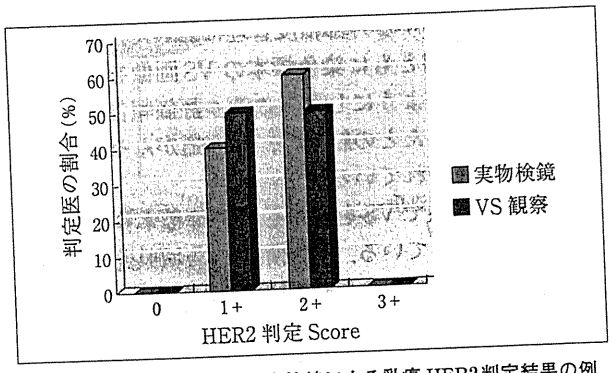

図2　VS画像および実物標本検鏡による乳癌HER2判定結果の例
乳癌HER2の判定を15人の病理医に行ってもらい，その判定結果の一例を示す．実物のHER2免疫染色標本を検鏡した結果もVS画像の観察による判定でも，1＋と判定する病理医の人数と2＋と判定する病理医の数に差はなく，判定が分かれた．また実物検鏡とVS画像観察での判定に大差はみられなかった．

てもらった．判定する症例数は15例で，事前の評価でScore 0～3＋の各段階の症例を用いた．判定基準として「HER2アトラス，トラスツズマブ病理部会」を用いた．HER2の判定はVS画像の観察によるものと実際の免疫染色標本の検鏡によるものと両方を行ってもらい，VSと実物検鏡との対比を行った．15人の病理医による評価結果を得るためにかかった時間はVSを用いた方法では約1ヵ月であったが，実物の免疫染色標本を検鏡する方法では約6ヵ月と大きな差があった．実物の免疫染色を検鏡する方法では1人の病理医が1セットの標本を検鏡し，その後，次の病理医に郵送するということを繰り返さなくてはならないが，VSによる方法では複数の病理医がサーバー内の免疫染色画像を同時に観察することが可能で，標本の占有が起こらず，また郵送の必要がないなどから判定結果を得る時間の短縮となった．病理医によるHER2判定の結果であるが，15例のうちの1例に対する判定結果を**図2**に示す．この症例はScore 1＋と判定する病理医とScore 2＋と判定する病理医に大きく別れた．1＋の判定では分子標的薬の治療対象とならないが，2＋の判定では，FISHを行い分子標的薬治療の対象となることがあり，治療方針に大きく影響する結果となった．このように病理医間での判定に差を認めたが，一つの要因としては病理経験年数が関係し，経験年数の少ない病理医で誤った判定が多い傾向がみられた．また**図2**にみられるようにVSによる判定結果と，実際の免疫染色標本を検鏡して得られた判定結果に大きな差はなく，VSでも精度管理を十分に行うことが可能であることが示された．この検討から治療に直結するような免疫染色の判定でも病理医間で判定に見過ごすことのできない差異の存在することが明らかとなり，病理診断において精度管理が重要であることが示された．精度管理では，その結果を対象となった個別の病理医にできるだけ迅速に伝える必要がある点で，精度管理にはVSを利用することは大変意義があることと思われる．精度管理の重要性からは1年に1～2回程度の精度管理を行うことが望まれるが，この点では全国規模で精度管理を行うよりも，小回りが利くという点では県単位程度でVSを利用した精度管理を行うことが現実的と考えられる．

V．VSの問題点と今後の展望

ウィルヒョウ（Virchow）の時代から150年以上，病理医は光学顕微鏡を覗き込み病理診断を行ってきたが，VSの登場によって病理医は顕微鏡の前から解き放たれ，パソコンやタブレット端末があればどこでも好きな時間に病理組織標本を観察することが可能な時代となってきた．現在，病理医の絶対的な不足が問題となっているが，その解決策の一つとして，家庭に入ったり，あるいは定年で第一線を退いた病理医に，VS画像を介した病理診断・病理診断コンサルテーションに参画してもらい，病理医不足を補っていくことが考えられる．この在宅勤務（テレワーク）を実現し，潜在的な力を引き出すためには地域内での病理医関係者の密接なコミュニティーが必要となり，そのためにも県単位の病理医関係者の学術集会の開催などは重要と考えられる．こうした業務形態を医療機関として裏づけていく制度的課題にも取り組む必要がある．

ところで従来の光学顕微鏡による検鏡に慣れた病理医にとってVSを利用した病理診断はレンズの狭い視野から広いモニター画面に変わることで違和感やその操作性の違いから，VS画像で病理診断を行っていこうとする意識や自信をもつことができるかはVSの病理診断への利用普及に重要である．この点に関して茨城県内の病理医を対象にアンケート調査を行ったので，その結果を紹介する．茨城県では茨城病院病理医の会が年2回開催され，検討症例の組織画像をVSで事前閲覧できるようになっている．このVS画像閲覧をしてもらいVSに対するアンケート調査を行った結果，VS画像を閲覧した場所は，約70％の病理医は勤務先での医療施設で閲覧していたが，約30％の病理医は自宅で閲覧しており，どこでも閲覧できるVSの利点が生かされていた．VS画像を閲覧するための操

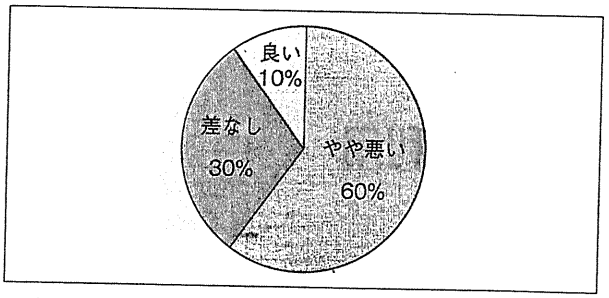

図3　VS画像の画質に対するアンケート調査結果　20人の病理医の回答で，VS画像の画質は実物の病理組織標本を検鏡するのと比較して，60%の病理医がやや悪いが気にならない程度，30%は差がみられないと回答していた．悪いと回答したものはいなかった．

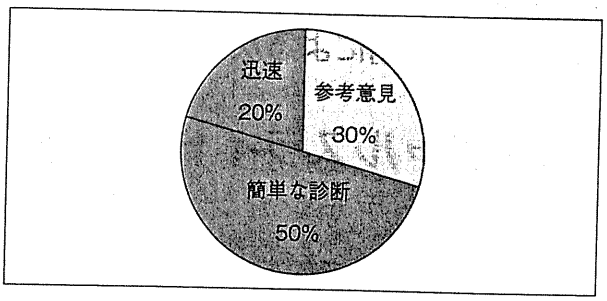

図4　VS画像による病理診断の可能性に対するアンケート調査結果　20人の病理医の回答で，VS画像を基にした病理診断を日常行うことが可能かを調査したところ，参考意見を述べるにとどまるとしたものが30%，胃炎等の簡単な病理診断までは可能としたものが50%，迅速まで可能としたものが20%であった．

作方法については操作が煩雑との回答者はなく，操作方法は全員が理解しやすいとの意見であった．VSの画質についても実物を検鏡するのと大差はみられないとの意見が多かった（図3）．VS画像での病理診断の可能性については図4で示すように簡単な胃炎程度の診断はできるとの回答者と術中迅速診断できるとの回答者の合計で70%であり，VS画像での病理診断が現時点の病理医でもある程度可能であることが示された．しかし30%の病理医はVS画像での病理診断に自信をもてずにいることもあり，今後のVSでの病理診断の普及にはVS画像への慣れが必要と考えられた．

VS画像での病理診断を行う上での問題点としてVS機器やインターネット環境にも改善の余地があるが[7]，一方で，VS画像を観察する側のパソコン，特にモニターの性能（解像度，明るさ，コントラストなど）は病理診断の質に影響を及ぼす可能性があると思われ，今後，VS本体の性能の向上を目指すだけでなく，画像を受け取る側のパソコン・モニターの信頼性・標準化も進めていく必要があると考えられる[8,9]．さらに，円滑なVS診断支援体制の運用のためには，こうしたVSやこれを取り巻く技術環境の向上と併せ，VSだけに依存せず，不安や疑問があれば常にガラス標本に戻れる運用上のバックアップ体制を整えておくことも当面必要ではないかと思われる．

現在，医学教育では，従来の光学顕微鏡を用いた病理組織学教育・実習がVSを利用して行われるようになってきている．VSを利用することで医学生の病理学へのイメージが変わり，病理診断ワークスタイルの変化が，病理医を志す若手医師の数の増加に結びつくことも期待される．VSでの実習に慣れた医学生が病理医になることで，将来，病理診断はVSで行うことが一般的となるかもしれない．しかし当面，病理医や医療資源の限られた地域ではVSをうまく利用することで，質の高い病理診断を行っていくことが可能と思われるが，そのために周辺機器を含めたVS機器やインターネット環境のさらなる改良整備が重要である．最後にこれらの機器改良やインフラの整備以上に，地域内での病理医の連携がVSをうまく活用して行くための要であり，そのためのコミュニティーづくりが最も重要である．

文　献

1) 松野吉宏：バーチャルスライドの応用—厚労省が目指すもの．Medical Technology 2008, 8：801-803
2) Costello, S.S., Johnston, D.J., O'Shea, D.G. et al.：Development and evaluation of the virtual pathology. J Med Internet Res 2003, 5：e11
3) 中村眞一，菅井　有，及川守康 他：virtual slideを用いたテレパソロジー．病理と臨床 2006, 24：373-378
4) 澤井高志，冨地信和，東福寺幾夫：遠隔病理診断．病理学と社会．病理と臨床 2009, 27（臨増）：64-72
5) 白石泰三，中野　洋，渡邊昌俊：テレパソロジーネットワークの構築とその問題点．別冊・医学のあゆみ　テレパソロジー2002．2002, 63-66
6) 東　学，丸川活司，中村厚志 他：北海道地方におけるバーチャルスライドを活用した病理組織染色外部精度管理報告．医学検査 2010, 59：835-841
7) 宇月美和，澤井高志：現状におけるバーチャルスライドの問題点と今後の開発．Medical Technology 2008, 8：835-838
8) Gilbertson, J.R., Ho, J., Parwani, A.V. et al.：Primary histologic diagnosis using automated whole slide imaging：a validation study. BMC Clin Pathol 2006, 6：4
9) 小野定康：テレパソロジーにおけるハード面での問題点．別冊・医学のあゆみ　テレパソロジー2002．2002, 44-47

DOI:10.1111/j.1365-2303.2010.00774.x

Cytodiagnosis through use of a *z*-axis video by volunteer observers: a promising tool for external quality assessment

K. Yamashiro*, M. Tagami†, K. Azuma‡, A. Nakamura§, O. Kato†, K. Taira*, M. Azuma*, H. Takeda* and H. Suzuki*

*Division of Pathology, Hokkaido Cancer Center, Kikusui 4-2, Shiroishi-ku, Sapporo, Japan, †Cytology Center, Hokkaido Cancer Society, Sapporo, Japan, ‡Department of Pathology, Sapporo Medical University Hospital, Sapporo, Japan and §Department of Laboratory, Sapporo Municipal Hospital, Sapporo, Japan

Accepted for publication 31 March 2010

K. Yamashiro, M. Tagami, K. Azuma, A. Nakamura, O. Kato, K. Taira, M. Azuma, H. Takeda and H. Suzuki

Cytodiagnosis through use of a *z*-axis video by volunteer observers: a promising tool for external quality assessment

Objective: This study examined whether cytological diagnosis through the use of a video, which shows the changing depth of focus in the microscopic field, described as a *z*-axis video, is useful compared with a still image.

Methods: From 17 cytology preparations of fine needle aspiration of the breast, we made six *z*-axis videos per case. A frame exhibiting the characteristic features was then extracted from each video and saved as a representative still image. One hundred and twenty-eight volunteer cytotechnologists were randomly divided into two groups of video observers and still image observers. The participants were asked to make a diagnosis of benign, indeterminate, suspicious or malignant without having any clinical information other than the age of the patient. Diagnoses were categorized as 'recommended' or 'unacceptable' according to degree of correlation with histology.

Results: The number of definitive diagnoses of 'benign' or 'malignant' were increased in video observers, and indeterminate or suspicious categories were decreased ($P = 0.013$). The distribution of diagnostic categories in three of the 17 cases was significantly different; the distribution in the remaining cases was similar between the two groups. The *z*-axis video observers may have selected the definite diagnoses with confidence because they observed valuable microscopic findings by 'focusing through observation'. The average number of 'recommended' diagnoses by individual observers was significantly higher in the video observer group than in the still image observer group ($P = 0.016$). In contrast, the average number of 'unacceptable' diagnoses was significantly lower ($P = 0.019$).

Conclusions: A *z*-axis video is easy to obtain and is therefore expected to become a powerful diagnostic modality for the external quality assessment of clinical cytology and even in the field of primary cytodiagnosis.

Keywords: cytology, fine needle aspiration, breast, external quality assessment, video, *z*-axis, microscopy, telepathology, telecytology

Introduction

Digital image technologies have been applied in various medical situations, and sharing microscopic images via the internet has established the practical field of telepathology.[1,2] Although the application of this process to cytology has been put forward as telecytology, its use is not yet widespread.[3–11] Some have noted that a still image hinders the growth of telecytology[12] as it merely provides an image of a single focus plane, whereas the cells and cell clusters have obvious three-dimensional structures.

However, new devices such as virtual microscopy,[13–16] with *z*-stack capability or multiframe video

Correspondence:
Katsushige Yamashiro, Division of Pathology, Hokkaido Cancer Center, Kikusui 4-2, Shiroishi-ku, Sapporo 003-0804, Japan
E-mail: yamasiro@sap-cc.go.jp

imaging along the z-axis,[17] have been introduced and may solve those problems. This study investigated whether a video cell image with changing depth of focus is useful for the external quality assessment of the cytological diagnosis compared with a still image.

Methods

Observation objects

Seventeen cytology specimens from fine needle aspirations of the breast, which had been stored in the Division of Pathology in the Hokkaido Cancer Center, were selected for this study. All glass slides were stained with the Papanicolaou method. Many specimens had been directly spread onto glass slides and some had been prepared by the SurePath® liquid-based cytology method. The diagnosis of each case has been confirmed by histological examination. Six microscopic field images showing distinct cytological features were obtained per case on video while changing the depth of focus (described as a z-axis video) with an objective lens of ×10 and ×40.

The microscopic video images were recorded with an IIDC camera, Scorpion 20SOC (Point Grey Research, Richmond, BC, Canada), which was controlled by Astro IIDC 4.0 (ASC Inc., Calgary, AB, Canada). The depth of microscopic focus was changed by hand with as constant a speed as possible. The movie files were trimmed and encoded by QuickTime Pro (Apple Inc., Cupertino, CA, USA). Finally, z-axis videos were prepared as QuickTime movie files with 24-bit depth RGB colour, 1024 × 768 pixels, 15 fps and H.264 video compression. The average size of the z-axis videos was 5.7 MB, ranging from 1.3 to 11.3 MB. A frame expressing the most characteristic feature was saved from each video file as a representative still image.

The recommended cytodiagnosis of each case was determined prior to the study from four diagnostic categories: benign, indeterminate, suspicious and malignant. 'Indeterminate' approximately corresponds to the term 'atypical' often adopted in other countries. The details of correlation between diagnostic categories and histological diagnosis is shown in Table 2.

Observers

Two hundred and sixty-eight cytotechnologists living in Hokkaido, Japan, were invited to participate voluntarily in this study. There were 175 responses, and 128 cytotechnologists agreed to be volunteer observers. After preparing the distribution of participants' characteristics, consisting of their career as a cytotechnologist, experience in breast aspiration cytology and familiarity with digital imaging technology, they were randomly assigned to two observer groups.

Compact disks containing z-axis videos were provided to one observer group, while the other observer group received a CD containing the still images. A z-axis video observer was provided with 'focusing through observation' using QuickTime Player, which could be played forwards and backwards repeatedly. QuickTime Player was a free download from the website of Apple Inc., and was provided on the CD for Windows users. Neither the computer or the monitor was prepared for the observers and they observed z-axis videos or still images on their own monitors.

The observers in each group were asked to select a diagnosis for each case out of the four diagnostic categories, without any other clinical information except the age of the patient. The observers were forbidden to consult with their colleagues or friends about the cases.

Evaluation methods

The distributions of the diagnostic categories selected by observers were compared between the two observer groups. In addition, the observer's diagnostic category was also compared with the recommended diagnosis determined in advance. If the observer's diagnostic categories were consistent with the histological diagnosis of benign or malignant, they were evaluated as 'recommended' (for some specific histological diagnoses 'intermediate' was included in this category with 'benign'), and if they disagreed by two or more categories they were evaluated as 'unacceptable' (Table 2). The Intercooled Stata 8.2 software package for Macintosh (StataCorp LP, College Station, TX, USA) was used to compute the statistics.

Study management

The study plan was evaluated by the Ethics Review Committee of the Hokkaido Cancer Center and approved in June 2009. An administration group independent of the researchers managed the study data and none of the researchers could access the personal data, including who participated in the study.

Results

One hundred and twenty-seven replies were obtained from volunteer observers, and 125 answers without any inadequate descriptions were selected for the analysis. The details of the observers are shown in Table 1. The characteristics of the observers, including their time in cytopathology practice, were established on the basis of information provided by themselves in the application forms used for their participation in the study. Table 2 shows the results of observers' diagnoses. The upper row shows the number of still image observers and the lower row shows the number of *z*-axis video observers. The overall distribution of diagnostic categories of the two observer groups was significantly different ($P = 0.013$). The *z*-axis video observers made many definitive diagnoses of 'benign' or 'malignant' and few indefinite diagnoses of 'indeterminate' or 'suspicious', in comparison with the still image observers. The same tendency was observed in three cases (cases F, G and P), while the differences between the two observer groups were not statistically significant in the remaining 14 cases.

Both 'recommended' and 'unacceptable' diagnoses were examined to evaluate diagnostic accuracy and the same three cases (F, G and P) were found to be different between the two observer groups. Table 3 shows the frequency of observers for the number of recommended diagnoses. The mode of recommended diagnoses of the *z*-axis video observers was 15, while that of the still image observers was 13. The mean of the video observers was 13.65 and that of the still image observers was 12.63. The difference was statistically significant ($P = 0.016$). The table also shows the frequency of unacceptable diagnoses. The modes of the *z*-axis video observers and still image observers were 1 and 2, respectively. The mean of the video observers was 1.21 and that of the still image observers was 1.65. The difference was statistically significant ($P = 0.019$).

The cytological characteristics of cases F, G and P were reviewed (Figures 1–3). It was difficult to determine the cytodiagnosis depending only on the nuclear atypia. Changing the depth of focus revealed the bilayer structure containing myoepithelial cells in case F, and confirmed the intracytoplasmic lumina in case G. A cribriform pattern was easily detected by 'focusing through observation' in case P.

The technical appraisals by observers were collected as a questionnaire (Table 4). There were 96 answers from the participants. The time required for determining diagnoses by a *z*-axis video observer was significantly prolonged compared with a still image observer ($P = <0.001$). Though there were quite a few observers who found difficulty in working with a *z*-axis video ($P = 0.086$), the operation of image files was generally accepted favourably by each observer group.

Discussion

The use of digital images has increased in the field of clinical cytology and still image telecytology has been introduced.[3–12] Virtual microscope scanners in clinical cytology have also been developed and telecytology using virtual microscopy will be available in the near future.[13–16] However, the applicability of this diagnostic modality remains unclear. On the other hand, the use of video images with changing depth of focus along the *z*-axis is an attractive modality. A previous report suggested that *z*-axis video observation was superior to still image observation for diagnostic accuracy.[17] However, that report might be regarded as a technical introduction to this new device without any definitive conclusions, because the results were derived from a small number of participants who were

Table 1. Features of observers (based on self-assessment)

	Career as cytotechnologist (years)			Experience in breast FNA cytology			Familiarity with digital images of cytology		
	< 3	3–10	> 10	> 1 case every few weeks	Inter-mediate	Occasional cases per year	> 1 case every few weeks	Inter-mediate	Occasional cases per year
Still image observers	9	14	39	32	17	13	8	11	43
z-axis video observers	9	15	39	34	16	13	9	12	42

Table 2. Comparison of distribution of cytodiagnoses by two observer groups

Case	Histological diagnosis	Cytodiagnosis by observers				*P*-value
		Benign	Indeterminate	Suspicious	Malignant	
A	Intracystic papilloma**	26	27	7	2*	0.306
		33	26	2	2*	
B	Fibroadenoma	61	0	0	1*	0.135
		60	3	0	0	
C	Ductal carcinoma	0	1*	6	55	0.401
		0	4*	6	53	
D	Malignant lymphoma	0	0	0	62	0.496
		0	0	2	61	
E	Ductal carcinoma	2*	14*	28	18	0.57
		4*	10*	26	23	
F	Fibroadenoma	27	19	10*	6*	**0.02**
		40	19	3*	1*	
G	Lobular carcinoma	1*	2*	22	37	**0.007**
		0	5*	7	51	
H	Phyllodes tumour**	49	6	5	2*	0.907
		51	7	4	1*	
I	Lobular carcinoma	0	0	8	54	0.352
		0	2*	9	52	
J	Ductal carcinoma	0	1*	8	53	0.622
		1*	1*	5	56	
K	Mucinous carcinoma	2*	6*	7	47	0.464
		0	7*	5	51	
L	Ductal carcinoma	1*	3*	3	55	0.795
		0	3*	3	57	
M	Intracystic papilloma**	17	34	6	5*	0.938
		19	31	7	6*	
N	Ductal carcinoma	0	3*	13	46	0.849
		0	2*	12	49	
O	Lactational change**	26	20	6	10*	0.96
		29	20	5	9*	
P	Ductal carcinoma	6*	28*	15	13	**0.003**
		7*	14*	10	32	
Q	Ductal carcinoma	1*	3*	8	50	0.171
		0	1*	3	59	
Total		219	167	152	516	**0.013**
		244	155	109	563	

Upper row: number of still image observers. Lower row: number of z-axis video observers.
*Unacceptable, deviation from 'recommended cytodiagnosis' by 2 + categories.
**Cases in which 'recommended cytodiagnosis' includes 'benign' and 'indeterminate'.
P-value: according to Pearson's chi-squared test

all quite familiar with digital cell images and the images used were diverse. Therefore, the current study was designed to address those issues.

The distributions of diagnostic categories between *z*-axis video observers and still image observers were significantly different in the current study and it was suggested that the *z*-axis video observers tended to select a more definitive diagnosis. Closer examinations of individual cases revealed no difference in 14 of 17 cases. However, the diagnostic categories in three cases were clearly different between two groups. Still image observers' diagnoses tended to remain indefinite because the nuclear atypia of cells in these three cases was obscure, whereas *z*-axis video observers selected the definite diagnoses with confidence because they easily obtained valuable microscopic

Table 3. Number of recommended and unacceptable cytodiagnoses by each observer

Group	Observers	Mode	Mean	Standard deviation	95% confidence interval		*P*-value (*t*-test)
*Recommended cytodiagnosis**							
Still image	62	13	12.63	2.39	12.02	13.24	0.016
z-axis video	63	15	13.65	2.27	13.08	14.22	
*Unacceptable cytodiagnosis***							
Still image	62	2	1.65	1.12	1.36	1.93	0.019
z-axis video	63	1	1.21	0.94	0.97	0.80	

*Recommended cytodiagnoses (includes some 'indeterminate' with 'benign', see Table 3).
**Unacceptable, deviation from recommended cytodiagnosis by two or more categories.

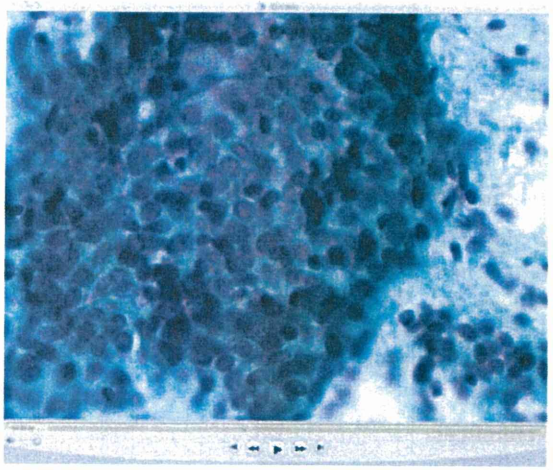

Figure 1. A z-axis video of case F, fibroadenoma. There is a cell cluster containing piled large nuclei accompanied by myoepithelia, showing so called 'bilayer structure'.

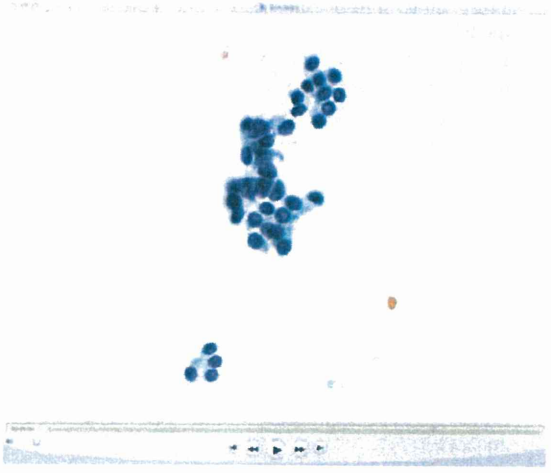

Figure 2. A z-axis video of case G, lobular carcinoma. There are cell clusters containing small but high N/C ratio hyperchromatic nuclei, also sometimes showing intracytoplasmic lumina.

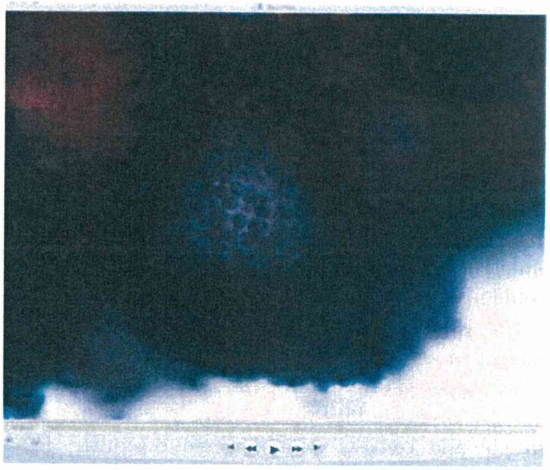

Figure 3. A z-axis video of case P, ductal carcinoma. There is a cell cluster containing many globular acellular spaces, showing so called 'cribriform pattern'. However, the nuclei are relatively small and uniform.

findings such as bilayered structure, cribriform pattern or intracytoplasmic lumina, by 'focusing through observation'. Similarly, the *z*-axis video observers were more likely to make recommended diagnoses than the still image observers, and less likely to make unacceptable diagnoses with the same three cases (cases F, G and P).

The experiment randomly assigned participants to the two observer groups to avoid maldistribution of observer characteristics. The study has proved the expectations of quality assessment and education suggested in the previous report, based on objective data. Therefore, a *z*-axis video has an advantage over still images as a tool for external quality assessment.

The major objectives of external quality assessment of clinical cytology are to monitor and improve the consistency of cytodiagnosis and to attain these goals

Table 4. Results of questionnaire concerning technical appraisal of image observation by participants

	How did you feel about working with the image files?				How long did it take you to decide the diagnoses?			
	Easy	Some difficult	Difficult	*P*-value	< 1 hour	1–2 hours	> 2 hours	*P*-value
Still image observers	46	3	0	0.0857	24	19	6	<0.001
z-axis video observers	37	9	1		6	24	17	

P-value: according to Pearson's chi-squared test.

it is necessary to observe cell images in various situations.[18–20] Glass slides are the optimal tools for this purpose, but there are considerable limitations to their use. Printed images or digital files on the internet have been used in some large-scale evaluations instead of glass slides. However, many have objected to the limitations posed by observation of a single focus plane image of cells, which have a three-dimensional structure.[12] Therefore, a video image with changing depth of focus might be adopted to achieve that aim. Of course, glass slides cannot be replaced by *z*-axis videos in all situations of external quality assessment, but they could frequently provide a viable alternative.

The current study provided CDs containing cell images to observers. However, *z*-axis videos can be downloaded via the internet easily because the average size of movie files was 5.7 MB. The equipment is simple and consists of a personal computer, IIDC camera, relay lens, microscope and some application software. It costs less than $4 000 US dollars for all of the components other than the microscope. Furthermore, it takes only 2 minutes to prepare a *z*-axis video file including recording, trimming, resizing, encoding and verifying. Therefore, the external quality assessment of clinical cytology through the use of *z*-axis videos can be carried out immediately and remotely.

As clinical cytology examines three-dimensional structures, virtual microscope scanners have recently made dramatic improvements. Z-stacked images can be made quickly, and browsing devices can smoothly handle the images. Though there are still many hurdles to clear before virtual microscopy contents can be freely disseminated via the internet, these obstacles are all expected to be overcome in time. However, *z*-axis video is a currently acceptable tool for external quality assessment and education, and its application should therefore be recommended.

This study does not question the validity of still image telecytology. It only showed the experimental results obtained from a study consisting of volunteer observers and did not deny the significance of still image observation itself. On the contrary, some still image observers showed an excellent diagnostic accuracy. Still image telecytology continues to be based on the expertise and ability of observers and technicians. The use of *z*-axis videos in the future is therefore expected to support their efforts, either directly or indirectly, as the need arises.

Acknowledgments

We heartily thank the cytotechnologists who willingly participated in this experiment and performed a difficult task. The study was supported by a Grant-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare of Japan.

References

1. Weinstein RS. Prospects for telepathology. *Hum Pathol* 1986;**17**:433–4.
2. Kayser K. Interdisciplinary telecommunication and expert teleconsultation in diagnostic pathology: present status and future prospects. *J Telemed Telecare* 2002;**8**:325–30.
3. Marsan C, Vacher-Lavenu MC. Telepathology: a tool to aid in diagnosis and quality assurance in cervicovaginal cytology. *Cytopathology* 1995;**6**:339–42.
4. O'Brien MJ, Takahashi M, Brugal G *et al.* Digital imagery/telecytology. International Academy of Cytology Task Force summary. Diagnostic cytology towards the 21st century: an international expert conference and tutorial. *Acta Cytol* 1998;**42**:148–64.
5. Galvez J, Howell L, Costa MJ, Davis R. Diagnostic concordance of telecytology and conventional cytology for evaluating breast aspirates. *Acta Cytol* 1998;**42**:663–7.
6. Yamashiro K, Kawamura N, Matsubayashi S *et al.* Telecytology in Hokkaido Island, Japan: Results of primary telecytodiagnosis of routine cases. *Cytopathology* 2004;**15**:221–7.
7. Jialdasani R, Desai S, Gupta M *et al.* An analysis of 46 static telecytology cases over a period of two years. *J Telemed Telecare* 2006;**12**:311–4.

8. Ayatollahi H, Khoei A, Mohammadian N *et al.* Telemedicine in diagnostic pleural cytology: A feasibility study between universities in Iran and the USA. *J Telemed Telecare* 2007;**13**:363–8.
9. Glatz K, Willi N, Glatz D *et al.* An international telecytologic quiz on urinary cytology reveals educational deficits and absence of a commonly used classification system. *Am J Clin Pathol* 2006;**126**:294–301.
10. Kerr SE, Bellizzi AM, Stelow EB, Frierson HF Jr, Policarpio-Nicolas ML. Initial assessment of fine- needle aspiration specimens by telepathology: Validation for use in pathology resident-faculty consultations. *Am J Clin Pathol* 2008;**130**:409–13.
11. Archondakis S, Georgoulakis J, Stamataki M *et al.* Telecytology: a tool for quality assessment and improvement in the evaluation of thyroid fine-needle aspiration specimens. *Telemed J E Health* 2009;**15**: 713–7.
12. Pantanowitz L, Hornish M, Goulart RA. The impact of digital imaging in the field of cytopathology. *Cytojournal* 2009;**6**:6.
13. Gagnon M, Inhorn S, Hancock J *et al.* Comparison of cytology proficiency testing: Glass slides vs. virtual slides. *Acta Cytol* 2004;**48**:788–94.
14. Marchevsky AM, Khuranna R, Thomas P *et al.* The use of microscopy for proficiency testing in gynecologic cytopathology: A feasibility study using Scan Scope. *Arch Pathol Lab Med* 2006;**130**:349–55.
15. Dee FR, Donnelly A, Radio S *et al.* Utility of 2-D and 3-D virtual microscopy in cervical cytology education and testing. *Acta Cytol* 2007;**51**:523–9.
16. Mori I, Nunobiki O, Ozaki T, Taniguchi E, Kakudo K. Issues for application of virtual microscopy to cytoscreening, perspectives based on questionnaire to Japanese cytotechnologists. *Diagn Pathol* 2008;**3**(Suppl 1):S15.
17. Yamashiro K, Taira K, Matsubayashi S *et al.* Comparison between a traditional single still image and a multiframe video image along the *z*-axis of the same microscopic field of interest in cytology: which does contribute to telecytology? *Diagn Cytopathol* 2009;**37**:727–31.
18. Wells CA. Quality assurance in breast cancer screening cytology: a review of the literature and a report on the U.K. *Eur J Cancer* 1995;**31**:273–80.
19. Cross P, Shiell A. Cervical cytology EQA – the Northern experience. *Cytopathology* 2001;**12**:94–103.
20. Slater DN, Hewer EM, Melling SE, Rice S. External quality assessment in gynaecological cytology: the Trent Region experience. *Cytopathology* 2002;**13**:206–19.

Gleason Score Correlation Between Biopsy and Prostatectomy Specimens and Prediction of High-grade Gleason Patterns: Significance of Central Pathologic Review

Kentaro Kuroiwa, Taizo Shiraishi, and Seiji Naito, for the Clinicopathological Research Group for Localized Prostate Cancer Investigators

OBJECTIVES To investigate the significance of dedicated central pathologic review for Gleason score (GS) correlation between the biopsy and radical prostatectomy (RP) specimens and the prediction of high-grade Gleason patterns. A discrepancy in the GS between the biopsy and RP specimens has been reported.

METHODS The Clinicopathological Research Group for Localized Prostate Cancer disease registry collated the data from 1629 patients who had undergone RP from 1997 to 2005. All biopsy and RP specimens were retrospectively re-evaluated by 2 central uropathologists according to the International Society of Urological Pathology consensus. The GS correlation between the biopsy and RP specimens and the presence of high-grade Gleason patterns (4 or 5) were recorded. The GS was categorized into 5 groups (2-4, 5-6, 3 + 4, 4 + 3, and 8-10).

RESULTS Central review significantly increased the exact concordance rate and decreased the undergrading and overgrading rates between the biopsy and RP specimens compared with local review ($P < .05$ for all). In each GS or prostate-specific antigen group, the central review biopsy GS had a significantly greater exact concordance rate with the RP specimen GS compared with the local review biopsy GS ($P < .05$ for all). Regarding high-grade Gleason patterns in the RP specimens, central review showed significantly greater sensitivity, positive predictive value, and negative predictive value than local review ($P < .05$ for all).

CONCLUSIONS We have demonstrated that central review using the International Society of Urological Pathology consensus improves the GS correlation and better predicts high-grade Gleason patterns compared with local review. We recommend central pathologic review by dedicated uropathologists for multi-institutional studies using data from prostate biopsy and RP specimens. UROLOGY 77: 407–411, 2011. © 2011 Elsevier Inc.

The Gleason grading system, proposed by Gleason[1] and represented as the Gleason score (GS) for each case, is the most widely used histologic grading system for prostate cancer. The GS in both biopsy and radical prostatectomy (RP) specimens is a powerful prognostic factor.[2,3] Accurate GS correlation between the biopsy and RP specimens is mandatory for preoperative estimation of the disease and for the planning treatment of each patient. However, the biopsy GS has been reported to have been undergraded in 18%-60% and overgraded in 6%-25% of specimens compared with the RP specimen GS.[4-11] Investigator error is one important factor for the discrepancy; thus, pathologic assessment by dedicated uropathologists might improve the GS correlation between the biopsy and RP specimens. Modern GS assessment according to the 2005 International Society of Urological Pathology (ISUP) consensus, reflecting contemporary changes regarding prostate cancer and the Gleason grading system, has shown better GS correlation than the previous assessment.[12] Pathologic assessment by dedicated uropathologists in a single academic institution has also shown better GS correlation than outside assessment.[6,11] However, the usefulness of pathologic assessment by dedicated uropathologists using the ISUP consensus for a large RP series from multiple institutions has not yet been studied.

Although high-grade Gleason patterns (4 or 5) in RP specimens, either a primary/secondary pattern or a ter-

This study was funded by the Clinical Research Foundation.

From the Department of Urology, Kyushu University Graduate School of Medical Sciences, Fukuoka, Japan; and Department of Pathologic Oncology, Mie University Graduate School of Medicine, Tsu, Japan

Reprint requests: Kentaro Kuroiwa, M.D., Ph.D., Department of Urology, Kyushu University Graduate School of Medical Sciences, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582 Japan. E-mail: humeiten@hotmail.com

Submitted: April 30, 2010, accepted (with revisions): May 22, 2010

0090-4295/11/$36.00
doi:10.1016/j.urology.2010.05.030

tiary pattern, have been reported to be related to a poor outcome, it remains unclear how effectively the biopsy GS determined by pathologic assessment by dedicated uropathologists will predict for high-grade Gleason patterns in the RP specimens.[13-15]

In the present, large-scale, multicenter study, we used the pathologic assessment by dedicated uropathologists according to the ISUP consensus for the biopsy and RP specimens from a large RP series with high-grade biopsy GSs using data from the Clinicopathological Research for Localized Prostate Cancer (CRPC) disease registry. The CRPC collates data from patients with clinically localized prostate cancer accrued from 108 academic and community practices throughout Japan. From 1997 to 2005, approximately 5000 patients with clinically localized prostate cancer who had undergone RP were consecutively enrolled into the CRPC registry after obtaining institutional review board approval from each institution.

MATERIAL AND METHODS

Patient Population

According to the CRPC data, the pathologic slides of the biopsy and RP specimens were available for 1650 patients with Stage cT1c-T3 disease and no preoperative therapy at 48 institutions that agreed to send the pathologic slides for central review. After excluding 21 patients (1.3%) without cancer cells in the biopsy specimens by central review, 1629 patients constituted the final cohort for the present study. In all patients, the diagnosis was made by systemic biopsy (≥6 cores). A total of 365 patients (22.4%) had only 6 cores taken at biopsy; 760 patients (46.7%) had ≥10 cores on taken at biopsy. The median number of biopsy cores taken was 8 (range 6-33). All RP specimens were processed using the whole mount technique at each institution. Preoperative information, including the serum prostate-specific antigen levels, and the original pathologic reports were available for all patients. The clinical stage was determined from the digital rectal examination findings and assigned according to the 2002 American Joint Committee on Cancer staging system.

Pathologic Analysis

The biopsy GS of each patient's original pathologic report was recorded as the local review biopsy GS. All pathologic slides and the biopsy and RP specimens were sent to, and reviewed by, 2 dedicated uropathologists (K.K. and T.S.) who were unaware of the original pathologic reports of each patient. In addition, the 2 uropathologists were unaware of the results from the biopsy specimens of each patient when reviewing the matching RP specimens, because the review of the RP specimens was separated from the review of the biopsy specimens. The Gleason pattern was assigned as the central review biopsy and RP GS according to the modified Gleason grading system using the ISUP consensus.[16] The GS was categorized into 5 groups (2-4, 5-6, 3 + 4, 4 + 3, and 8-10). For the biopsy specimens with multiple positive cores, a global GS was recorded, because the GS of each core was not available in most (>95%) of the original pathologic reports. For central review, the reporting rules for a secondary pattern occupying <5% and a tertiary pattern conformed to the ISUP consensus.[16] For the RP specimens, the global GS considering the entire tumor within the prostate as 1 lesion was recorded. A tertiary Gleason pattern in the RP specimens was not reflected as a primary or secondary pattern on the final RP GS. The presence of high-grade Gleason patterns (4 or 5), including tertiary patterns, in the RP specimens was recorded.

Statistical Analysis

Spearman's rank correlation coefficients for the GS in the biopsy and RP specimens were generated. The chi-square test was used for the comparison of the exact GS concordance rate between the local and central pathologic review and for the sensitivity, specificity, positive predictive value, and negative predictive value for the depiction of high-grade Gleason patterns. Two-sided *P* values were calculated; the significance level was set at 5%. All statistical analyses were performed using the Statistical Package for Social Sciences, version 17.0 (SPSS, Chicago, IL).

RESULTS

Clinical Characteristics

For the 1629 patients whose CRPC data were analyzed, the median age was 65 years (range 44-84), and the median prostate-specific antigen level 8.0 ng/mL (range 0.5-85.9). Of the 1629, patients, 1058 (64.9%) had Stage cT1c disease.

GS in Biopsy and RP Specimens

By central review, no patient (0%) had GS 2-4 disease in the biopsy specimens compared with 107 patients (6.6%) who had GS 2-4 by local review. Of the 107 patients with local review biopsy GS of 2-4, central review found a biopsy GS of 5-6, 3 + 4, 4 + 3, and 8-10 in 66 (61.7%), 35 (32.7%), 4 (3.7%), and 2 (1.9%), respectively. In the other GS groups, the distribution of the central biopsy GS was 5-6 in 545 (33.5%), 3 + 4 in 602 (37.0%), 4 + 3 in 257 (15.8%), and 8-10 in 225 (13.8%). The corresponding distribution by local review for the biopsy GS was 687 (42.2%), 379 (23.3%), 192 (11.8%), and 264 (16.2%; Table 1). Of the patients with a biopsy GS of 5-6, 3 (0.6%) of 545 by central review and 138 (20.1%) of 602 by local review had GS 5. Exact concordance between the local and central biopsy GS was observed for 841 patients (51.6%). The undergrading and overgrading rate for local review was 32.6% and 15.8%, respectively. Spearman's rank correlation coefficient for local biopsy GS and central biopsy GS was 0.607. The central review RP GS distribution for GS 5-6, 3 + 4, 4 + 3, and 8-10 was 423 (26.0%), 675 (41.4%), 363 (22.3%), and 168 (10.3%), respectively.

GS Correlation Between Biopsy and RP Specimens

Table 2 lists the correlation between the local review biopsy GS and central review RP GS. The exact concordance rate and the concordance rate within ±1 GS group was 41.3% (672 of 1629) and 81.7% (1331 of 1629), respectively. The undergrading and overgrading rate for

Table 1. Biopsy Gleason score correlation between local review and central review

Local Review Biopsy GS	Central Review Biopsy GS (n)					Exact Concordance Rate (%)	Local Review	
	2-4	5-6	3 + 4	4 + 3	8-10		Undergrading Rate (%)	Overgrading Rate (%)
2-4 (n = 107)	0	66	35	4	2	0.0	100.0	0.0
5-6 (n = 687)	0	388	233	50	16	56.5	43.5	0.0
3 + 4 (n = 379)	0	64	225	62	28	59.4	23.7	16.9
4 + 3 (n = 192)	0	13	60	84	35	43.8	18.2	38.0
8-10 (n = 264)	0	14	49	57	144	54.5	0	45.5
Total (n = 1629)	0	545	602	257	225	51.6	32.6	15.8

GS, Gleason score.

Table 2. Gleason score correlation between local review biopsy and central review prostatectomy specimens

Local Review Biopsy GS	Central Review RP GS (n)					Exact Concordance Rate (%)	Undergrading Rate in Biopsy (%)	Overgrading Rate in Biopsy (%)
	2-4	5-6	3 + 4	4 + 3	8-10			
2-4 (n = 107)	0	42	48	14	3	0.0	100.0	0.0
5-6 (n = 687)	0	282	286	97	22	41.0	59.0	0.0
3 + 4 (n = 379)	0	73	204	86	16	53.8	26.9	19.3
4 + 3 (n = 192)	0	16	65	85	26	44.3	13.5	42.2
8-10 (n = 264)	0	10	72	81	101	38.3	0.0	61.7
Total (n = 1629)	0	423	675	363	168	41.3	39.3	19.5

RP, radical prostatectomy; GS, Gleason score.

Table 3. Gleason score correlation between central review biopsy and prostatectomy specimens

Central Review Biopsy GS	Central Review RP GS (n)					Exact Concordance Rate (%)	Undergrading Rate in Biopsy (%)	Overgrading Rate in Biopsy (%)
	2-4	5-6	3 + 4	4 + 3	8-10			
2-4 (n = 107)	0	0	0	0	0	—	—	—
5-6 (n = 687)	0	335	173	27	10	61.5	38.5	0.0
3 + 4 (n = 379)	0	83	391	113	15	65.0	21.3	13.8
4 + 3 (n = 192)	0	2	76	160	19	62.3	7.4	30.4
8-10 (n = 264)	0	3	35	63	124	55.1	0.0	44.9
Total (n = 1629)	0	423	675	363	168	62.0	21.9	16.1

Abbreviations as in Table 2.

the biopsy specimens was 39.3% and 19.5%, respectively. Of the 107 patients with a biopsy GS of 2-4, all had an RP GS of ≥5-6, including 65 patients (60.1%) with a RP GS of ≥7. Spearman's rank correlation coefficient for the local biopsy GS and central RP GS was 0.459.

Table 3 lists the correlation between the central biopsy GS and the central RP GS. The exact concordance rate and the concordance rate within ±1 GS group was 62.0% (1010 of 1629) and 94.4% (1537 of 1629), respectively. The undergrading and overgrading rate for the biopsy specimens was 21.9% and 16.1%, respectively. Central review had a significantly greater exact concordance and lower undergrading and overgrading rates than did the local review ($P < .05$ for all). Spearman's rank correlation coefficient for central biopsy GS and central RP GS was 0.687. In each GS group, the central review biopsy GS (GS 5-6, 61.5%; 3 + 4, 65.0%; 4 + 3, 62.3%; and 8-10, 65.1%) had a significantly greater exact concordance rate than did the local review biopsy GS (GS 5-6, 41.0%; 3 + 4, 53.8%; 4 + 3, 44.3%; and 8-10, 38.3%; $P < .05$ for all). In each prostate-specific antigen group, the central review biopsy GS (<4.0 ng/mL, 56.6%; 4.1-10 ng/mL, 64.1%; 10.1-20 ng/mL, 60.7%; and >20 ng/mL, 56.4%) had a significantly greater exact concordance rate than the local review biopsy GS (<4.0 ng/mL, 56.6%; 4.1-10 ng/mL, 64.1%; 10.1-20 ng/mL, 60.7%; and >20 ng/mL, 56.4%; $P < .05$ for all).

High-Grade Gleason Patterns (4 or 5)

The number of patients with Gleason pattern 4 or 5 in the biopsy GS as a primary or secondary pattern was 846 (51.9%) in the local review and 1084 (66.6%) in the central review.

Overall, 1371 patients (84.2%) had Gleason pattern 4 or 5 on RP specimens on the central pathology review of the RP specimens. Of these, 1206 (88.0%) had Gleason pattern 4 or 5 as the primary or secondary pattern. The remaining 165 (12.0%) with RP GS 3 + 3 had a high-grade Gleason pattern of <5% on the RP specimens.

Table 4 lists the correlation of high-grade Gleason patterns between the biopsy GS and RP specimens. The central review GS had significantly greater sensitivity and a significantly greater positive and negative predictive values ($P < .05$ for all).

Table 4. High-grade Gleason patterns (4 or 5) in biopsy Gleason score and prostatectomy specimens

Review	High-Grade GP in Biopsy GS	High-Grade GP in RP Specimens (n)		Sensitivity	Specificity	PPV	NPV
		Positive	Negative				
Local	Positive	797	49	0.581	0.810	0.942	0.140
	Negative	574	206				
Central	Positive	1052	32	0.767	0.876	0.970	0.415
	Negative	319	226				
P value				<.001	.053	.003	<.001

GP, Gleason pattern; NPV, negative predictive value; PPV, positive predictive value; other abbreviations as in Table 2.

COMMENT

In the pretreatment setting for prostate cancer in which clinicians can only use the biopsy information for histologic grade, a more accurate GS correlation between the biopsy and RP specimens must result in more precise evaluation of the disease, regardless of the treatment type planned. However, studies investigating the GS correlation between the biopsy and RP specimens have shown considerable discrepancy—especially of undergrading in biopsy specimens.[4-11] Although the number of patients involved in these studies has varied from 28 to 1455, very few men had high-grade biopsy GSs.[8-10] The present study included the largest number of patients with high-grade biopsy GS (local review 264, central review 168) for investigating the correlation of the GS between the biopsy and RP specimens. Pathology error and sampling error are thought to be the main reasons for the discrepancy.

Steinberg et al[11] previously reported that pathologists at an academic center had a better GS correlation than those at community sites. According to their recent study of 1455 patients, Fine and Epstein[6] reported that the exact GS concordance rate was improved in both community sites (from 34% to 70%) and an academic center (multiple pathologists; from 58% to 76%) compared with the rate in their older study. The effects of education and pathologists' efforts in the United States might have contributed to this improvement.

The present study had some differences from that conducted by Fine and Epstein.[6] First, each Gleason pattern was assigned according to the ISUP consensus, which was published in 2005 after their study period (2002-2003). Second, we used the global GS, considering the entire tumor within the prostate as 1 lesion for both the biopsy and the RP specimens because the GS of each core was not available in most (>95%) of the original pathologic reports. The use of the global GS should be considered a weakness of the present study. In the study by Fine and Epstein,[6] the RP GS was recorded from the dominant tumor or highest grade tumor. However, it was not clearly reported whether the global or highest core GS had been used for the biopsy specimens. Although almost all preoperative nomograms have used the highest core grade of the given case when multiple cores with different GSs are present, and urologists have tended to use the greatest GS to determine their treatment plan, some clinicians might use the global GS. ISUP did not actually specify that the highest core GS should be used for the biopsy GS in each case.[2,16,17] Third, the present study included significantly more patients with greater biopsy and RP specimens than the previous study. In the present study, 67% of the biopsy and 74% of the RP specimens had a GS of ≥7 compared with the previous 26% and 23%, respectively.[6,11] This might have resulted from patient selection bias and ethnic differences in the patients with prostate cancer, because the present cohort of patients underwent RP at academic or community institutions in Japan.[18] In addition to the differences in the distribution of GS, the division of GS 7 into 3 + 4 and 4 + 3 might explain the relatively low exact concordance rate in our study. When GS 3 + 4 and 4 + 3 were combined as 1 entity, the exact concordance rate was high (73.6%) in the present study. However, a GS of 3 + 4 and that of 4 + 3 have different biologic behavior and should not be combined into 1 category.[19]

Reflecting contemporary changes regarding prostate cancer and the Gleason grading system, the ISUP proposed a modified Gleason grading system in 2005.[16] The ISUP consensus has been reported to minimize biopsy undergrading and improve the GS correlation compared with the previous system.[12] In the present study, including patients who underwent RP from 1997 to 2005, biopsy GS 2-4 was originally diagnosed at each institution in 14.6% of all patients compared with 1.6% in another study.[6] ISUP recommended that a GS 2-4 should rarely, if ever, be considered, because of the poor correlation with the RP GS. Most expert uropathologists would not have assigned a GS of 2-4 even before the ISUP consensus.[20] In our study, all locally reviewed biopsy GS 2-4 specimens were upgraded by the central review and 61% actually had a RP GS of ≥7, including 3 patients with a RP specimen GS of 8-10. In addition, no RP specimens in the present study was graded with a GS of 2-4. For the GS categories other than 2-4, we also showed that central review using the ISUP consensus gave a more accurate GS correlation than local review, including biopsy GS 8-10. However, the exact concordance rate was far from perfect (100%) and was less satisfactory even when a central review using the ISUP consensus was done. The actual GS of each patient can be apparent only after RP has been performed. We believe this is an advantage for RP compared with other

treatment modalities that offer patient surveillance and adjuvant treatment according to the biopsy GS only.

High-grade Gleason patterns, either a primary/secondary pattern or a tertiary pattern, in RP specimens have been related to a poor outcome.[13-15] We have demonstrated that the central review biopsy GS using the ISUP consensus is superior to the local review biopsy GS in terms of predicting high-grade Gleason patterns in the RP specimens. It has been reported that the highest core GS has the largest effect on a significant upward shift of the biopsy GS among the reporting rules of the ISUP consensus.[21] Because we used a global biopsy GS for the central review, the difference in the interpretation of each Gleason pattern between the local review and central review might explain our results for high-grade Gleason patterns.

CONCLUSIONS

This is the first study to investigate the significance of dedicated pathologic reassessment using the ISUP consensus for biopsy and RP specimens from academic and community practices. Central pathologic review resulted in a more accurate GS correlation and prediction of high-grade Gleason patterns. We believe that more educational effort is needed for both pathology and urology communities to disseminate the ISUP consensus. We recommend central pathology review by dedicated uropathologists for a study of prostate biopsy and RP specimens from patients at multiple institutions, although the central review will cost more and is time-consuming. We should carefully interpret multicenter study data that have not included a central review. In addition, the exact concordance rate was far from perfect (100%) and was not satisfactory even when a central review using the ISUP consensus was done. Also, the actual GS of each patient can be apparent only when RP has been performed.

Acknowledgments. To all the patients for their participation in the present study; to the doctors, pathologists, and other support staff at the institutes of the CRPC; to the CRPC facilitators: Hideyuki Akaza, University of Tsukuba; Yoichi Arai, Tohoku University; Shiro Baba, Kitasato University; Shin Egawa, Jikei University; Yoshihiko Hirao, Nara Medical University; Susumu Kagawa, University of Tokushima; Hiroshi Kanetake, Nagasaki University; Masaru Murai, Keio University; Yoshiaki Nose, Kyushu University; Osamu Ogawa, Kyoto University; Makoto Ohori, Tokyo Medical University; Shinichi Ohshima, National Center for Geriatrics and Gerontology; Akihiko Okuyama, Osaka University; Taiji Tsukamoto, Sapporo Medical University; Michiyuki Usami, Osaka Medical Center for Cancer and Cardiovascular Diseases; Hidetoshi Yamanaka, Kurosawa Hospital; Ken Goto, Kyushu University; and Hirofumi Koga, Kyushu University.

References

1. Gleason DF. Classification of prostatic carcinomas. *Cancer Chemother Rep.* 1966;50:125-128.
2. Kattan MW, Eastham JA, Stapleton AM, et al. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *J Natl Cancer Inst.* 1998;90:766-771.
3. Kattan MW, Wheeler TM, Scardino PT. Postoperative nomogram for disease recurrence after radical prostatectomy for prostate cancer. *J Clin Oncol.* 1999;17:1499-1507.
4. Bostwick DG. Gleason grading of prostatic needle biopsies: Correlation with grade in 316 matched prostatectomies. *Am J Surg Pathol.* 1994;18:796-803.
5. Djavan B, Kadesky K, Klopukh B, et al. Gleason scores from prostate biopsies obtained with 18-gauge biopsy needles poorly predict Gleason scores of radical prostatectomy specimens. *Eur Urol.* 1998;33:261-270.
6. Fine SW, Epstein JI. A contemporary study correlating prostate needle biopsy and radical prostatectomy Gleason score. *J Urol.* 2008;179:1335-1338.
7. King CR. Patterns of prostate cancer biopsy grading: trends and clinical implications. *Int J Cancer.* 2000;90:305-311.
8. Lattouf JB, Saad F. Gleason score on biopsy: is it reliable for predicting the final grade on pathology? *BJU Int.* 2002;90:694-698.
9. Paulson DF. Impact of radical prostatectomy in the management of clinically localized disease. *J Urol.* 1994;152:1826-1830.
10. San FIF, DeWolf WC, Rosen S, et al. Extended prostate needle biopsy improves concordance of Gleason grading between prostate needle biopsy and radical prostatectomy. *J Urol.* 2003;169:136-140.
11. Steinberg DM, Sauvageot J, Piantadosi S, et al. Correlation of prostate needle biopsy and radical prostatectomy Gleason grade in academic and community settings. *Am J Surg Pathol.* 1997;21:566-576.
12. Helpap B, Egevad L. The significance of modified Gleason grading of prostatic carcinoma in biopsy and radical prostatectomy specimens. *Virchows Arch.* 2006;449:622-627.
13. Nanda A, Chen MH, Renshaw AA, et al. Gleason pattern 5 prostate cancer: further stratification of patients with high-risk disease and implications for future randomized trials. *Int J Radiat Oncol Biol Phys.* 2009;74:1419-1423.
14. Pan CC, Potter SR, Partin AW, et al. The prognostic significance of tertiary Gleason patterns of higher grade in radical prostatectomy specimens: a proposal to modify the Gleason grading system. *Am J Surg Pathol.* 2000;24:563-569.
15. Whittemore DE, Hick EJ, Carter MR, et al. Significance of tertiary Gleason pattern 5 in Gleason score 7 radical prostatectomy specimens. *J Urol.* 2008;179:516-522.
16. Epstein JI, Allsbrook WC Jr, Amin MB, et al. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma. *Am J Surg Pathol.* 2005;29:1228-1242.
17. Rubin MA, Bismar TA, Curtis S, et al. Prostate needle biopsy reporting: how are the surgical members of the Society of Urologic Oncology using pathology reports to guide treatment of prostate cancer patients? *Am J Surg Pathol.* 2004;28:946-952.
18. Naito S, Kuroiwa K, Kinukawa N, et al. Validation of Partin tables and development of a preoperative nomogram for Japanese patients with clinically localized prostate cancer using 2005 International Society of Urological Pathology consensus on Gleason grading: data from the Clinicopathological Research Group for Localized Prostate Cancer. *J Urol.* 2008;180:904-909.
19. Makarov DV, Sanderson H, Partin AW, et al. Gleason score 7 prostate cancer on needle biopsy: is the prognostic difference in Gleason scores. 4 + 3 and 3 + 4 independent of the number of involved cores? *J Urol.* 2002;167:2440-2442.
20. Epstein JI. Gleason score 2-4 adenocarcinoma of the prostate on needle biopsy: a diagnosis that should not be made. *Am J Surg Pathol.* 2000;24:477-478.
21. Kuroiwa K, Uchino H, Yokomizo A, et al. Impact of reporting rules of biopsy Gleason score for prostate cancer. *J Clin Pathol.* 2009;62:260-263.