

一般口演

一般口演20

医療データ分析8（生体情報）

2018年11月25日(日) 09:00 ~ 10:30 G会場 (5F 504+505)

[4-G-1-5] AIによる画像診断に向けた、腎病理糸球体画像の所見一致度の評価

○山口 亮平¹, 堂本 裕加子², 宇於崎 宏³, 清水 章⁴, 長田 道夫⁵, 嶋本 公德⁶, 河添 悦昌^{1,6}, 大江 和彦^{1,6} (1.東京大学大学院 医学系研究科 医療情報学分野, 2.東京大学大学院 医学系研究科 人体病理学・病理診断学分野, 3.帝京大学医学部 病理学講座, 4.日本医科大学 解析人体病理学, 5.筑波大学 腎・血管病理学, 6.東京大学医学部附属病院 企画情報運営部)

<背景> 近年、医療画像の診断において Deep Learningをはじめとする人工知能関連技術が導入されてきているが、その学習には質の高い正解ラベル付きデータが必要である。ラベル付けを行うためには正解となる基準が必要であるが、腎生検の糸球体病理画像には、そのような基準はない。今回我々は、腎生検の糸球体病理画像における所見項目について、定義を明確にした上で医師が所見付けを行い、その一致度を評価した。<方法> PAS染色でつけるべき12種類の糸球体所見項目ならびにその定義を作成した。IgA腎症に限定した所見定義であるOxford分類を参考に、一部所見項目の追加と定義変更を行った。東大病院の腎病理デジタルスライド画像から糸球体画像を抽出し、その中からランダムに選択した100枚の画像を対象とし、5人の専門医が所見付けを行った。所見付けはお互いの所見が見えない状態で行った。所見付け終了後、5人のうちのすべての2人の組み合わせにおける所見一致度を κ 係数として算出し、各所見項目における平均値を評価に用いた。<結果> 各所見項目において κ 係数の平均値は0.16から0.76まで大きくばらついていた。もともとOxford分類に採用されている所見項目は高い傾向にあり、sclerosisやcellular crescentなどの所見項目は κ 係数の平均値は0.4(中等度の一致)を超えていたが、他のmesangiolysisやadhesionなどの所見項目の平均値は低かった。いくつかの所見項目においては、質の高い正解ラベルつきデータを得るには不十分な κ 係数であり、所見項目の定義に改善の余地があると考えられた。今後、各所見項目において高い一致率が得られるよう検討が必要である。

AIによる画像診断に向けた、腎病理系球体画像の所見一致度の評価

山口亮平¹、河添 悦昌^{1,2}、嶋本公德²、堂本裕加子³、宇於崎宏⁴、清水章⁵、長田道夫⁶、大江和彦^{1,2}

1. 東京大学大学院医学系研究科 医療情報学分野、2. 東京大学医学部附属病院 企画情報運営部、

3. 東京大学大学院医学系研究科人体病理学・病理診断学分野、4. 帝京大学医学部病理学講座、

5. 日本医科大学 解析人体病理学、6. 筑波大学 腎・血管病理学

Evaluation of concordance of findings descriptors of glomerular images in renal biopsy for medical imaging diagnostic system by artificial intelligence

Ryohei Yamaguchi¹, Yoshimasa Kawazoe^{1,2}, Kiminori Shimamoto², Yukako Shintani-Domoto³,

Hiroshi Uozaki⁴, Akira Shimizu⁵, Michio Nagata⁶, Kazuhiko Ohe^{1,2}

*1 Department of Biomedical Informatics, Graduate School of Medicine, The University of Tokyo,

*2 Department of Healthcare Information Management, The University of Tokyo Hospital,

*3 Department of Pathology, Graduate School of Medicine, The University of Tokyo,

*4 Department of Pathology, Teikyo University School of Medicine,

*5 Department of Pathology, Nihon University School of Medicine,

*6 Kidney and Vascular Pathology, Faculty of Medicine, University of Tsukuba

<Background> In recent years, artificial intelligence-related technologies, such as Deep Learning, have been introduced in medical image diagnosis. While qualifying criteria must be used to label Deep Learning data, there are no criteria for the glomerular pathology images of renal biopsies. We developed qualifying criteria for descriptors of glomerulus. We then scored them based on those criteria and evaluated the agreement of the scores. <Method> We used Periodic acid-Schiff (PAS) staining to choose 12 glomerular descriptors and then we defined them. We changed some of the descriptors based on Oxford classification, which is a criterion only used for IgA nephropathy. The glomerular images were extracted from the renal pathology digital slides of The University of Tokyo Hospital, and 100 images were hand-selected. Five doctors scored 12 descriptors for each image without knowing one another's scores. We then calculated the inter-doctor agreement of the scores using the kappa coefficient for all combinations of two out of five doctors. <Result> The average kappa coefficient value ranged from 0.16 to 0.76 according to the descriptors. The descriptors, such as sclerosis or cellular crescent, had high kappa scores. In the future, the descriptor criteria should be improved to ensure better agreement for inter-doctor scoring.

Keywords: Artificial Intelligence, Renal Biopsy, Concordance, Glomerulus

1. 緒論

腎糸球体疾患は採血検査や尿検査、超音波検査などの非侵襲的な検査のみでは診断に至らない事が多く、腎生検を施行してはじめて正確な診断や治療方針を決定できることが多い。実際、腎生検を施行することで 34%の患者で治療方針が変更になったという報告¹⁾があるように、腎生検は臨床的に重要な検査である。生検腎を用いた腎病理診断は、腎臓内科医および病理医が、病理スライドに含まれる微小な糸球体を顕微鏡で一つ一つ目視により観察し、糸球体の形態的な特徴から所見付けを行うものである。このように、極めて労力を要する作業であるものの、それをコンピュータによってサポートする診断支援システムは存在しないため、その登場が期待される。

近年、医療画像の診断において Deep Learning をはじめとする人工知能関連技術が導入されてきている²⁾。腎病理診断においても、例えば Deep Learning を用いて全ての糸球体に所見を自動的に付与するシステムを構築し医師の見落としを防ぐ、といった有用性が考えられる。ここで、Deep Learning の学習には質の高い正解ラベル付きデータセットが必要になるが、そのようなラベル付けを行うためには正解を規定する基準が重要となる。病理診断におけるこの基準の要件としては、

高い再現性を持って所見付けを行うことができることに加え、病理スライドを評価する時点では疾患が定まっているとは限らないため、疾患に依存しない所見項目のセットを用意する必要があることが挙げられる。腎病理診断の分野では、IgA 腎症に対する Oxford 分類³⁾や移植腎に対する Banff 分類⁴⁾、巣状糸球体硬化症に対する Columbia 分類⁵⁾といった疾患を限定した所見付け基準が存在する。これらの既存の分類基準は、予後に関係する、もしくは鑑別診断が変わるなど臨床的に重要であり、かつ所見付けをした際に、異なる医師間で付与する所見の一致度が高いという二つの要件を満たしている。しかしながら、疾患を限定しない場合に高い所見の一致度を保証するものではない。一方で、単一の疾患に限定しない所見付け基準として NEPTUNE (The multicenter Nephrotic Syndrome Study Network) が提案する NEPTUNE descriptor⁶⁾があり、その所見一致度の評価は行われているものの、その対象疾患は「巣状糸球体硬化症」「微小変化群」「膜性腎症」に限られているため、この分類基準においても疾患を限定しない場合に同様の一致度が得られる保証はない。

2. 目的

Deep Learning による腎糸球体疾患の自動所見付けの実現に

向けて、疾患を限定しない糸球体病理画像に対して、医師間で所見スコアリングの一致度が高くなるような所見項目とそのスコアリングの基準のセット(以下、所見定義と呼ぶ)を開発する。開発した所見定義をもとに所見付けを行い、医師間の所見スコアリングの一致度(以下、所見一致度と呼ぶ)を評価する。本研究の所見付けタスクは、1枚の糸球体画像に対して、複数の所見項目のスコアリングを全て付与するものである。

3. 方法

東京大学医学部付属病院(東大病院)の所有する腎病理デジタルスライドから糸球体画像を100枚選択し、それらを用いて所見項目及びそのスコアリング基準をmodified Delphi法によって作成する。医師間での所見一致度をCohenのkappa係数を用いて算出し、所見定義の妥当性を検討する。以下、方法の詳細について述べる。

3.1 画像選択

東大病院において2010年から2017年までの期間に得られた腎病理デジタルスライドから、PAS染色のスライド200枚を選択した。各スライドから、主となる糸球体が一個含まれるように長方形によって糸球体領域を抽出し、png形式で一枚の画像として保存し、8058枚の糸球体画像を得た。それら画像から、目視によって100枚を選択し、本研究における所見付けの対象とした。100枚の画像には、典型的な例だけでなく、判断が難しく所見が分かれる可能性のある画像も含めた。

3.2 所見付けへの参加医師

5人の医師が所見付けに参加した。いずれも本論文の著者及び共著者であり、うち1人が腎臓内科専門(医師A)、2人が病理専門(医師B、C)、2人が腎病理専門(医師D、E)である。

3.3 所見項目とスコアリング基準の作成

PAS染色で判断可能であり、かつ臨床的に重要と思われる糸球体所見の項目をOxford分類や病理学の成書をもとに選択し、それらのスコアリング基準の原案を筆頭著者が作成した。5人の参加医師で原案を吟味した上で一時的な所見定義案を作成し、予備実験としてその定義に従い100枚の糸球体画像に所見付けを行った。実際に所見付けを行う上での問題点を議論した後、所見定義を推敲し、今回の実験で用いる最終所見定義とした。

3.4 所見付け及びその所見一致度の評価

開発した所見定義に従って5人の参加医師が所見付けを行なった。所見付けには、筆頭著者が所属する研究室で開発された専用のウェブベースのソフトウェア(図1)を用いて、他の医師の所見スコアリングを見ることができない環境で行なった。所見をつける上で不適切であると参加医師が判断したものを除くすべての画像に対して、全ての所見項目にスコアを必ず付与することとした。

参加医師による所見付け終了後、全員のスコアを集計した。その際、誰か一人でも不適切と判断した画像を解析から除外した。残りの画像に対して、所見項目ごとに、5人のうちのすべての2人の組み合わせにおける所見一致度をCohenのkappa係数を用いて算出し、その平均値を各所見項目の評価値とした。Kappa係数は-1から1までの値を取りうるものであり、1は完全一致、-1は完全に不一致、0は偶然の一致と変わらない事を表す。一般的にkappa係数が0~0.4まではpoor、0.4~0.6はmoderateとされている。腎病理の分野では

Banff分類に見られるようにkappa係数が0.4を超えているものが再現性のある所見として採用される事が多く、本研究でもkappa係数の平均値が0.4以上のものを一致度の高い所見項目と判断した。

4. 結果

4.1 定義された所見項目とスコアリング基準

本研究で開発した所見定義を表1に示す。最終的に選択された所見項目は12項目であり、このうち、陽性陰性の2値で表現されるものが9項目、3値以上で表現されるものが3項目であった。さらに、画像としては適切であるが所見によってはスコアを判断できないケースがあるため、全ての所見項目に対して「判別不能」をとりうる値として追加した。

本研究で開発した所見定義のセットには、Oxford分類に採用される所見項目を12項目中9項目と多く含むものであった。しかしながら、本研究はIgA腎症以外の疾患も対象とするため、Oxford分類の所見定義のままでは、所見を正しく評価できないような事例がみられた。例えば、図2に示す糸球体画像において、楕円で囲まれる領域は本来メサンギウム基質であったが、何らかの疾患によって別の物質に置き換わってしまっている。このような領域で、メサンギウム基質が増加しているという所見である‘Increased mesangial matrix’を評価しようとした場合、Oxford分類の‘Increased mesangial matrix’そのままのスコアリング基準では所見を正しく評価することができない。というのも、Oxford分類が対象とするIgA腎症では、このような基質代替性により出現する領域が想定されないからである。よって、今回のスコアリング定義においては「基質代替性領域は評価対象外である。」という条件を追加した。このように、Oxford分類から引用した所見も、一部定義の追加、変更を行なった(表1における赤字の部分が、追加および変更点を示す)。

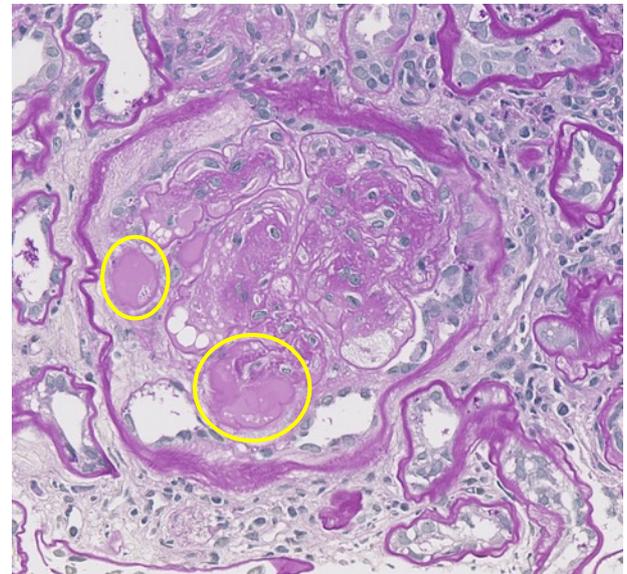


図2 Oxford分類そのままでは評価できない項目が含まれる糸球体画像の例

本来のメサンギウム基質は別の物質に置き換わっており(楕円)、本来のメサンギウム基質が増加するという所見を表す‘Increased mesangial matrix’は従来のままでは評価できない。

4.2 所見一致度

各所見項目ごとに、「判別不能」を含むものと含まないものとでそれぞれ一致度を算出した。表 2 に判別不能を含む kappa 係数の分布を示す。各所見項目において kappa 係数の平均値は 0.21 から 0.52 まで異なっていた。Kappa 係数の平均値が 0.4 を超えていたもの、すなわち一致度が高いと判断されたものは、「Sclerosis」、「Cellular Crescent」、「Collapse/Ischemic」の 3 項目であった。

次に、表 3 に所見ごとに誰か一人でも「判別不能」とした画像を除外して算出した kappa 係数の分布を示す。判別不能画像を除いた場合、kappa 係数の平均値が 0.4 を超えていたものは、「Increased mesangial matrix」、「Sclerosis」、「Polar Vasculosis」、「Afferent/Efferent Arteriolar Hyalinosis」、「Cellular Crescent」の 5 つであった。「Afferent/Efferent Hyalinosis」及び「Polar Vasculosis」の所見は、判別不能を除くことで kappa 係数は上昇する傾向にあった。逆に、「Collapse/Ischemic」の所見は、判別不能を除くことで kappa 係数は低下する傾向にあった。

5 考察

本研究は Oxford 分類などを元にして、任意の糸球体疾患における糸球体を評価するための所見定義を開発した。以下、開発した所見定義を用いて行った、5 人の医師間の所見付けスコアリングの一致度に関する考察を述べる。

5.1 Kappa 係数が高い所見項目と低い所見項目

判別不能画像を除く除かないに関わらず平均 kappa 係数が 0.4 を超えたものとして、「Sclerosis」及び「Cellular Crescent」の二つの所見があった。この二つの所見の一致度が高い理由として、もともと Oxford 分類に規定されているという他に、腎機能的な予後に直結しやすい所見であるため、腎病理に関わる多くの医師が日常臨床において特に重要視していることから、今回の実験でも参加医師間で高い一致度を示したものと考えられた。一方、「Mesangial Hypercellularity」と「Increased mesangial matrix」の二つの所見項目における平均 kappa 係数は 0.33~0.38 であり、一致度が高いとは判断されなかった。315 個の糸球体に関して kappa 係数を評価した NEPTUNE study⁴⁾における「Mesangial Hypercellularity」の平均 kappa 係数は 0.54~0.64 であり、本研究の実験結果よりも高い値を示すものであった。我々の「Mesangial Hypercellularity」は 4 値に分類するものであり、NEPTUNE においては 2 値に分類するものであるため単純に比較することはできないが、この違いの原因として、NEPTUNE は 3 種類の疾患に限定している一方、本研究は、疾患を限定しない糸球体を評価対象としたことから、その多様性が kappa 係数の低下に繋がったと考えられた。

5.2 判別不能の有無による影響

「判別不能」画像を入れた場合と、所見ごとに「判別不能」を除いた場合の差について述べる。「判別不能」を省くことで、「Afferent/Efferent Arteriolar Hyalinosis」および「Polar Vasculosis」は kappa 係数が上昇する傾向にあり、特に「Afferent/Efferent Arteriolar Hyalinosis」で顕著であった。これらの所見は、画像内に「輸入・輸出細動脈」および「血管極」という構造物が見られた場合のみ陽性か陰性か判断できるものである。判別不能を除外した場合に kappa 係数が上昇するような所見項目は、単に分類クラスの数が一つ減るといった影響の他に、構造物が画像内に確認できさえすれば、所見スコアリング自体はばらつきにくい性質を持つものと考えられた。一

方、「判別不能」を省くことで kappa 係数が低下した所見項目は、「Collapse/Ischemic」であった。本来、判別不能が省かれた場合はクラス数が一つ減るため kappa 係数は上昇しやすいはずであるが、それにも関わらず判別不能を除外すると kappa 係数が上昇するという事はつまり、「ある画像に対して「Collapse/Ischemic」が判別不能かどうか」に関しては一致するが、判別が可能である場合の所見スコアリング自体は一致しないということであるため、このような所見はスコアリングの基準が不十分である可能性が考えられた。

次に、「判別不能」を入れた場合と「判別不能」を除いた場合とで、どちらの kappa 係数がより重要であるかを検討した。すでに述べたとおり、腎臓病理の所見では kappa 係数の平均が 0.4 以上であることが求められ、kappa 係数の値だけで見るならば判別不能を除いた解析の方がクラス数が減少する分、所見一致度は高くなりやすい。しかしながら、「判別不能」を入れて算出した kappa 値のほうが、臨床的には重要であると考えられる。なぜならば、糸球体画像においてある所見項目が「判別不能」であるとは、「ある所見項目のスコアを決める上で、その対象となる構造物が糸球体画像に含まれないため判別不能である」という情報を含意するからである。例えば、ある糸球体画像において「Mesangial Hypercellularity」が判別不能であるということは、その糸球体に本来存在するはずのメサンギウム領域が残っていないということ意味、メサンギウム領域がなくなっていくような臨床病態が想定される。そのため、判別不能を入れた所見一致度の方がより重要であると考えられる。

5.3 Kappa 係数の低下要因

医師間で同一の画像に対する所見スコアリングに差異が見られること、つまり kappa 係数が低下する要因の一つとして、医師のキャリアの違いが考えられた。IgA Nephropathy Classification Working Group の 2016 年の報告⁷⁾によると、261 人の IgA 腎症のコホート study⁸⁾において、ワーキンググループの中心となる医師に比べて、そうでない医師は、「Mesangial Hypercellularity」、「Endocapillary Proliferation」の所見に対して、2~3 倍の頻度でそれらの項目を陽性にする傾向があった。本研究に参加した 5 人の医師においても、キャリアが大きく異なっていることから、それが kappa 係数の低下に繋がった可能性が考えられた。他の要因として、今回の実験では PAS 染色のみを対象としたことが考えられた。通常一つの糸球体を評価する際には、HE や PAM といった複数の染色画像を合わせて所見を判断する事が多い。今回行った実験では単一の染色画像のみを用いて通常と異なる方法で行ったため、kappa 係数の低下に繋がった可能性が考えられた。また、より重要な要因として、本研究で開発した所見定義のスコアリング基準には、所見間の依存関係を明示しなかったことが考えられた。例えば、ある糸球体画像において「Sclerosis」の所見が Global であるということは、全ての血管腔が閉塞しているということになるため、血管腔に関する所見である「Endocapillary Proliferation」は判別不能とするべきであるが、その依存関係のルールをスコアリング基準に明示しなかったため、医師によってはこれを正しく判断できず、kappa 係数の低下に繋がった可能性が考えられた。

5.4 開発した所見定義の意義

現在、任意の糸球体疾患に対する所見定義は存在しない。よって、Deep Learning といった技術をもって任意の糸球体疾患疾患の画像の所見分類に取り組もうにも、付けた所見の正しさを保証できない。今回作成した所見定義リストを改定し高

い所見一致度を得られれば、その所見定義リストに従って所見をつけることでその正しさを保証でき、Deep Learningをはじめとする人工知能関連技術による糸球体病理画像の所見分類の実用化に資すると考えられる。

6. 結論

腎糸球体疾患の自動所見付けの実現に向けて、疾患の種類を限定しない所見定義を開発した。その定義に基づき、疾患を限定しない糸球体画像について 5 名の医師による所見付けを行なった結果、‘Sclerosis’、‘Cellular Crescent’、‘Collapse/Ischemic’の3項目は kappa 係数の平均が 0.4 を超えたが、他はそれ未満であった。所見一致度を上げるために、所見定義のさらなる改善が必要である。

7. 倫理的配慮・謝辞

本研究は東京大学大学院医学系研究科・医学部倫理委員会による承認を得て実施した (承認番号 11455)。本研究は厚生労働科学研究費補助金 (H28-ICT-一般-010) の助成を受けた。

8. 参考文献

- 1) Kitterer, D. et al., 2015. Diagnostic impact of percutaneous renal biopsy. *Clinical Nephrology*, 84(6), pp.311-322
- 2) Terunuma T, Tokui A, Sakae T. Novel real-time tumor-contouring method using deep learning to prevent mistracking in X-ray fluoroscopy. *Radiol Phys Technol* [Internet]. Springer Singapore; 2018;11(1):43-53. Available from: <https://doi.org/10.1007/s12194-017-0435-0>
- 3) Roberts IS, Cook HT, et al. Working Group of the International IgA Nephropathy Classification Working Group. The Oxford classification of IgA nephropathy: pathology definitions, correlations, and reproducibility. *Kidney Int.* 2009; 76:546-556.
- 4) Adam B, Randhawa P, Chan S, et al. Banff initiative for quality assurance in transplantation (BIFQUIT): reproducibility of polyomavirus immunohistochemistry in kidney allografts. *Am J Transplant.* 2014; 14:2137-2147.
- 5) D'Agati VD, Fogo AB, Bruijn JA, Jennette JC (2004) Pathologic classification of focal segmental glomerulosclerosis: a working proposal. *Am J Kidney Dis* 43:368-382
- 6) Barisoni L, Troost JP, Nast C, Bagnasco S, Avila-Casado C, Hodgkin J, Palmer M, Rosenberg A, Gasim A, Liensziewski C, Merlino L, Chien HP, Chang A, Meehan SM, Gaut J, Song P, Holzman L, Gibson D, Kretzler M, Gillespie BW, Hewitt SM. Reproducibility of the NEPTUNE descriptor-based scoring system on whole-slide images and histologic and ultrastructural digital images. *Mod Pathol.* 2016 Jul;29(7):671-84.
- 7) Hernan trimarchi et al. Oxford Classification of IgA nephropathy 2016: an update from the IgA Nephropathy Classification Working Group. *Kidney International* Volume 91, Issue 5, May 2017, Pages 1014-1021
- 8) Coppo R, Lofaro D, Camilla RR, et al. Risk factors for progression in children and young adults with IgA nephropathy: an analysis of 261 cases from the VALIGA European cohort. *Pediatr Nephrol* 2017; 32: 139-150.

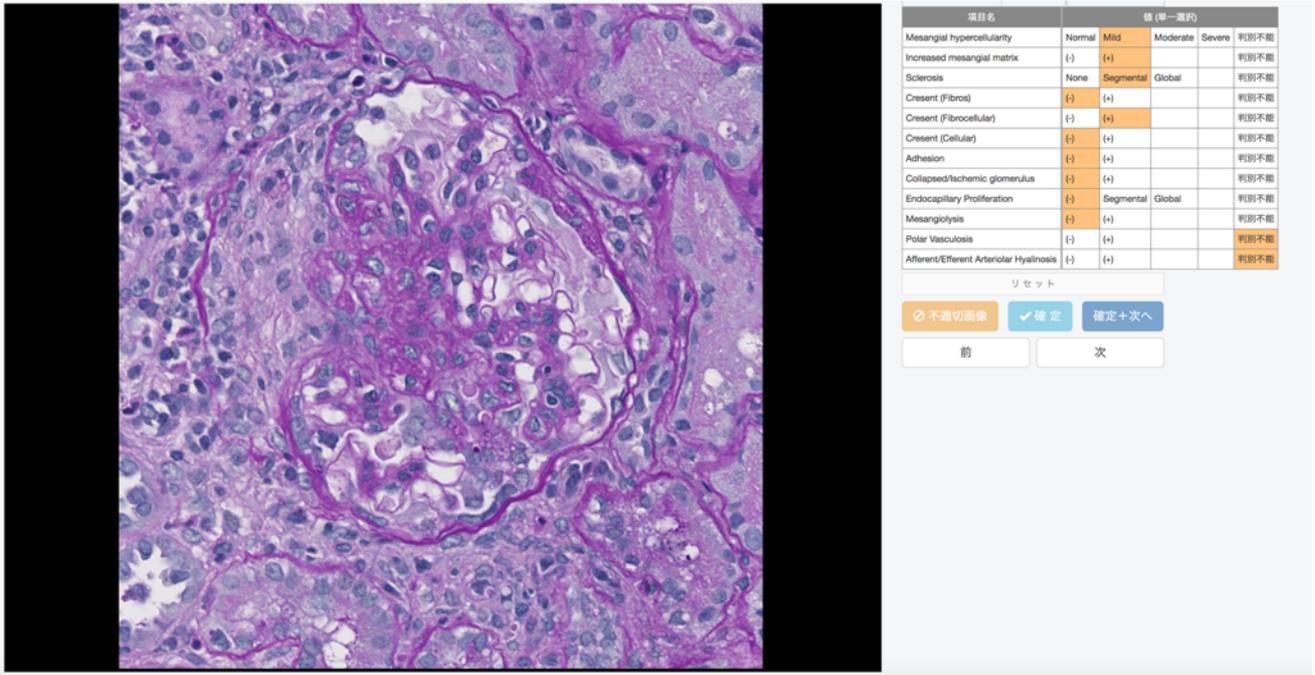


図 1 所見付けを行うためのウェブベースのソフトウェア

画面左: 所見付け対象となる糸球体画像。一枚の画像に主となる糸球体を 1 つ含む。

画面右: 各所見項目に対するスコアリングを行うユーザーインターフェース。「不適切画像」でない限り、全 12 種類の所見を評価しなければ確定することができない。

項目名	選択可能な値				Oxford分類に含まれる	スコアリング基準	
1 Mesangial hypercellularity	Normal	Mild	Moderate	Severe	判別不能	○	もともと細胞増殖の多いメサンギウム領域(血管種以外)で一領域あたりのメサンギウム細胞をカウントし、以下の基準で所見をつける。 4未満:normal,4-5:mild,6-7:moderate,8以上:severe メサンギウム基質が完全に残っていない場合は、評価不能と判断する。
2 Increased mesangial matrix	(-)	(+)			判別不能	○	メサンギウム領域において、基質面積が細胞2つの領域を越えれば(+)と判断する。 基質代替性領域はPAS染色では評価対象外である。メサンギウム基質が完全に残っていない場合は、判断不能とする。
3 Sclerosis	None	Segmental	Global		判別不能	○	細胞外基質により血管腔が閉塞したもの。泡沫細胞やヒアリノーシスの有無を問わない。 わずかも血管腔が空いていればsegmental,全ての血管腔が詰まっていればglobalと判断する
4 Adhesion	(-)	(+)			判別不能	○	糸球体係蹄とポウマン囊が接しているもの。ただし、管外病変及びsclerosisの場所で接しているものは含まない。 Crescentで接している部分があっても、それは評価しない
5 Collapsed/ischemic glomerulus	(-)	(+)			判別不能	○	糸球体毛細血管が硬化以外で虚脱しているもの。ポウマン囊の肥厚の有無や線維化の有無は問わない。 わずかも虚脱係蹄がみられれば、陽性とする。係蹄が一切残っていない時は判断不能とする
6 Polar Vasculosis	(-)	(+)			判別不能	○	血管極部に、輸入輸出細動脈以外の血管が増生している場合に、(+)と判断する。血管壁の硝子化は必須ではない。 血管極が見えないものは「判断不能」とする。
7 Endocapillary Proliferation	(-)	Segmental	Global		判別不能	○	血管腔内に細胞が増えており、血管内腔が狭小化しているもの。細胞増殖が見られないものはつけない。 糸球体全体に見られる場合はglobal,一部に見られる場合はsegmentalと所見をつける。 血管腔が一つも画像に確認できないときは、判断不能とする。
8 Afferent/Efferent Arteriolar Hyalinosis	(-)	(+)			判別不能	○	輸入もしくは輸出細動脈にPAS染色でヒアリン様物質の沈着があれば陽性と判断する。程度は問わない。 輸入動脈および輸出動脈、どちらか片方であれば陽性とする。画像内に輸入動脈も輸出動脈もみられない場合は判断不能とする。
9 Mesangiolytic	(-)	(+)			判別不能	○	メサンギウム細胞が変性しており、かつ、メサンギウム基質構造が融解している部分があれば(+)と判断する。 メサンギウム基質が一切残っていない場合は、判断不能とする。
10 Crescent(Fibrous)	None	Fibrous			判別不能	○	ポウマン囊の周の10%以上を占めている管外病変を定義する。三層以上細胞が重なっており細胞のしめる割合が50%以上のものをcellular,基質が90%以上を占めるものをfibrous,それ以外をfibrocellular 合併例がある時はそれぞれ別個につける
11 Crescent(Fibrocellular)	None	Fibrocellular			判別不能	○	
12 Crescent(Cellular)	None	Cellular			判別不能	○	

表 1 本研究で開発した所見定義

左より、所見項目名、スコアの取りうる値、Oxford 分類に含まれるかどうか、スコアリングの基準を表す。

もともと Oxford 分類に含まれている所見定義から、変更を加えた部分は赤字で表す。

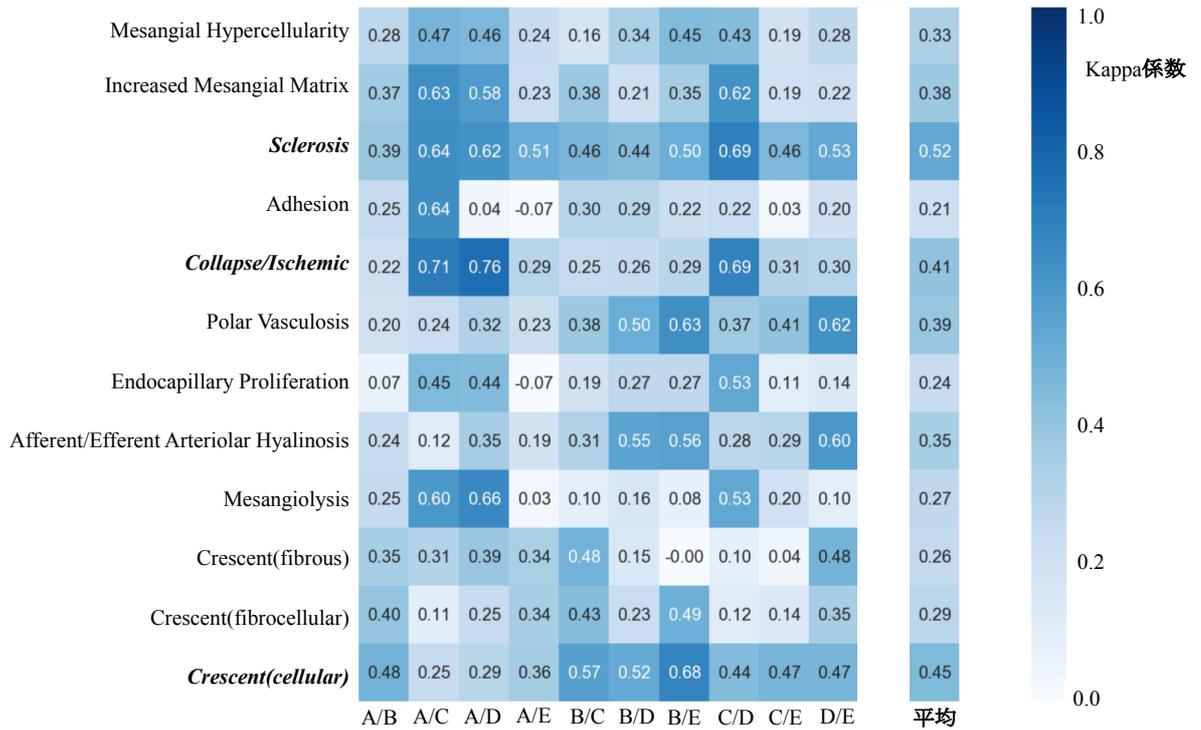


表 2 Kappa 係数結果

縦軸は、所見項目を表す。

横軸の A~E は医師 A~E に対応し、A/B は医師 A と医師 B との所見の一致を見ていることを意味する。

イタリック体で記載される所見は、平均 kappa 係数が 0.4 を超えるものである。

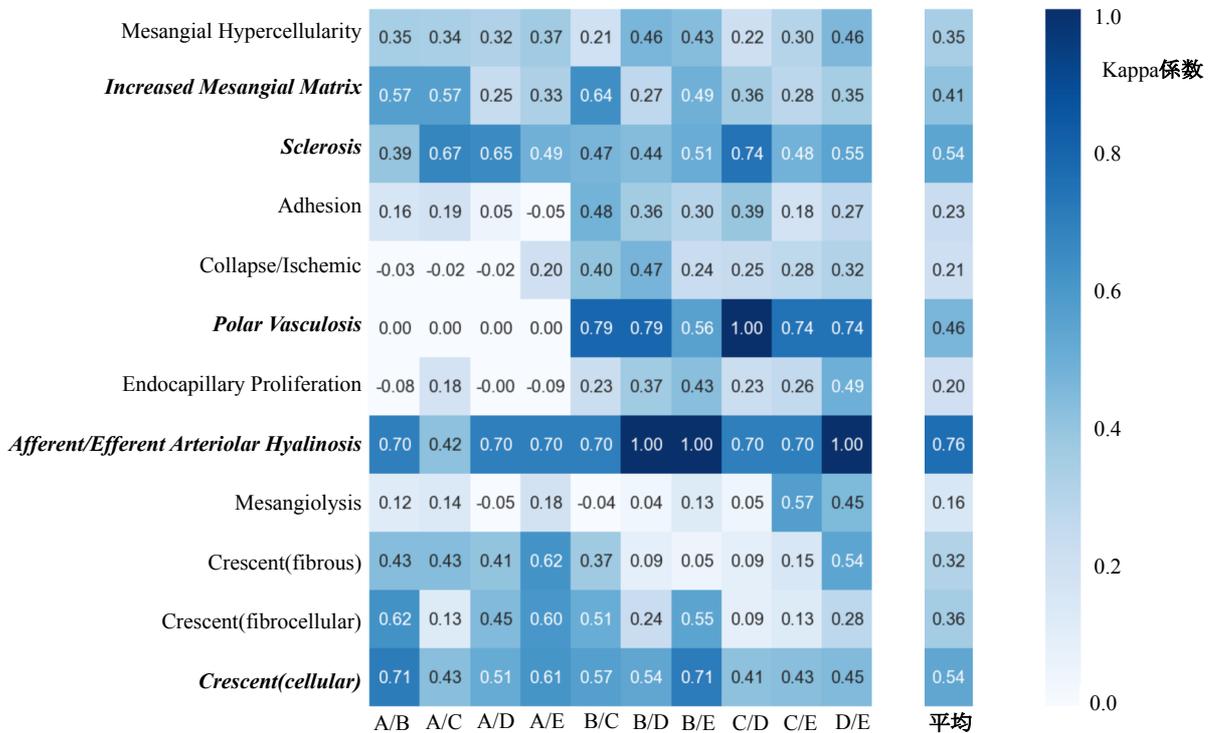


表 3 Kappa 係数結果(所見ごとに判別不能画像を除いたもの)

縦軸は、所見項目を表す。

横軸の A~E は医師 A~E に対応し、A/B は医師 A と医師 B との所見の一致を見ていることを意味する。

イタリック体で記載される所見は、平均 kappa 係数が 0.4 を超えるものである。