

Journal of Nephropathology



The resolution of immunofluorescent pathological images affects diagnosis for not only artificial intelligence but also human

Kensaku Takahashi^{ID}, Shinji Kitamura*^{ID}, Kazuhiko Fukushima, Yizhen Sang, Kenji Tsuji, Jun Wada

Department of Nephrology, Rheumatology, Endocrinology and Metabolism, Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Sciences, Okayama-shi, Okayama, Japan

ARTICLE INFO

Article type:
Original Article

Article history:
Received: 10 February 2021
Accepted: 9 March 2021
Published online: 2 April 2021

Keywords:
Immunofluorescent image
Renal pathology
Artificial intelligence
Deep learning
IgA nephropathy

ABSTRACT

Introduction: For human, the resolution of images is important for diagnosis. Many clinical applications of artificial intelligence have been studied, however there are few reports on the difference in diagnosis between humans and artificial intelligence on the point of the renal pathological image resolution.

Objectives: We examined whether the resolution of renal pathological images affects diagnosis of artificial intelligence and human.

Patients and Methods: From 885 renal biopsy patients, we collected renal IgA immunofluorescent pathological images that resolution is 4, 16, 32, 64, 128, 256 and 512 pixels for each patient, and divided into training data set and validation data set, and created optimum deep learning models for each resolution. To compare with artificial intelligence nephrologist also tried to diagnose by using the same validation data set images.

Results: We inputted IgA immunofluorescent pathological images into each optimum model. Human could not identify specific staining site with four pixels images, however, each resolution optimum model showed high accuracy, average over 80%. The each accuracy was observed higher depending on the resolution. The area under the curve (AUC) showed higher diagnosis ratio depending on the resolution, too. Nephrologist performed high diagnosis sensitivity depending on resolution images as same as artificial intelligence. However, nephrologists' diagnosis observed large variations in specificity depending on resolution. These results suggested that the resolution might affect specificity for human not artificial intelligence

Conclusion: The resolution of images might be important for not AI but human on the point of specificity.

Implication for health policy/practice/research/medical education:

We examined how the image resolution affects the diagnosis not only artificial intelligence but also nephrologists in this study. The differences between human and artificial intelligence is specificity on different resolution image diagnosis. The resolution of images might be important for not artificial intelligence but human on the point of specificity.

Please cite this paper as: Takahashi K, Kitamura S, Fukushima K, Sang Y, Tsuji K, Wada J. The resolution of immunofluorescent pathological images affects diagnosis for not only artificial intelligence but also human. J Nephropathol. 2021;10(3):e26. DOI: 10.34172/jnp.2021.26.

Introduction

Detection and early treatment of chronic kidney diseases (CKD) have been important because not only to prevent the deterioration of kidney function but also to prevent the progression of cardiovascular event. Among CKD, chronic nephritis syndrome is representative renal disease, which causes proteinuria and hematuria, and

renal function gradually declines. IgA nephropathy is the most common type of chronic nephritis in Asia. On IgA nephropathy, glycan insufficient IgA forms an immune complex, deposits in the para-mesangial region in glomeruli, causes inflammation, urinary protein and urine occult blood. The IgA nephropathy is diagnosed from renal biopsy specimens by the immunofluorescence

*Corresponding author: Shinji Kitamura,
Email:kitamura@okayama-u.ac.jp

staining images, light microscope images and electron microscope images. On IgA nephropathy, the diagnosis was performed only from pathological images because it is necessary to observe the deposition of IgA immune complex in glomeruli. Thus, the pathological diagnostic images are very important for definitive diagnosis.

On the other hand, various technologies have emerged in the medical world, and the medical environment is changing. Artificial intelligence has already entered the fields of medical research and clinical practice, and various applications have been made (1). Artificial intelligence is expected to deal with big data, and contribute to real world. Image diagnosis is one of the fields that artificial intelligence is often applied (2). Image analysis of artificial intelligence is divided into three categories; image classification (3), image detection (4), and image segmentation (5). AI image diagnosis from renal pathological images has been attempted. Hermsen et al reported that deep learning could assess renal histopathological images in kidney tissue on the point of image segmentation (6). Ginley et al reported that artificial intelligence could diagnose segmentation and classification of diabetic glomerulosclerosis (7). There are many reports that artificial intelligence could contribute to diagnose from medical images. For human, the resolution of medical images are important because the high resolution images give much information for human to diagnose clearly. However, there has been few reports that discuss how the image resolution affect the differences between human and artificial intelligence.

Objectives

We studied whether the different resolution of immunofluorescent images of IgA nephropathy affect to diagnosis for human and artificial intelligence.

Patients and Methods

Patients information

This study is a retrospective observational study. As for informed consent, the contents of the research are posted on our department homepage and in the hospital, and public informed consent is given. Kidney samples consisted of needle biopsy samples from patients that hospitalized to undergo the examination of renal biopsy in Okayama university hospital from January 2008 to May 2018. The number of renal biopsy patients is 885 from 2012-2018 in Okayama university hospital. During 2012-2018, the number of IgA nephropathy renal biopsy patients is 162 in Okayama university hospital. For comparison, 723 non-IgA nephropathy patients' images were collected during 2012-2018. We excluded 463 patients because of renal transplantation episode biopsy patients, complicated multiple nephropathy and uncertain diagnosis. After that the exclusion, we enrolled

260 non-IgA nephropathy patients for comparison. The representative validation images from the patients with IgA nephropathy were shown in Figure S8 (Supplementary file 1), and representative validation images from the patients with non-IgA nephropathy patients were shown in supplemental figures S9. Almost 40% of renal biopsy patients are IgA nephropathy patients (Table S1, Supplementary file 1). The diagnosis of the renal disease was decided by the nephrologists with the discussion of conference in the light of medical history, physical information, laboratory data and pathological findings including of immunofluorescence, light microscopy and electron microscopy images.

Renal immunofluorescent images

The kidney samples were obtained by renal biopsy from the patients administrated in Okayama University hospital. We prepared frozen tissues from renal biopsy samples, and cut at 4 µm in a cryostat. We stained the frozen sections by fluorescein isothiocyanate (FITC)-conjugated antibodies in moist chamber for one hour. FITC-conjugated goat anti-human IgA was purchased from MP biomedicals and FITC-conjugated goat anti-human IgG from medical and biological laboratories Co., LTD. The images were obtained by fluorescence microscopes (Olympus, Japan).

Data preprocessing

The images data were obtained as JPEG or Tiff file. We changed the resolution of the files from 2776 × 2074 pixel to 4 × 4, 16 × 16, 32 × 32, 64 × 64, 128 × 128, 256 × 256 and 512 × 512 pixel. After the conversion, we converted the JPEG file to PNG file to analyze.

Deep learning

The programming language (python), the environment (Visual studio code, Microsoft) and software, neural network console (Sony Inc.) were used. This neural network console application software contains two main features. One is easy editing by dragging and dropping. This could design a neural network with layers of each parameter. Another feature is that it has an automatic structure search function. This is the ability to automatically find higher performance and lighter neural network structures. Therefore, complicated functions and tuning work are conducted automatically with neural network console software. We described Neural Network Console previously (8). Input is IgA immunofluorescent images converted previously described. And the renal pathological image is classified into training images and validation images at the ratio of 8: 2. The validation images showed (Figures S1 and S2). We performed supervised training for deep learning with convolutional neural network (CNN).

Ethical issues

The research followed the tenets of the Declaration of Helsinki. This study was approved by the committee of Okayama university (approved#ken1908-008). Accordingly, written informed consent was taken from all participants before any intervention. This study is a retrospective observational study, not clinical trial. As for informed consent, the contents of the research are posted on our department homepage and in the hospital. Public informed consent is given. As this retrospective observation study, the committee approved public informed consent.

Statistical analysis

Statistical analysis was performed by JMP (SAS Institute Inc. version 11.0.0 for Windows software). Statistical significance was defined by one-way analysis of variance (ANOVA) with student *t* test. Data are shown as the mean \pm SE. Significance was defined as $P < 0.05$. In addition, to determine the cut-off value of the diagnosis ratio, a receiver operating characteristic (ROC) curve was constructed using statistical analysis software JMP.

Results

Overview of schema of deep learning

We collected IgA nephropathy images from 162 patients and non-IgA nephropathy images from 260 non-IgA nephropathy patients (Figure 1). The diagnosis of the patients showed in Table S1 (Supplementary file 1). An overview of the computational schema is showed (Figure 2). We inputted IgA immunofluorescent images on seven kinds of resolution types, 4×4 , 16×16 , 32×32 , 64×64 , 128×128 , 256×256 and 512×512 pixels. The construction of model is that middle layer is composed by three layers, the CNN layers is composed by convolutional, pooling and function layers. Each image analyzed, and connected to output. The renal pathological IgA images are classified into training images and validation images at the ratio of 8: 2. We used the software, neural network console provided from Sony Inc. This software is to automatically add or deletes some layers and adjust parameters to get optimum result. Using this software, we performed supervised training for deep learning. This software structured optimum model automatically.

AI could diagnose IgA nephropathy from IgA renal immunofluorescent pathological images

We created some models for each resolution images, from 4×4 pixels images to 512×512 pixels images. The representative images are shown (Figure 3 a-g). The ROC curve for each resolution image are shown (Figure 3 h-n). In the regard of 4×4 pixels images (Figure 3a), although it is a fairly blurry images, artificial intelligence could diagnose that the AUC of created superior models

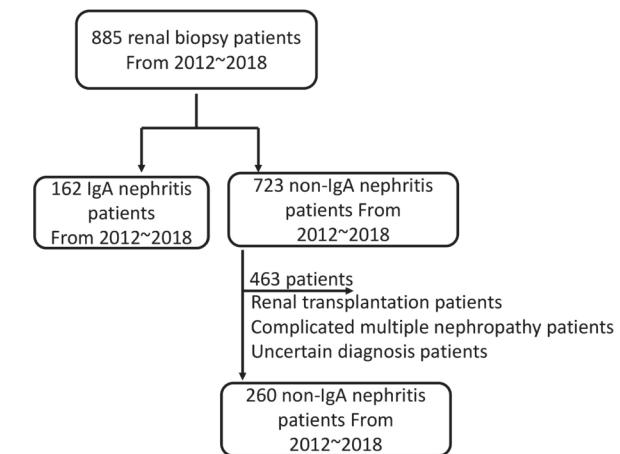


Figure 1. Enrolment. All patients underwent renal biopsy in Okayama University Hospital.

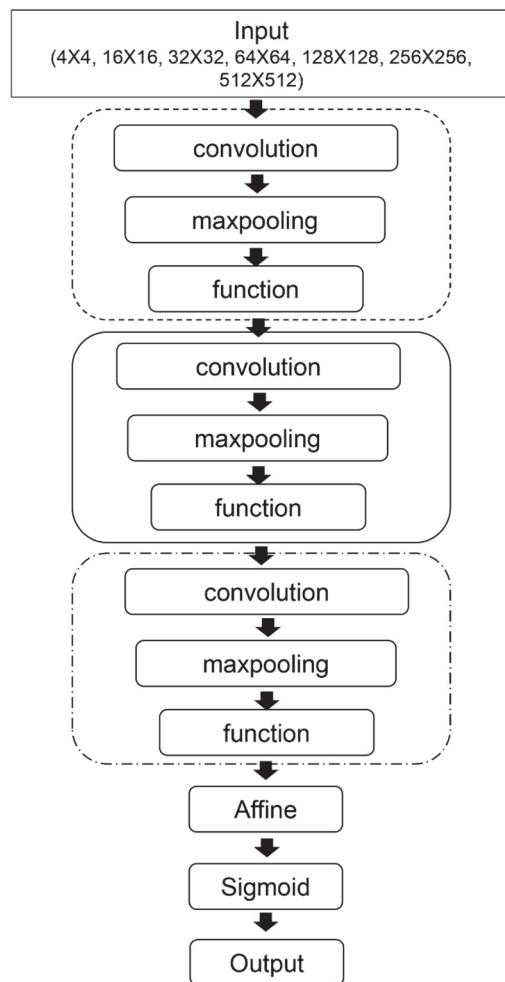


Figure 2. The overview of CNN model. We used input data IgA renal immunofluorescent images with the resolution 4×4 , 16×16 , 32×32 , 64×64 , 128×128 , 256×256 , 512×512 pixel. The basic structure model is a three-layer structure. Middle layer consists of convolution, pooling and function layer.

was 0.91111 (Figure 3h; ROC curve). Representative model showed in Figure S3. Next, in the regard of 16×16 pixels images (Figure 3a), the AUC (*area under the curve*) of created superior models was 0.93255 (Figure 3i; ROC curve). Representative model showed in Figure S4. Next, in the regard of 32×32 pixels images (Figure 3c), the AUC of created superior models was 0.93242 (Figure 3j; ROC curve). Representative model showed in Figure S5. Next, in the regard of 64×64 pixels images (Figure 3d), the AUC of created superior models was 0.94868 (Figure 3k; ROC curve). Representative model showed in Figure S6. Next, in the regard of 128×128 pixels images (Figure 3e), the AUC of created superior models was 0.94281 (Figure 3l; ROC curve). Representative model showed in Figure S7. Next, in the regard of 256×256 pixels images (Figure 3f), the AUC of created superior models was 0.97581 (Figure 3m; ROC curve). Representative model showed in Figure S8. Next, in the regard of 512×512 pixels images (Figure 3g), the AUC of created superior models was 0.98904 (Figure 3n; ROC curve). Representative model showed in Figure S9. These data indicated that artificial intelligence is automatically extracted as features from limited image information, and the resolution of images affects the diagnosis. Deep learning could diagnose IgA nephropathy higher with high resolution images. However, there are so many times to analyze the images with high resolution images.

The diagnostic comparison between nephrologist and artificial intelligence

To compare the differences between human and artificial intelligence, we tested diagnosis by nephrologist using the same validation data set. With 32 pixels images, nephrologists could diagnose the IgA nephropathy, however, the diagnosis ratio is a little inferior compared to artificial intelligence (Figure 4a). On the other hand, with 256 pixels images, nephrologists could diagnose the IgA nephropathy, however the diagnosis ratio is a little inferior compared to artificial intelligence, too (Figure 4b). The nephrologist usually comprehensively diagnose IgA nephropathy with the information of clinical course, laboratory data, pathological images and so on. As same as artificial intelligence, nephrologist diagnosed higher sensitivity depending on resolution (Figure 4c). However, we could observe the tendency for specificity to increase as the resolution increased (Figure 4d). The differences in expert experience may lead to differences in specificity even if the low resolution images are hard for human eye to identify specific staining site. These results suggested that high resolution images are easier to identify and distinguish for human in image diagnosis, especially affected the specificity for human.

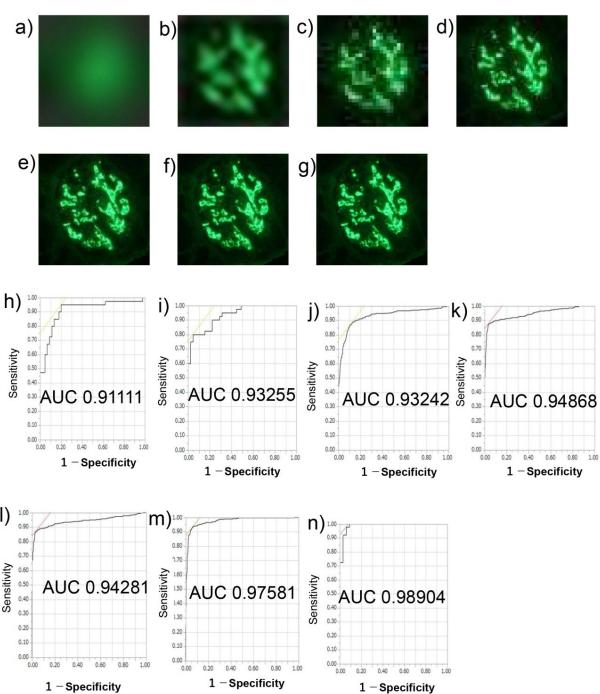


Figure 3. The AUC curve for each resolution images. (a-g) different resolution images from same IgA immunofluorescent image, (a) 4×4 pixels IgA immunofluorescent image, (b) 16×16 pixels IgA immunofluorescent image, (c) 32×32 pixels IgA immunofluorescent image, (d) 64×64 pixels IgA immunofluorescent image, (e) 128×128 pixels IgA immunofluorescent image, (f) 256×256 pixels IgA immunofluorescent image, (g) 512×512 pixels IgA immunofluorescent image, (h-n) ROC curve for each resolution image, (h) ROC curve for 4×4 pixels IgA immunofluorescent image, (i) ROC curve for 16×16 pixels IgA immunofluorescent image, (j) ROC curve for 32×32 pixels IgA immunofluorescent image, (k) ROC curve for 64×64 pixels IgA immunofluorescent image, (l) ROC curve for 128×128 pixels IgA immunofluorescent image, (m) ROC curve for 256×256 pixels IgA immunofluorescent image, (n) ROC curve for 512×512 pixels IgA immunofluorescent image.

Discussion

Artificial intelligence is reported to predict the occurrence of acute kidney injury (9), and assume the glomerular filtration rate from CT image (10). On the pathology, artificial intelligence could diagnose the metastases superior to human eye (11). As same as metastases pathology, diagnosis of IgA nephropathy is made by renal pathological, especially immunofluorescent images. This is because IgA nephropathy will be diagnosed when coarse granular deposits are found in the mesangial region in glomeruli by immunostaining of IgA and C3 from the sample of the renal biopsy. We have studied how artificial intelligence can intervene in such pathological diagnosis. In this study, we examined the difference in diagnosis due to the difference in image resolution.

Artificial intelligence can automatically extract features from images of any resolution that human is difficult to distinguish. On the image of CT and MRI, it is reported to be better to diagnose for AI with the high-resolution

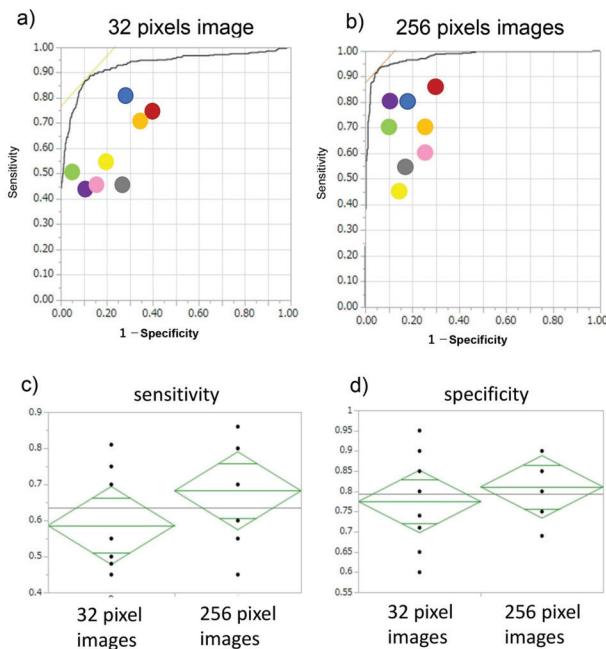


Figure 4. The difference of the diagnosis between AI and nephrologists. (a) ROC curve for 32 pixel image model with 8 nephrologists diagnosis, (b) ROC curve for 256 pixel model with 8 nephrologists. Each different color circle points out each nephrologist diagnosis. (c) Nephrologist sensitivity; diagnosis sensitivity between 32 pixel images and 256 pixel images, (d) Nephrologist specificity; diagnosis specificity between 32 pixel images and 256 pixel images.

images (12,13). It is reported that the model trained with higher resolution showed higher accuracy (14). On the other hand, it is indicated that accuracy was not affected by the difference of resolution (15). Other report applied the resolution of 32×32 pixels and obtained an AUC of 0.86 on poly detection (16). However, the resolution of more than 224×224 pixels is believed to generally use because it is necessary for enough information. Our data suggested that high resolution images show high accuracy and sensitivity not only for human but also artificial intelligence.

In image diagnosis of medicine, it is easier for human eyes to make a diagnosis with a high-resolution image. In this study, the 4×4 pixels images could not lead to diagnose on human eyes. However, artificial intelligence could diagnose the features from the low-resolution images. Artificial intelligence converts each image into a digital signal, calculates them with a function, and creates a calculation formula that matches the correct answer. Therefore, the diagnosis in that process is also called a black box problem. Therefore, it was assumed that the difference of diagnosis depending on the resolution could be different from nephrologists, but the result in this study shows higher resolution of images leads to higher diagnostic accuracy. These results suggest that

high-resolution images are also desirable for artificial intelligence because the high-resolution images provide more information. In this study, we input images of different resolutions to the model optimum for each resolution. The model optimum for high resolution could make a diagnosis well with low-resolution images, however the model optimum for low resolution could not with high-resolution images. It is similar to the fact that an adult could deal with a low degree of knowledge well, on the other hand, a child couldn't with a high degree of knowledge.

In addition, nephrologists' diagnosis observed large variations in specificity depending on resolution. For experienced nephrologists, even if the low resolution images were difficult to observe, they will expect to complement where difficult-to-observe part according to experience and lead to high diagnosis, but for inexperienced nephrologist, it was difficult to complement and advance diagnosis might be made only when the resolution is high. The difference is reflected in the specificity, further examination is needed to resolve the differences.

Conclusion

The high resolution of images provided more information to human and artificial intelligence, and affected the diagnosis. In this study, the high resolution of images affected for artificial intelligence to diagnosis, and affected for human large variations in specificity depending on resolution. The differences might be from the differences in the analysis and processing system between human and artificial intelligence. Taken together, the resolution of images might be important to diagnose for not only for human but also artificial intelligence.

Limitations of the study

In this study limitation, diagnosis of artificial intelligence was performed from only immunofluorescence images. In addition, the images were adjusted for easy observing. Moreover, human doctors diagnose comprehensive judgments with clinical course, clinical findings and clinical images, such as light microscope images and electron microscope images. To compare the differences between human and artificial intelligence, further examination is needed.

Authors' contribution

SK is the principal investigator of this study. SK designed, conceptualized the computational detections, defined the immunofluorescent images computational features, designed network architecture, analyzed the results, receiving funding and wrote the manuscript. KTA managed and collected the images and clinical data,

and wrote the manuscript. KF, YS and KTU assisted to collect the images and clinical data. JW conceived the overall research scheme, coordinated with the study team. All authors participated in preparing the final draft of the manuscript and critically evaluated the intellectual contents. All authors have read and approved the content of the manuscript and confirmed the accuracy or integrity of any part of the work.

Conflicts of interest

The authors report no conflicts of interest in this work.

Ethical considerations

Ethical issues (including plagiarism, data fabrication, double publication) have been completely observed by the authors.

Funding/Support

This research was supported by Yukiko Ishibashi Memorial foundation, Wesco Scientific Promotion Foundation and joint research with Kioxia Corporation, Japan.

Supplementary Materials

Online Supplementary file 1 contains Table S1 and Figures S1-S9.

References

1. Yang C, Kong G, Wang L, Zhang L, Zhao M. Big data in nephrology: Are we ready for the change? *Nephrology*. 2019;24:1097-102. doi: 10.1111/nep.13636.
2. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical image analysis using convolutional neural networks: a review. *J Med Syst*. 2018;42:226. doi: 10.1007/s10916-018-1088-1.
3. Pang S, Du A, Orgun MA, Yu Z. A novel fused convolutional neural network for biomedical image classification. *Med Biol Eng Comput*. 2019;57:107-21. doi: 10.1007/s11517-018-1819-y.
4. Sheehan S, Mawe S, Cianciolo RE, Korstanje R, Mahoney M. Detection and classification of novel renal histologic phenotypes using deep neural networks. *Am J Pathol*. 2019;189:1786-1796. doi: 10.1016/j.ajpath.2019.05.019.
5. Wang S, Yang DM, Rong R, Zhan X. Pathology Image analysis using segmentation deep learning algorithms. *Am J Pathol*. 2019;189:1686-98. doi: 10.1016/j.ajpath.2019.05.007.
6. Hermsen M, de Bel T, den Boer M, Steenbergen EJ, Kers J, Florquin S, et al. Deep Learning-based histopathologic assessment of kidney tissue. *J Am Soc Nephrol*. 2019;30:1968-79. doi: 10.1681/ASN.2019020144.
7. Ginley B, Lutnick B, Jen KY, Fogo AB, Jain S Rosenberg A, et al. Computational segmentation and classification of diabetic glomerulosclerosis. *J Am Soc Nephrol*. 2019;30:1953-67. doi: 10.1681/ASN.2018121259.
8. Kitamura S, Takahashi K, Sang Y, Fukushima K, Tsuji K, Wada J. Deep learning could diagnose diabetic nephropathy with renal pathological immunofluorescent images. *Diagnostics*. 2020;10:466. doi: 10.3390/diagnostics10070466.
9. Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A Clinically Applicable Approach to Continuous Prediction of Future Acute Kidney Injury. *Nature*. 2019;572:116-9. doi: 10.1038/s41586-019-1390-1.
10. Park J, Bae S, Seo S, Park S, Bang JI, Han JH, et al. Measurement of Glomerular Filtration Rate Using Quantitative SPECT/CT and Deep-learning-based Kidney Segmentation. *Sci Rep*. 2019;9:4223. doi: 10.1038/s41598-019-40710-7.
11. Benjnordi BE, Veta M, Diest PJ, Ginneken BV, Karssemeijer N, Litjens G, et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA*. 2017;318:2199-2210. doi: 10.1001/jama.2017.14585.
12. Chaudhari AS, Fang Z, Kogan F, Wood J, Stevens KJ, Gibbons EK, et al. Super-resolution Musculoskeletal MRI Using Deep Learning. *Magn Reson Med*. 2018;80:2139-2154. doi: 10.1002/mrm.27178.
13. Akagi M, Nakamura Y, Higaki T, Narita K, Honda Y, Zhou J, et al. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *European Radiology*. 2019;29:6163-6171. doi: 10.1007/s00330-019-06170-3.
14. Sahlsten J, Jaskari J, Kivinen J, Turunen L, Jaanio E, Hietala K, et al. Deep learning fundus image analysis for diabetic retinopathy and macular edema grading. *Sci Rep*. 2019;9:10750. doi: 10.1038/s41598-019-47181-w.
15. Urban G, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology*. 2018;155:1069-1078.e8. doi: 10.1053/j.gastro.2018.06.037.
16. Park SY, Sargent D. Colonoscopic polyp detection using convolutional neural networks. *SPIE Medical Imaging*. 2016;9785:978528. doi: 10.1117/12.2217148.

Copyright © 2021 The Author(s); Published by Society of Diabetic Nephropathy Prevention. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.