

1 研究の要旨

背景：病理スライドガラス標本をデジタル化した Whole Slide Imaging (以下 WSI とする) による病理一次診断 (= 最終診断) が可能であることを証明する研究は国外において複数認められている。すでにヨーロッパでは WSI による病理一次診断が実施されている。米国でも、WSI 技術に対して米国臨床病理医協会 (College of American Pathologists: CAP) からガイドラインが発表されると同時に、2016 年にはその使用をアメリカ食品医薬品局 (U.S. Food and Drug Administration: FDA) が承認する可能性が高いといわれている。しかしながら、日本からは WSI による病理一次診断の有用性を示す根拠となる報告がなく、デジタルパソロジーに関して日本は世界に大きく遅れをとっている。

目的：多施設共同研究によって WSI による病理一次診断が、診断精度の観点から、光学顕微鏡によるスライドガラスでの病理一次診断に劣らず有用であるかどうかを検討する。

対象と方法：研究者は病理診断業務に専ら 10 年以上従事し、日本病理学会の定める病理専門医資格を保有する医師とする。参加者には検討前に、CAP ガイドラインに基づいて WSI 診断のためのトレーニングセット 60 症例を行い、全問正解することを要件とする。参加施設において偏りの無い生検検体および 5 ブロック以下の手術検体連続 100 症例を集積し、HE 染色標本の WSI にて病理 1 次診断を行う。スライドガラス標本をデジタル化するバーチャルスライドスキャナは各施設にすでに設置された機器を用いて、対物レンズ倍率 20 倍もしくは 40 倍でスキャンを行う。WSI の閲覧についてはモニタによる画質の劣化を避けるために、解像度 3840 × 2160 ピクセル以上のいわゆる 4K モニタを用いることとし、操作するパーソナルコンピュータに関しては 4K モニタ投影可能なスペックを満たすこととする。2 週間のウォッシュアウト期間を空けた後、HE 染色スライドガラス標本の光学顕微鏡による病理診断を行う。診断者内における WSI 診断と光学顕微鏡によるスライドガラス標本診断の一致率は Concordance：一致した症例、Minor concordance：一致していないが治療方針や生命予後に影響しない症例、Major discrepancy：治療方針や生命予後に影響する症例の 3 段階評価とした。不一致率などを算出したのち、診断結果および WSI 診断と光学顕微鏡によるスライドガラス標本診断の不一致症例について合議の下、診断の妥当性と不一致の原因について検討する。

結果：全 9 施設から 900 症例 1070 検体が集積された。一致率については Concordance：1023 検体 (95.6%)、Minor discrepancy：37 検体 (3.5%)、Major discrepancy：9 検体 (0.9%) であった。合議によって、Major discrepancy 9 検体のうち 8 症例が光学顕微鏡によるスライドガラス標本診断が妥当と判断された。臓器別、検体採取方法、診断の相違点などに特記すべき傾向は認めなかった。また、施設別の不一致率検討結果では、Minor discrepancy が 0.0-8.8%、Major discrepancy が 0.0-3.0% であり、施設毎の一致率はほぼ同等であった。

結論：生検と小手術の HE 標本における WSI 診断と光学顕微鏡によるスライドガラス標本診断の病理診断不一致率は極めて低く、一定の条件下であれば WSI 診断による病理一次診断が有用であると考える。今後、特殊染色や免疫染色を加えた評価や、手術材料に対する検討も必要である。

2 研究の背景

病理スライドガラス標本をデジタル化した Whole Slide Imaging (以下 WSI とする) を観察することにより、病理一次診断 (= 最終診断) を行うことが可能であることを証明する研究は国外において複数認められている(1-13)。既にヨーロッパでは昨年 6 月に実証データに基づき、バーチャルスライドシステムに対して医療機器としての CE マークが承認された事を受けて、WSI による病理一次診断が実施されている。米国では、2013 年 5 月に病理一次診断を見越した WSI 診断に対して米国臨床病理医協会 (College of American Pathologists: CAP) からガイドラインが発表されている(14) と同時に、アメリカ食品医薬品局 (U.S. Food and Drug Administration: FDA) もバーチャルスライドシステムを医療機器として承認すべく調査を開始しており、2016 年に WSI 診断が承認される可能性が高いといわれている。一方、本邦では、WSI 診断の精度に関する検証データや論文報告等がほとんどなく、欧米に大きく水を開けられているのが現状である。

また、WSI には様々な可能性が秘められている。デジタル技術である WSI による診療が実施された場合、放射線画像診断においてデジタル技術が導入された事例と同様の技術革新が起こることが予想される。すなわち、デジタルデータとしてストレージ管理とデジタル化による遠隔診断であり、これを発展させるためのハード面及びソフト面での技術発展等である。将来的に必ず訪れる顕微鏡から WSI への切り替えに遅れを取ることは、経済の観点からも好ましくない。

もう一点は診断精度の担保に関する可能性である。申請者が厚生労働省難治性疾患克服研究事業びまん性肺疾患に関する調査研究班報告として 2010 年に報告した如く、間質性肺炎の病理診断では、呼吸器を専門としない病理医の一致率は、値で 0.13 と極めて低く、専門性の高い領域における診断者間の一致率は低いことが確認された(15)。WSI 技術によって急速に進むと思われる遠隔診断は、ダブルチェックとして診断精度を改善する可能性を持ち、専門性の高い疾患での診断一致率を向上させることが確実であるとともに、現在水面下にて発生していると推測される過大医療を未然に防ぐことで医療費を削減することに繋がると考える。本邦に負のガラパゴス化を産まない為にも、早急に WSI による一次診断の実用性および安全性を検討する国内研究が必要であり、その後、WSI を用いたダブルチェックの必要性と有用性を証明していくこととなる。

3 研究の目的/意義

当該研究の目的は、WSIによる病理一次診断が光学顕微鏡によるスライドガラスでの病理診断に劣っていないことを検証することであり、統計学的に解析可能な十分な症例数を得ることが必要であり、多施設共同研究とした。さらに多施設共同研究では、単一の施設で得られるデータなどの偏りを除くことが出来るというメリットもある。WSI診断は日本病理学会認定の病理専門医で、日本デジタルパソロジー研究会所属の病理医が行い、実際の臨床検体を用いて、実臨床に極めて近いデータの取得を目指すと同時に、推奨されるスペックをコンセンサスにて決定し、日本におけるWSI一次診断のガイドライン作成を目指す。当該研究はまずWSIによる一次診断の実現性と安全性を多施設にて示し、実施可能な推奨スペックも同時に検討する。これにより、来年度から診断者間一致率の検討やダブルチェックの必要性の研究へとつながるものとする。

日本における病理医の現状を鑑みた場合、診断病理医の育成は急務であるが、緊急に解決することの困難なテーマと言える。病理後継者は持続的に減少しており、病理医の高齢化が進む一方である現状は、全国的に大きく問題視されている。長崎県を例に挙げると、日本病理学会のホームページ上では26名の病理専門医を確認することが出来るが、現役で実働の病理医は16名しか存在しない。単純に病理専門医の数でその充足度を予測できないのが病理医の現状である。申請者は2006年からこの課題を極めて深刻と受け取り、解決に取り組んでいるが、この中で、デジタル病理技術による技術革新が、本邦の病理診断レベルの向上と病理医育成に必須と考えた。

病理学においてデジタル分野と分子病理分野は今後の大きな医療発展の鍵となる2大テーマであり、日本が世界に遅れをとることはグローバルヘルスにおいて、大きな経済チャンスを逃すことにもなる。

4 研究の方法

4.1 研究デザイン

後ろ向きコホート研究

4.2 研究概要

WSI による一次診断の為にトレーニングセット 60 症例を全分担者並びに協力者が行う。長崎大学病院の症例を用いて、Philips IntelliSite Pathology Solution Ultra Fast Scanner (Philips Health, Netherlands)を用いて対物レンズ 40 倍にてデジタル化を行う。分担者及び協力者はトレーニングセットによる診断を行い、診断が全問正解となった時点で研究に参加する。

各施設において、分担者及び協力者は、特に偏りの無い生検検体および5ブロック以下の手術検体連続100症例を集積し、HE染色標本のWSIにて一次診断を行う。2週間のウォッシュアウト期間を空けた後、HE染色スライドガラス標本にて診断を行う。診断者内におけるWSI診断とガラス標本診断の一致率、不一致率などを算出したのち、診断結果およびWSI診断とスライドガラス診断の不一致症例のWSIをハードディスクに保存して、長崎大学病院に返送する。長崎大学は不一致症例のWSIと診断結果をとりまとめ、ハードディスクにて各施設に再配布し、各分担者及び協力者が独自に検討した後に全体会議を行い、診断の妥当性と不一致の原因について検討する。

4.3 症例数

4.3.1 トレーニングセット

20臓器(口腔、唾液腺、乳腺、食道、脳、胃、腸、胆嚢、肝臓、膵臓、肺、甲状腺、子宮、骨髄、リンパ節、卵巣、骨軟部、皮膚、膀胱、前立腺)各3症例(腫瘍2例、非腫瘍1例)、計60症例とする。

4.3.2 検討症例

各施設100症例を目安として、12施設全1200症例とする。

4.4 対象選択・除外基準

4.4.1 トレーニングセット選択基準

平成 27 年 1 月 1 日から 3 月 1 日の間に長崎大学病院病理診断科・病理部にて診断された症例。

生検検体内から広く 20 臓器（口腔、唾液腺、乳腺、食道、脳、胃、腸、胆嚢、肝臓、膵臓、肺、甲状腺、子宮、骨髄、リンパ節、卵巣、骨軟部、皮膚、膀胱、前立腺）の腫瘍 2 例、非腫瘍 1 例を基準に偏りなく選定する。症例数は 2013 年の CAP ガイドライン(14)によって、診断の一致度を検討するために推奨されている症例数による。

4.4.2 検討症例選択基準

実際の検討に用いる症例は、平成 27 年 1 月 1 日から平成 27 年 12 月 31 日の間に各施設で診断された症例を対象とし、年齢、性別及び、入院或いは外来は不問とする。

研究期間内において各施設の判断で連続 100 症例を抽出。症例数は 2013 年の CAP ガイドライン(14)によって、診断の一致度を検討するために推奨されている症例数による。

4.4.3 除外基準

本研究の目的は HE 染色に対する診断であるため、診断は HE 染色によるものに限定し、特殊染色や免疫染色、遺伝子検査を要したものは検討から除外する。また、スキャナの性能評価では無く、WSI による診断評価が目的であるため、スキャンエラーによって標本全体に焦点が合っていない症例も除外する。

4.5 デジタル化

各施設が保有するスライドスキャナーを用いて、デジタル化する。対物レンズ倍率は 20 倍もしくは 40 倍とするが、各施設に一任する。使用するスキャナと対物レンズ倍率は以下に列挙する。

- Philips IntelliSite Pathology Solution Ultra Fast Scanner (Philips Health, Amsterdam, NETH), 40 倍
- Aperio AT2 scanner (Leica Biosystems, San Diego, CA, USA), 20 倍
- Nanozoomer 2.0-HT C-9600-13 (Hamamatsu, Hamamatsu, Shizuoka, JPN), 20 倍
- Nanozoomer 2.0-RS C-10730-13 (Hamamatsu, Hamamatsu, Shizuoka, JPN), 20 倍
- Nanozoomer 2.0-RS C-10730-13 (Hamamatsu, Hamamatsu, Shizuoka, JPN), 40 倍
- VS120 -L100-J (Olympus Corporation, Tokyo, JPN), 40 倍

- VS800 (Olympus Corporation, Tokyo, JPN), 40 倍
- FINO (CLARO, Hirosaki, Aomori, JPN), 40 倍

4.6 情報収集

下記情報を各施設の病理診断システムより収集する。

- 年齢
- 性別
- 臓器
- 採取方法

4.7 評価項目

- WSI 診断時評価項目
 - 病理組織診断
 - 確診度 (90%以上、50-90%、<50%の3段階評価)
 - 診断難易度 (平易、やや困難、困難の3段階評価)
 - 確定診断或いは記載診断か?
 - フォーカスの問題点
- WSI 診断から 2 週間のウォッシュアウト時間を空けたのちに行ったスライドガラスによる病理組織診断
 - WSI 及びスライドガラス診断の診断者内一致度 (Concordance, Minor discrepancy, Major discrepancy の3段階評価)
 - 診断の妥当性: 診断者が不一致と判断した症例については consensus により WSI 診断或いはガラス診断のどちらが妥当かどうかを判定する。

4.8 統計解析

- 診断者内における WSI 診断とスライドガラス診断の診断一致度
- 一致度別の症例特性 (臓器、採取方法、WSI 診断時評価項目)

5 利益相反

本研究にかかわる研究者等は各施設における臨床研究に係る利益相反管理指針の規定に従って必要事項を申告し、各施設における利益相反審査委員会の審査を受け承認を得るものとする。

6 結果

6.1 症例背景

6.1.1 年齢（全体）

集積した症例は9施設より900症例1070検体であった。集積された検体の年齢については欠側値:203検体、データが得られたものは0-10歳:8検体(0.9%)、11-20歳:15検体(1.7%)、21-30歳:42検体(4.8%)、31-40歳:91検体(10.5%)、41-50歳:166検体(19.1%)、51-60歳:137検体(15.8%)、61-70歳:181検体(20.9%)、71-80歳:179検体(20.6%)、81歳以上:48検体(5.5%)であった(Figure 1, Table1)

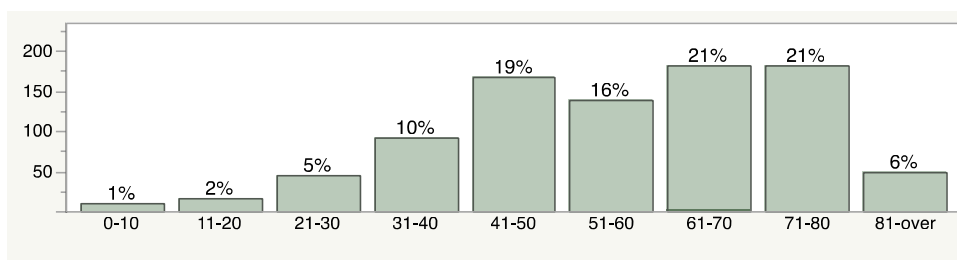


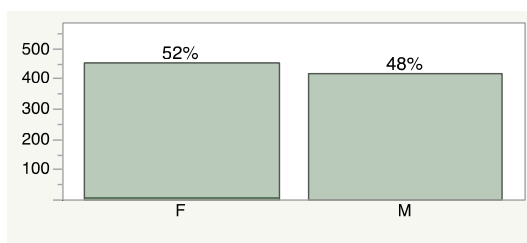
Figure 1 年齢分布（全体）

年齢	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-over	合計
N	8	15	42	91	166	137	181	179	48	867
%	0.9%	1.7%	4.8%	10.5%	19.1%	15.8%	20.9%	20.6%	5.5%	

Table 1 年齢分布（全体）

6.1.2 性別（全体）

性別については欠側値:208検体、男450検体(47.8%)、女412検体(52.2%)であった(Figure 2, Table 2)



性別	F	M	合計
N	450	412	862
%	52.2%	47.8%	

Table 2 性別（全体）

Figure 2 性別（全体）

6.1.3 臓器（全体）

臓器については胃:258検体(24.1%)、結腸:240検体(22.4%)、子宮頸部:97検体(9.1%)、乳腺:62検体(5.8%)、皮膚:57検体(5.3%)、食道:38検体(3.6%)、前立腺:37検体(3.5%)、子宮体部:29検体(2.7%)、口腔:28検体(2.6%)、尿路系:27検体(2.5%)、直腸:21検体(2.0%)、骨髄:16検体(1.5%)、咽頭/十二指腸/小腸(それぞれ):15検体(1.4%)、胆膵:13症例(1.2%)、女性付属器:11症例(1.0%)、軟部組織:10検体(0.9%)、虫垂/肺(それぞれ):9検体(0.8%)、リンパ節/喉頭/心血管系(それぞれ):8検体(0.7%)、扁桃:7検

体 (0.7%)、子宮内容物：5 検体 (0.5%)、骨：4 検体 (0.4%)、中枢神経系/肝/唾液腺/肛門 (それぞれ)：3 例 (0.3%)、外陰部/腎 (それぞれ)：2 検体 (0.2%)、甲状腺/上皮小体/体腔液/腸間膜/涙囊/脾臓/腔 (それぞれ)：1 例 (0.1%) であった (Figure 3, Table 3)。

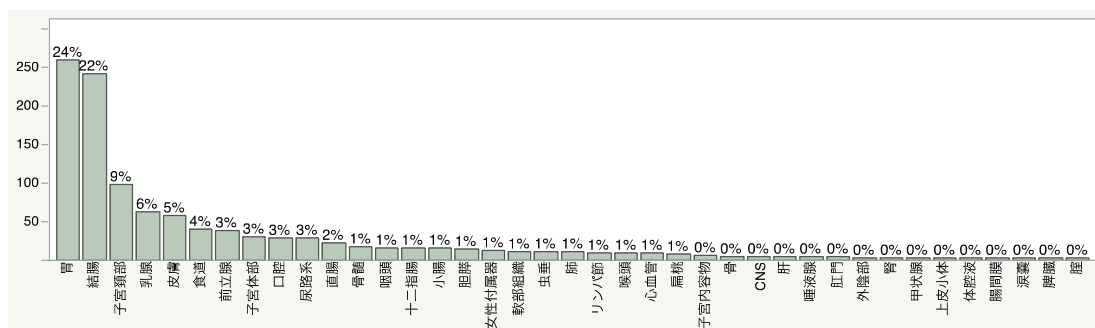


Figure 3 臓器分布 (全体)

臓器	胃	結腸	子宮頸部	乳腺	皮膚	食道	前立腺	子宮体部	口腔	尿路系	直腸	骨髄	咽頭	十二指腸	小腸	胆膵	女性付属器	軟部組織	虫垂	肺	
N	258	240	97	62	57	38	37	29	28	27	21	16	15	15	15	13	11	10	9	9	
%	24.1%	22.4%	9.1%	5.8%	5.3%	3.6%	3.5%	2.7%	2.6%	2.5%	2.0%	1.5%	1.4%	1.4%	1.4%	1.2%	1.0%	0.9%	0.8%	0.8%	
臓器	リンパ節	喉頭	心血管	扁桃	子宮内容物	骨	中枢神経系	肝	唾液腺	肛門	外陰部	腎	甲状腺	上皮小体	体腔液	腸間膜	涙囊	脾臓	腔	-	合計
N	8	8	8	7	5	4	3	3	3	3	2	2	1	1	1	1	1	1	1	-	1070
%	0.7%	0.7%	0.7%	0.7%	0.5%	0.4%	0.3%	0.3%	0.3%	0.3%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	-	

Table 3 臓器分布 (全体)

6.1.4 採取方法（全体）

採取方法については biopsy:716 検体(66.9%)、small surgery (EMR/ESD):97 検体(9.1%)、needle biopsy:68 検体(6.4%)、excision biopsy:67 検体(6.3%)、surgery:54 検体(5.0%)、curettage:29 検体(2.7%)、polypectomy:20 検体(1.9%)、aspiration:17 検体(1.6%)、others:2 検体(0.2%)であった (Figure 4, Table 4)。

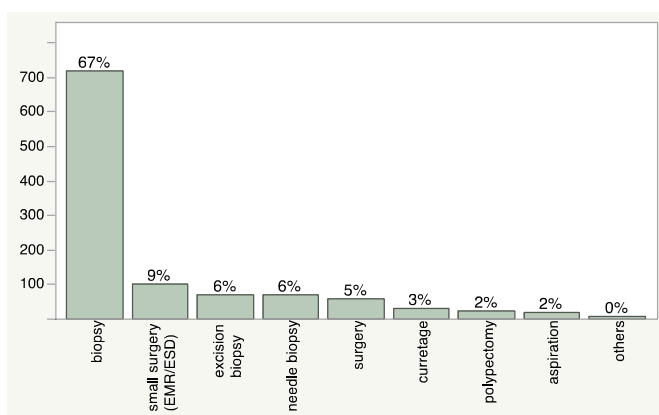


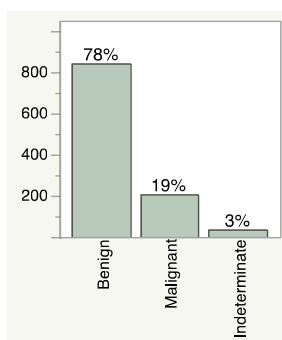
Figure 4 検体採取方法（全体）

採取方法	biopsy	small surgery (EMR/ESD)	needle biopsy	excision biopsy	surgery	curettage	polypectomy	aspiration	others	合計
N	716	97	68	67	54	29	20	17	2	1070
%	66.9%	9.1%	6.4%	6.3%	5.0%	2.7%	1.9%	1.6%	0.2%	

Table 4 検体採取方法（全体）

6.1.5 良悪性（全体）

良性病変(Benign)は 839 検体 (78.4%)、悪性病変(Malignant)は 199 検体 (18.6%)、良悪性判定不能(Indeterminate)が 32 検体 (3.0%)であった (Figure 5, Table 5)。



	Benign	Malignant	Indeterminate	合計
N	839	199	32	1070
%	78.4%	18.6%	3.0%	

Table 5 良悪性の割合（全体）

Figure 5 良悪性割合（全体）

6.1.6 腫瘍性病変か否か（全体）

非腫瘍性病変(Non-neoplastic)は 596 検体（55.7%）、腫瘍性病変(Neoplastic)は 438 検体（40.9%）、判定不能(Indeterminate)が 36 検体（3.4%）であった（Figure 6, Table 6）。

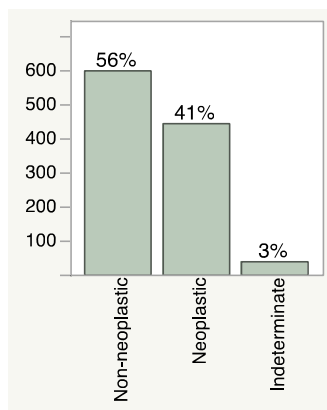


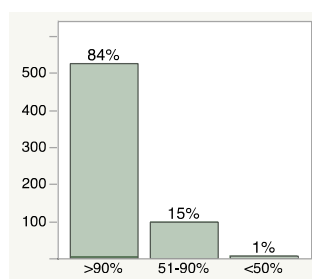
Figure 6 腫瘍性病変か否か（全体）

	Non-neoplastic	Neoplastic	Indeterminate	合計
N	596	438	36	1070
%	55.7%	40.9%	3.4%	

Table 6 腫瘍性病変か否か（全体）

6.1.7 診断の確診度（全体）

回答が得られたのは 620 検体（欠測値 450 検体）。90%以上の確診度を持って診断した検体は 522 検体（84.2%）、51-90%の確診度を持って診断した検体は 94 検体（15.2%）、確診度が 50%に満たないものが 4 検体（0.6%）であった（Figure 7, Table 7）



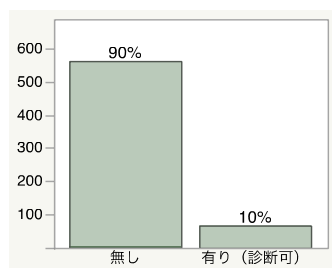
確診度	>90%	50-90%	<50%	合計
N	522	94	4	620
%	84.2%	15.2%	0.6%	

Table 7 診断の確診度（全体）

Figure 7 診断の確診度（全体）

6.1.8 WSI 焦点の問題（全体）

回答が得られたのは 620 検体（欠測値 550 検体）。症例選択基準として焦点に問題があって診断できなかったものは対象に含まれていない。焦点に問題が無かったものは 556 検体（89.7%）、診断可能だが焦点に問題があったものは 64 検体（10.3%）であった（Figure 8, Table 8）。



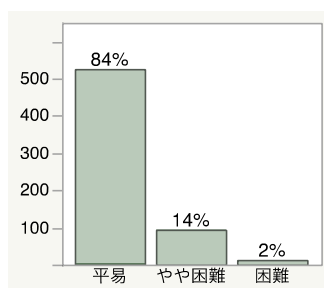
Focus 問題	無し	有り（診断可）	合計
N	556	64	620
%	89.7%	10.3%	

Table 8 WSI 焦点の問題（全体）

Figure 8 WSI 焦点の問題（全体）

6.1.9 診断難易度（全体）

回答が得られたのは 620 検体（欠測値 450 検体）。診断の難易度として平易と回答されたのは 521 検体（84.0%）、やや困難と回答されたのは 89 検体（14.4%）、困難と回答されたのは 10 検体（1.6%）であった（Figure 9, Table 9）。



難易度	平易	やや困難	困難	合計
N	521	89	10	620
%	84.0%	14.4%	1.6%	

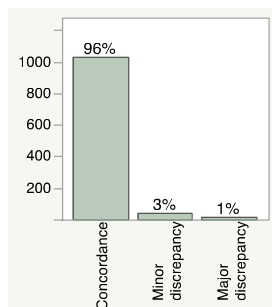
Table 9 診断難易度（全体）

Figure 9 診断難易度（全体）

6.2 不一致症例検討

6.2.1 一致度（全体）

診断の一致度としては、Concordance：1023 検体（95.6%）、一致していないが治療方針や生命予後に影響しない Minor discrepancy:37 検体（3.5%）、治療方針や生命予後に影響する Major discrepancy：9 検体（0.9%）であった（Figure10, Table10）。



一致度	Concordance	Minor discrepancy	Major discrepancy	合計
N	1023	38	9	1070
%	95.6%	3.6%	0.8%	

Table 10 一致度（全体）

Figure 10 一致度（全体）

6.2.2 一致度（施設別）

施設別の一致度を Table 11 に示す。Concordance の割合は 89.2-99.0%、Minor discrepancy の割合は 0.0-8.8%、Major discrepancy の割合は 0.0-3.0% であり、施設間での大きな違いは認めない。

施設	Concordance N(%)	Minor discrepancy N(%)	Major discrepancy N(%)	計
A	195(96.5)	6(3.0)	1(0.5)	202
B	132(89.2)	13(8.8)	3(2.0)	148
C	95(97.9)	2(2.1)	0(0.0)	97
D	93(93.0)	4(4.0)	3(3.0)	100
E	98(98.0)	2(2.0)	0(0.0)	100
F	99(99.0)	1(1.0)	0(0.0)	100
G	121(96.8)	3(2.4)	1(0.8)	125
H	99(99.0)	0(0.0)	1(1.0)	100
I	91(92.9)	7(7.1)	0(0.0)	98

Table 11 一致度（施設別）

6.2.3 臓器（不一致症例）

診断の不一致が見られた症例は13臓器であった（Table 12）。Major discrepancyは結腸6検体（結腸の2.5%）、食道1検体（食道の2.6%）、子宮頸部1検体（食道の1.0%）、胃1検体（胃の0.4%）で生じた。Minor discrepancyは結腸12検体（結腸の5.0%）、子宮頸部8検体（子宮頸部の8.2%）、胃7検体（胃の2.7%）、直腸2検体（直腸の9.5%）、虫垂1検体（虫垂の11.1%）、肺1検体（肺の11.1%）、尿路系1検体（尿路系の3.7%）、口腔1検体（口腔の3.6%）、子宮体部1検体（子宮体部の3.4%）、前立腺1検体（前立腺の2.7%）、皮膚1検体（皮膚の1.8%）、乳腺1検体（乳腺の1.6%）で生じた。

	結腸	食道	子宮頸部	胃	虫垂	肺	直腸	尿路系	口腔	子宮体部	前立腺	皮膚	乳腺
Concordance(%)	222 (92.5)	36 (94.7)	88 (90.7)	250 (96.9)	8 (88.9)	8 (88.9)	19 (90.5)	26 (96.3)	27 (96.4)	28 (96.6)	36 (97.3)	56 (98.2)	61 (98.4)
Major discrepancy(%)	6(2.5)	1(2.6)	1(1.0)	1(0.4)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)
Minor discrepancy(%)	12(5.0)	1(2.6)	8(8.2)	7(2.7)	1(11.1)	1(11.1)	2(9.5)	1(3.7)	1(3.6)	1(3.4)	1(2.7)	1(1.8)	1(1.6)

Table 12 臓器（不一致症例）

6.2.4 採取方法（不一致症例）

診断の不一致が見られた採取方法は7方法であった（Table 13）。Major discrepancyはbiopsy 6検体（biopsyの0.8%）、EMR/ESD 2検体（EMR/ESDの2.1%）、polypectomy 1検体（polypectomyの5.0%）で生じていた。Minor discrepancyはbiopsy 23検体（3.2%）、EMR/ESD 9検体（EMR/ESDの9.3%）、needle biopsy 2検体（needle biopsyの3.0%）、excision biopsy 2検体（excision biopsyの2.9%）、curettage 1検体（curettageの3.4%）、surgery 2検体（surgeryの1.9%）で生じていた。

	Biopsy	small surgery (EMR/ESD)	polypectomy	curettage	needle biopsy	excision biopsy	surgery
Concordance(%)	687(95.9)	86(88.7)	19(95.0)	28(96.6)	65(97.0)	66(97.1)	53(98.1)
Major discrepancy(%)	6(0.8)	2(2.1)	1(5.0)	0(0.0)	0(0.0)	0(0.0)	0(0.0)
Minor discrepancy(%)	23(3.2)	9(9.3)	0(0.0)	1(3.4)	2(3.0)	2(2.9)	1(1.9)

Table 13 採取方法（不一致症例）

6.2.5 良悪性（不一致症例）

Major discrepancy は良性腫瘍 7 検体（良性腫瘍の 0.8%）、悪性腫瘍 2 検体（悪性腫瘍の 1.0%）で見られた。Minor discrepancy は良性腫瘍 23 検体（良性腫瘍の 2.7%）、悪性腫瘍 13 検体（悪性腫瘍の 6.5%）、判定不能病変 2 検体（判定不能病変の 6.3%）で見られた（Table 14）。

	Benign	Malignant	Indeterminate
Concordance	809(96.4)	184(92.5)	30(93.8)
Major discrepancy	7(0.8)	2(1.0)	0(0.0)
Minor discrepancy	23(2.7)	13(6.5)	2(6.3)

Table 14 良悪性（不一致症例）

6.2.6 腫瘍性病変か否か（不一致症例）

Major discrepancy では腫瘍性病変 5 検体（腫瘍性病変の 1.1%）、非腫瘍性病変 3 検体（非腫瘍性病変の 0.5%）、判定不能病変 1 検体（判定不能病変の 2.8%）で見られた。Minor discrepancy では腫瘍性病変 25 検体（腫瘍性病変の 5.7%）、非腫瘍性病変 8 検体（非腫瘍性病変の 1.3%）、判定不能病変 5 検体（判定不能病変の 13.9%）で見られた（Table 15）。

	Neoplastic	Non-neoplastic	Indeterminate
Concordance	408(93.2)	585(98.2)	30(83.3)
Major discrepancy	5(1.1)	3(0.5)	1(2.8)
Minor discrepancy	25(5.7)	8(1.3)	5(13.9)

Table 15 腫瘍性病変か否か（不一致症例）

6.2.7 診断確診度（不一致症例）

Major discrepancy は確診度 90%以上の症例で 2 検体（0.4%）、確診度 51-90%の症例で 3 検体（3.2%）認めた。Minor discrepancy は確診度 90%以上の症例で 11 検体（2.1%）、確診度 51-90%の症例で 7 検体（7.4%）認めた（Table 16）。

	>90%	51-90%	<50%
Concordance (%)	509(97.5)	84(89.4)	4(100)
Major discrepancy (%)	2(0.4)	3(3.2)	0(0)
Minor discrepancy (%)	11(2.1)	7(7.4)	0(0)

Table 16 診断確診度（不一致症例）

6.2.8 WSI 焦点の問題（不一致症例）

Major discrepancy では焦点に問題があった症例が 1 検体（1.6%）、問題がなかった症例が

4 検体 (0.7%) であった。Minor discrepancy では焦点に問題があった症例が 1 検体 (1.6%)、問題がなかった症例が 17 検体 (3.1%) であった (Table 17)。

	無し	有り (診断可)
Concordance (%)	535(96.2)	62(96.9)
Major discrepancy (%)	4(0.7)	1(1.6)
Minor discrepancy (%)	17(3.1)	1(1.6)

Table 17 WSI 焦点の問題 (不一致症例)

6.2.9 診断難易度 (不一致症例)

Major discrepancy は平易と判断された症例 4 検体 (0.8%)、やや困難とされた症例 1 検体 (1.1%) に認められた。また、Minor discrepancy は平易と判断された症例 5 検体 (1.5%)、やや困難とされた症例 10 検体 (11.2%) に認められた (Table 18)。

	平易	やや困難	困難
Concordance (%)	509(97.7)	78(87.6)	10(100.0)
Major discrepancy (%)	4(0.8)	1(1.1)	0(0.0)
Minor discrepancy (%)	8(1.5)	10(11.2)	0(0.0)

Table 18 診断難易度 (不一致症例)

6.2.10 対物レンズ倍率

Major discrepancy は対物レンズ倍率 20 倍の 2/427 検体 (0.5%)、対物レンズ倍率 40 倍の 7/643 検体 (1.1%) に認められた。また、Minor discrepancy は対物レンズ倍率 20 倍の 11/427 検体 (2.6%)、対物レンズ倍率 40 倍の 27/643 検体 (4.2%) に認められた (Table 19)。

対物倍率	20x	40x
Concordance	414 (97.0)	609 (94.7)
Major discrepancy	2 (0.5)	7 (1.1)
Minor discrepancy	11 (2.6)	27 (4.2)

Table 19 対物レンズ倍率 (不一致症例)

6.2.11 Major discrepancy 症例の詳細

Major discrepancy を認めた検体の WSI による診断とガラス標本による診断を Table 18 に示す。Consensus により WSI 診断或いはガラス診断のどちらが妥当か検討した結果、ガラスが妥

当だと判断された検体が 9 検体、WSI が妥当だと判定された症例が 1 検体であった (Table 20)

	臓器	WSI 診断	ガラス診断	倍率	妥当性
Case 1	胃	Poorly differentiated adenocarcinoma	Regenerative mucosa	40x	ガラス
Case 2	食道	Ulcer, r/o low grade intraepithelial neoplasia	High grade intraepithelial neoplasia	40x	ガラス
Case 3	結腸	Group 3, Low grade adenoma, Group 3	Group 1	40x	ガラス
Case 4	結腸	Group 1, no adenoma	Group 3, low grade adenoma	40x	WSI
Case 5	子宮頸部	Chronic cervicitis	CIN2	40x	ガラス
Case 6	結腸	Group 3, Tubular adenoma, low to high grade	Group 5, tubular adenocarcinoma	40x	ガラス
Case 7	結腸	Tubular adenoma, low to high grade	Carcinoma in adenoma	40x	ガラス
Case 8	結腸	Chronic colitis with granulomatous lesion, compatible with Crohn disease	Chronic colitis, non-specific	20x	ガラス
Case 9	結腸	Group 5, adenocarcinoma	Group 3, adenoma	20x	ガラス

Table 20 Major discrepancy 検体の診断

WSI 及びガラス診断での良悪性判定を Table 21 に示す。良悪性の判定が一致するのは 5 検体、良性から悪性に変更された検体が 3 検体、悪性から良性に変更された検体が 1 検体であった (Table 21)。

	WSI 良悪性判定	ガラス良悪性判定
Case 1	Malignant	Benign
Case 2	Benign	Malignant
Case 3	Benign	Benign
Case 4	Benign	Benign
Case 5	Benign	Benign
Case 6	Benign	Malignant
Case 7	Benign	Malignant
Case 8	Benign	Benign
Case 9	Malignant	Malignant

Table 21 WSI とガラス診断の良悪性判定比較 (Major discrepancy)

WSI 及びガラス診断での腫瘍性が否かの判定を Table 22 に示す。判定が一致するのは 4 検体、腫瘍性から非腫瘍性、あるいは非腫瘍性から腫瘍性に変更された検体がそれぞれ 2 検体ずつ、判定不能から腫瘍性に変更された検体が 1 検体であった。

	WSI 腫瘍性状判定	ガラス腫瘍性状判定

Case 1	Neoplastic	Non-neoplastic
Case 2	Indeterminate	Neoplastic
Case 3	Neoplastic	Non-neoplastic
Case 4	Non-neoplastic	Neoplastic
Case 5	Non-neoplastic	Neoplastic
Case 6	Neoplastic	Neoplastic
Case 7	Neoplastic	Neoplastic
Case 8	Non-neoplastic	Non-neoplastic
Case 9	Neoplastic	Neoplastic

Table 22 WSI とガラス診断の腫瘍性状判定比較 (Major discrepancy)

6.2.12 Minor discrepancy の詳細

Minor discrepancy を認めた検体の WSI による診断とガラス標本による診断を Table 21 に示す。Consensus により WSI 診断或いはガラス診断のどちらが妥当か検討した結果、ガラスが妥当だと判断された検体が 20 検体、WSI が妥当だと判定された症例が 17 検体であった (Table 23)

	臓器	WSI 診断	ガラス診断	倍率	妥当性
Case 10	子宮頸部	CIN2 > 1	CIN1	40x	WSI
Case 11	結腸	Group 3, high grade adenoma	Group 3, low grade adenoma	40x	WSI
Case 12	結腸	Group 3, low to high grade adenoma	Group 3, high grade adenoma	40x	WSI
Case 13	子宮頸部	Chronic cervicitis with squamous metaplasia > CIN1	CIN1	40x	WSI
Case 14	子宮頸部	Chronic cervicitis with squamous metaplasia > CIN1	CIN1	40x	WSI
Case 15	前立腺	Atypical glands	Benign prostate	20x	WSI
Case 16	結腸	Group 5, tubular adenocarcinoma	Group 3, high grade tubular adenoma	20x	WSI
Case 17	子宮頸部	Chronic cervicitis	CIN1	40x	WSI
Case 18	結腸	High grade tubular adenoma	Low grade tubular adenoma	40x	WSI
Case 19	子宮頸部	CIN2	CIN3	40x	WSI
Case 20	口腔	Squamous cell carcinoma, suspicious, edge(-)	Oral intraepithelial neoplasia, edge(-)	40x	WSI
Case 21	胃	Group 4, atypical glands	Group 5, adenocarcinoma	40x	WSI
Case 22	食道	High grade intraepithelial neoplasia	Low grade intraepithelial neoplasia	40x	WSI
Case 23	皮膚	Psoriasisiform dermatitis	Lichen planus-like keratosis	40x	WSI
Case 24	肺	Organizing acute lung injury	Non-diagnostic, hemorrhage and cellular interstitial pneumonia	40x	WSI
Case 25	胃	Group 4	Group 2	20x	WSI
Case 26	胃	Group 4	Group 2	20x	WSI
Case 27	乳腺	Ductal carcinoma in situ, high grade	Invasive carcinoma of no special type, high grade, suspicious,	40x	ガラス
Case 28	尿路系	Erosive mucosa, possible for malignancy	Erosive mucosa without malignancy	40x	ガラス
Case 29	胃	Group 1	Group 2	40x	ガラス
Case 30	胃	Group 1	Group 2	40x	ガラス
Case 31	結腸	Group 3, high grade adenoma	Group 3, low grade adenoma	40x	ガラス
Case 32	結腸	Group 4, high grade adenoma	Group 5, adenocarcinoma	40x	ガラス
Case 33	結腸	Group 3, sessile serrated adenoma/polyp	Group 1, hyperplastic polyp	40x	ガラス
Case 34	子宮頸部	Chronic cervicitis > CIN1	CIN1	40x	ガラス
Case 35	子宮頸部	CIN2	CIN3	40x	ガラス
Case 36	胃	Indefinite for neoplasia	Tubular adenocarcinoma, well differentiated	40x	ガラス
Case 37	結腸	Tubular adenoma, low grade	Tubular adenoma, low to high grade	40x	ガラス
Case 38	結腸	Hyperplastic polyp	Low grade serrated adenoma, edge(-)	20x	ガラス
Case 39	虫垂	Acute catarrhal appendicitis	Acute phlegmonous appendicitis	20x	ガラス
Case 40	子宮体部	Endometrioid carcinoma, G1	Endometrial intraepithelial neoplasia	20x	ガラス
Case 41	胃	Group 4, Adenocarcinoma, suspected	Group 2	40x	ガラス
Case 42	子宮頸部	CIN2	CIN3	40x	ガラス
Case 43	直腸	Group 4	Group 2, atypical epithelium	20x	ガラス
Case 44	直腸	Group 4	Group 2, atypical epithelium	20x	ガラス
Case 45	結腸	Group 5	Group 4	20x	ガラス
Case 46	結腸	Tubular adenoma, low to high grade	Carcinoma in adenoma	20x	ガラス
Case 47	結腸	Group 3, high grade adenoma	Group 3, low grade adenoma	40x	どちらも

Table 23 Minor discrepancy 検体の診断

7 考察

7.1 一致率評価

本研究は、日本より発信される WSI 診断の有用性を検討する最初かつ、最大の研究である。本研究では、診断一致率について 3 段階で評価した。すなわち、WSI 及びガラス標本での診断が一致したと考える Concordance、不一致ではあるが治療方針や生命予後に影響しない Minor discrepancy、治療方針や生命予後に影響すると考えられる Major discrepancy の 3 段階である。結果は、Concordance : 1023 検体 (95.6%)、Minor discrepancy: 38 検体 (3.6%)、Major discrepancy: 9 検体 (0.8%) であった。WSI 診断とガラス標本診断が、どの程度一致すれば WSI がガラス標本に劣らず有用なのであろうか。この点については CAP のガイドラインにも記載されておらず、明確な基準は示されていない。とはいうものの、WSI 診断とガラス標本診断の一致率に関する報告例が多数なされており、その有用性を示している。術中迅速診断検体を用いた Fallon らの報告では良悪性の判定において 96% の一致率があった事を示しているし(1)、同様に術中迅速診断を用いた検討では正診率が 98% とする報告や、89% とする報告がされている(2, 3)。WSI による一次診断に着目すると、Molnar らによる胃生検検体の検討では一致率が 95.1% と報告されているし(4)、稀少症例や難解症例を除いた 101 症例での検討において治療に影響を与える不一致の発生率を 4.4% とする報告や(5)、日常診療で得られた連続 607 症例での検討において WSI による不一致率を 1.65% とする報告(6)や 17 名の病理医による 3017 症例の検討では不一致率を 0.7% と報告している(7)。また、専門家へのコンサルテーションを要する難解症例についての報告もなされている。Wilbur らはコンサルテーション症例において依頼元医師と専門家との診断一致率を 85%、WSI と光学顕微鏡との一致率を 91% と報告しているし(8)、様々な臓器におけるコンサルテーション症例 217 症例に対する検討では不一致率が 0.9% であった(9)。これらの報告例と比較しても、本研究の結果は遜色ないものと思われる。

しかしながら、この不一致率が WSI と顕微鏡の違いによるものであるのか、単に診断者内の病理診断再現性に起因した問題であるのかと疑問は残る。病理診断の再現性について診断者間の一致率については多く報告されているものの、診断者内の一致率についての報告はほとんど無い。Raab らの報告では診断者内不一致率が 7% であり、major discrepancy は 5% であったと報告している(10)。また、Major discrepancy の割合を 6% とする報告や “error rates” を 2% とする報告もされている(11, 12)。さらに Bauer らは光学顕微鏡による再検鏡において Major discrepancy が 524 症例中 9 症例 (1.72%) があったと報告した(6)。これらの報告を鑑みても、本研究においてガラス標本診断における診断者内一致率の検討は行っていないが、Major discrepancy が 0.9% であったという結果は診断者内における不一致率としても矛盾しないと考ええる。

本研究では WSI の見え方や WSI に対する慣れに関する問題にも留意し、WSI 診断の際に使用

するモニタの解像度を 3840×2160 ピクセル以上の 4K モニタを用いることで、モニタの性能による画質劣化の可能性を排除した。また、WSI に対する不慣れを解決するために、WSI 診断に比較的習熟している施設を選択するとともに、CAP ガイドライン(14)に基づいて、検討前に WSI 診断のトレーニングを行う事でさらなる習熟度の改善を行った。このことで診断能力以外の問題要素については可能な限り排除できたと考える。

以上のことから WSI による一次診断はガラス標本による一次診断と同等の診断が可能と判断した。

7.2 不一致理由

病理診断において重要であることは治療方針や生命予後の正しい判断材料たることである。本研究において不一致が発生した検体については全参加者による合議を行い、Minor 或は Major discrepancy の再検討を行うとともに、WSI 或はガラス標本による診断のどちらが妥当であるかの検討を行った。Major discrepancy が生じた検体では、概ねガラス標本診断が妥当であるとの結果となり、合議において検討者の意見は比較的合致した。Snead にも WSI とガラス標本診断の不一致症例において、どちらの診断がより妥当であるかの検討を合議によって行っているが、不一致 21 症例中、9 症例は WSI 診断が、12 症例はガラス標本診断が妥当であったと報告している(7)。我々の結果とはやや異なるため、診断内容や検体情報、WSI の状況等について詳細に検討した。良悪性の判定や腫瘍性か否かの判定が異なった検体はそれぞれ約半数であった。本研究では検討者に対して診断時の確診度や診断の確診度や、検体の診断難易度について回答を求めた。Major discrepancy 検体は少ないこと、回答があったものが約半数だったこと等の問題はあっても特に傾向は認めなかった。臓器や採取方法についても特定の傾向を認めなかった。また、WSI の状況として焦点問題の有無や WSI スキャン時の対物レンズ倍率についても検討したが、特定の傾向を認めなかった。Major discrepancy とされた症例が少なく、症例数を増やした検討や臓器別に検討することで、Major discrepancy が生じる傾向を検討する必要があると考える。

一方、Minor discrepancy と評価された検体は、合議の結果、ガラスが妥当だと判断された検体が 20 検体、WSI が妥当だと判定された症例が 17 検体、合議によっても回答が出なかったものが 1 検体であった。妥当な診断を判断した合議の際にも意見が割れた検体がほとんどであった。臓器や採取方法による特異性、焦点問題の有無、診断の確診度や難易度に関しての特定の傾向は認めなかった。これらのことから、診断者間の一致率も低く、判定の困難な症例において Minor discrepancy が生じる可能性が示唆される。

7.3 施設別評価

本研究は多施設研究による検証であり、ガラス標本をデジタル化するスキャナは各施設保有のものを利用した。スキャナは5社8機種に及び対物レンズの倍率は20倍もしくは40倍であった。各施設の検証結果では、Minor discrepancyの割合は0.0-8.8%、Major discrepancyの割合は0.0-3.0%と多少のばらつきはあるものの、概ね良好な結果であった。施設毎の一致率は診断者毎の一致率を反映するとともに、スキャナや対物レンズ倍率の影響を反映すると考えられる。本研究において、診断者は全員が病理診断業務に専ら従事して10年以上の経験を有しており、日本病理学会が定める専門医資格保有である。さらにCAPガイドライン(14)に基づいて、検討前にWSI診断のトレーニングを行う事でさらなる習熟度の改善を行った。以上の点から、診断者の診断能力に著しい優劣は無く、診断能力による影響を最小限まで排除出来たと仮定すると、施設別一致率は対物レンズ倍率やスキャナの性能を反映されており、結果として対物レンズの倍率やスキャナの種類による診断一致率への影響はほとんど無いと考えた。対物レンズの倍率は極めて重要な問題であり、2つの点に關与する。1つ目はWSIを保管するストレージに関する問題である。デジタル化によって飛躍的な進歩を遂げた放射線画像と比べると画像1枚あたりの容量や1症例あたりの画像の枚数は病理がはるかに多いため、保存する媒体はかなりの容量が必要となることが予想される。診断の質を下げることなく、低倍率でのスキャンが可能であるなら、ファイルサイズを縮小することが出来、ストレージの負担も軽減できると考えられる。2つ目はスキャンに要する時間の問題である。単純にWSIによる一次診断を行うためには、Turn around time (TAT)にスキャン時間が加わるため、デジタル化に要する時間が圧縮できることに越したことはない。従って、低倍率スキャンはスキャン時間も圧縮できるため、診断一致率と対物レンズの倍率に相関がなかった事は有用であると考えられる。

しかしながら、低倍率でのデジタル化には一定の問題が存在する。低倍率では観察できない対象が存在することである。過去の報告では炎症細胞や造血系腫瘍、微生物、核分裂像など対物レンズ倍率20倍では観察できず、40倍ないし60倍でスキャンされていないと観察困難と報告されている(6, 7)。今回の検討では造血系腫瘍が少なかったことや、核分裂数の評価を行っていないが、例えば、*Helicobacter pylori*のような微生物は観察困難であった。対象症例や臓器によってスキャン倍率を変えるとといった検討も必要となる可能性が示唆される。

7.4 WSI の利点・欠点

今回の研究では検証していないが、WSI を用いた診断では診断者に大きな利点が生じる。既往検体の参照という点において大きな利点を得られる。通常、既往のガラス標本は保管庫から取り出して検鏡するため、標本の検索や閲覧までに一定の時間を要する。さらに、ガラス標本の場合は複数年経過すると退色して観察が困難になるものも多い。しかし、WSI であればサーバーにアクセスすることで容易に閲覧が可能であると共に、標本の退色もない。また、WSI には測定ツールが存在するので、腫瘍の深達度や大きさ、断端距離の測定などが極めて容易であり、かつ正確に行われる。例えば、胃癌粘膜切除検体の場合、その深達度を 100 μm 間隔で測定する必要があり、深達距離によっては病期分類が異なってしまう。さらに、粘膜下浸潤が 500 μm を越えるかどうかで、外科的胃切除術を追加するかどうかの判断を迫られる。また、セキュリティの問題はあるものの、Virtual Private Network による秘匿回線網を形成することが出来れば、難解症例をエキスパートにコンサルテーションしたり(13)、病理医が 1 人しかいない病院の病理診断をダブルチェックすることで診断精度を担保することも可能になると思われる。

欠点としては、WSI を管理する媒体の問題であろう。デジタル化によって飛躍的な進歩を遂げた放射線画像と比べると画像 1 枚あたりの容量や 1 症例あたりの画像の枚数は病理がはるかに多いため、保存する媒体はかなりの容量が必要となることが予想される。しかしながら、媒体の単価は実質的に下がり続けており、フィルムなどの磁気媒体の改良やクラウドサービスの出現など、技術的な発展により補填できる可能性が示唆される。もう 1 点は WSI を操作するデバイスの問題である。マウスによる操作が主体であり、対物レンズ 10 倍程度の中等度倍率での全視野観察では見落とされる領域が出現する可能性も否定できない。また、マウスによるドラッグアンドドロップ、スクロールホイールの回転を繰り返す事で操作するが、顕微鏡でのステージ操作と比して手への負担が極めて高い。各メーカーによる新しいデバイスの開発も見られることから近い将来に改善される事が期待される。

7.5 Limitation

本研究における limitation は対象選択によるものである。本研究では対象を日常診療で得られた症例から連続症例とし、5 ブロック以下の HE 染色のみとした。この影響で臓器分布としては日常診療と同様の頻度になったものの、検討が著しく少ない臓器が生じたことで、診断一致率の臓器特異性が検討出来ていない。また、手術検体や、特殊染色或は免疫染色の検討も出来ていない。さらに選択基準の影響で、自動的に難解症例が除外された結果となっており、難解症例についての検討も出来ていない。今後、臓器別の検討やコンサルテーション症例を集

積しての検討など追加の検討を行う必要がある。

7.6 WSI の今後の展望

日本病理学会において一次診断に対する WSI 利用は喫緊の問題であると理解されており、デジタルパソロジー検討委員会が設置され、運用やガイドライン策定に関する議論が始まったところである。また、今回の結果は一定の条件下においては WSI による一次診断が有用である事が示された。しかし、これにより WSI 診断が光学顕微鏡による病理診断を即座に置換可能と言うわけではない。初期費用の問題もさることながら、最大の問題は「格差」であると思われる。病理医間の「温度格差」により、WSI 一次診断に否定的な意見が少なくないことも問題であるが、「地域格差」が大きな障壁でもある。病理医不足が問題となっているヨーロッパやカナダで WSI 診断が認められている一方、先進国の中でも比較的病理医が充足しているアメリカ合衆国において、WSI 診断の承認が遅れており、これと同じような状況が日本国内でも起こっている。比較的病理医が充足している都市圏と、病理医の不足が深刻な地方医療圏では、WSI 診断が可能であると言うことの意味合いも異なる。WSI による病理診断が承認されれば、病理医不在の地域中核病院における病理診断や、一人病理医による診断に対する遠隔ダブルチェック、診断困難症例のエキスパートコンサルテーションなどが可能となり、病理医不足の補填や診断の精度管理が容易になることは確実であると考ええる。

重要なことは WSI による一次診断が可能であるという承認であり、つまり病理診断のツールの一つとして認められれば、必要な地域において、必要な病理医が利用することが可能となる。導入されれば、必要に応じた実証研究や、ネットワーキング、精度管理が加速度的に進行していくことが予想される。

8 結論

本研究は、日本で初めて行われた WSI の有用性を検討する多施設共同研究であり、海外からの報告も含めて最大規模の研究の一つである。結果、治療方針や生命予後を左右するような Major discrepancy は少なく、診断者内不一致の範疇で説明可能なレベルであり、光学顕微鏡による診断とも遜色のない診断ツールであると思われる。今後は臓器別あるいは難解症例の検討など詳細な検討を行うことが望ましいと考える。