approaches. These methods vary, from convolutional neural networks to variational auto-encoders, and have significant application in medical imaging analysis. In radiology practice, medical images are traditionally evaluated visually by trained physicians, for detection, characterization, and monitoring of disease. However, with the application of AI in radiology modalities, identification of images can be carried out with more accuracy.

One of the major challenges of AI radiology is the lack of trust by the radiologists, when it is regarding answers related to analysis of medical images. The reason being, many radiologists perceive it as a “black box” due to doubts regarding the unclear

process which gives a conclusive answer. Scientific research and test running of the software in the hospital can help strengthen the radiologist’s trust in AI. According to an example proposed by one study,³⁸ similar cases from training databases could be depicted for rendering more information about data and providing relevant insights to the physicians. For AI radiology to survive, it is important to have the trust of its users, i.e., radiologists. Also radiologists can play a vital role in identifying targeted clinical cases for which these AI integrated tools can be implemented to test their effectiveness and sensitivity in clinical practice.³⁹ They can also play a crucial role of preserving their expertise and keeping check on the drawbacks of over-reliance on technology.³⁹

In the future, imaging data may be associated more readily with non-imaging data, such as those of electronic medical records or other large data sets. Indeed, when applied to electronic medical record data, deep learning can assist in extracting patient presentations that may link to clinical predictions and improve clinical decision support systems. Machine learning may thus play more of a significant role in the prediction of treatment response and prognosis. Initial phases toward this type of work have already begun. For example, machine learning can accurately estimate brain tumor response to various therapies. Also, machine learning can be used in the prediction of longevity of patients by detecting characteristics representative of overall individual health.

This study has several limitations. First, it might have language bias, as only the studies that were published in English were included. Another limitation was that the studies varied in their methodology; therefore, meta-analysis was not done.

In conclusion, it is necessary for medical technologies to improve value with respect to the delivery of radiology services and medical care, for reduced time on tasks for radiologists, increased diagnostic certainty, mitigated costs of care with effective findings for patients, and faster availability of findings. Significant time and experience are required to establish whether these advantages have been met in the implemented technology and to understand comparative benefits, as with any new technological innovation. If machine learning and AI programs can be developed that are tolerant of various data acquisition protocols and work in various patient populations, they will have achieved the required outcomes. Nevertheless, success will need comprehensive understanding of the conditions under which a particular program is appropriate. Yet, the ultimate role of machine learning methods in radiology is still unclear, as is the influence these will eventually have on radiologists. What is apparent, however, is that machine learning and AI offer a powerful set of

tools to analyze image data that have considerable potency. The elevated interest in AI in radiomics and radiology in recent years suggests it may have a primary role in the near future.

Acknowledgments

The author is very thankful to all the associated personnel in any reference that contributed in/for the purpose of this research.

Author Contributions

The sole author is responsible for all the elements involved for the execution of this research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Rani Ahmad  <https://orcid.org/0000-0002-3635-819X>

References

- da Costa Leite C, Júnior EA, Buchpiguel CA, et al. Perspective of radiology in the next one hundred years. *Rev Med São Paulo* 2016;29:107–109.
- Murray TE, Halligan JJ, Lee MJ. Inefficiency, dignity and patient experience: is it time for separate outpatient diagnostics? *Br J Radiol* 2017;90:20170574.
- Kohli M, Prevedello LM, Filice RW, et al. Implementing machine learning in radiology practice and research. *Am J Roentgenol* 2017;208:754–760.
- Choy G, Khalilzadeh O, Michalski M, et al. Current applications and future impact of machine learning in radiology. *Radiology* 2018;288:318–328.
- Thrall JH, Li X, Li Q, et al. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. *J Am Coll Radiol* 2018;15:504–508.
- Erickson BJ, Korfiatis P, Akkus Z, et al. Machine learning for medical imaging. *Radiographics* 2017;37:505–515.
- Giger ML. Machine learning in medical imaging. *J Am Coll Radiol* 2018;15:512–520.
- Chen PH, Zafar H, Galperin-Aizenberg M, et al. Integrating natural language processing and machine learning algorithms to categorize oncologic response in radiology reports. *J Digit Imaging* 2018;31:178–184.
- Syeda-Mahmood T. Role of big data and machine learning in diagnostic decision support in radiology. *J Am Coll Radiol* 2018;15:569–576.
- Weese J, Lorenz C. Four challenges in medical image analysis from an industrial perspective. *Med Image Anal* 2016;33:44–49.
- Chartrand G, Cheng PM, Vorontsov E, et al. Deep learning: a primer for radiologists. *Radiographics* 2017;37:2113–2131.
- Ahmad LG, Eshlaghy AT, Poorebrahimi A, et al. Using three machine learning techniques for predicting breast cancer recurrence. *J Health Med Inform* 2013;4:124–130.
- Arsanjani R, Xu Y, Dey D, et al. Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population. *J Nucl Cardiol* 2013;20:553–562.
- Abdel-Zaher AM, Eldeib AM. Breast cancer classification using deep belief networks. *Expert Syst App* 2016;46:139–144.
- Utomo CP, Kardiana A, Yuliwulandari R. Breast cancer diagnosis using artificial neural networks with extreme learning techniques. *IJARAI* 2014;3:4–10.
- Chilamkurthy S, Ghosh R, Tanamala S, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet* 2018;392:2388–2396.
- Sun YV, Bielak LF, Peyser PA, et al. Application of machine learning algorithms to predict coronary artery calcification with a sibship-based design. *Genetic Epidemiol IGES* 2008;32:350–360.
- Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 2018;172:1122–1131.
- McBee MP, Awan OA, Colucci AT, et al. Deep learning in radiology. *Acad Radiol* 2018;25:1472–1480.
- Tajmir SH, Alkasab TK. Toward augmented radiologists: changes in radiology education in the era of machine learning and artificial intelligence. *Acad Radiol* 2018;25:747–750.
- Zacharaki EI, Wang S, Chawla S, et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magn Reson Med* 2009;62:1609–1618.
- Mazurowski MA, Buda M, Saha A, et al. Deep learning in radiology: an overview of the concepts and a survey of the state of the art with focus on MRI. *J Magn Reson Imaging* 2019;49:939–954.
- Chockley K, Emanuel E. The end of radiology? Three threats to the future practice of radiology. *J Am Coll Radiol* 2016;13:1415–1420.
- Zech J, Pain M, Titano J, et al. Natural language-based machine learning models for the annotation of clinical radiology reports. *Radiology* 2018;287:570–580.
- Halabi SS, Prevedello LM, Kalpathy-Cramer J, et al. The RSNA pediatric bone age machine learning challenge. *Radiology* 2019;290:498–503.
- Nguyen DH, Patrick JD. Supervised machine learning and active learning in classification of radiology reports. *J Am Med Inform Assoc* 2014;21:893–901.
- Saba L, Biswas M, Kuppili V, et al. The present and future of deep learning in radiology. *Eur J Radiol* 2019;114:14–24.

28. Chan S, Siegel EL. Will machine learning end the viability of radiology as a thriving medical specialty? *Br J Radiol* 2019;92:20180416.
29. Suominen H, Ginter F, Pyysalo S, et al. Machine learning to automate the assignment of diagnosis codes to free-text radiology reports: a method description. In: *Proceedings of the ICML/UAI/COLT workshop on machine learning for health-care applications*, 2008.
30. Bahl M, Barzilay R, Yedidia AB, et al. High-risk breast lesions: a machine learning model to predict pathologic upgrade and reduce unnecessary surgical excision. *Radiology* 2018;286:810–818.
31. Lugo-Fagundo C, Vogelstein B, Yuille A, et al. Deep learning in radiology: now the real work begins. *J Am Coll Radiol* 2018;15:364–367.
32. Zuccon G, Waghlikar AS, Nguyen AN, et al. Automatic classification of free-text radiology reports to identify limb fractures using machine learning and the snomed CT ontology. *AMIA Summits Transl Sci Proc* 2013;2013:300.
33. Wang S, Summers RM. Machine learning and radiology. *Med Image Anal* 2012;16:933–951.
34. Witten IH, Frank E. *Data mining: practical machine learning tools and techniques with Java implementations*. SIGMOD Rec 2002;31:76–77.
35. Bishop CM. *Pattern recognition and machine learning*. 1st ed. New York: Springer, 2006.
36. Al-Shayea QK. Artificial neural networks in medical diagnosis. *IJCSI* 2011;8:150–154.
37. Lakhani P, Prater AB, Hutson RK, et al. Machine learning in radiology: applications beyond image interpretation. *J Am Coll Radiol* 2018;15:350–359.
38. Ridley EL. New algorithm overcomes imaging AI challenges, <https://www.auntminnie.com/index.aspx?sec=sup&sub=aic&pag=dis&ItemID=124063> (2018, accessed 4 March 2019).
39. Rubin DL. Artificial intelligence in imaging: the radiologist's role. *J Am Coll Radiol* 2019;16:1309–1317.