Italy-Japan agreement and discrepancies in diagnosis of superficial gastric lesions

Carla Vindigni[1], Mario Marini[2], Gabriele Cevenini[3], Maria Raffaella Ambrosio[4], Monica Onorati[5], Giorgio Frosini[6], Takuji Gotoda[7], Hirokazu Taniguchi[8], Piero Tosi[9]

[1]Department of Human Pathology and Oncology, Division of Pathological Anatomy and Histopathology, University of Siena, Italy, [2]Gastroenterology and Gastrointestinal Endoscopy Unit, University Hospital, Siena, [3]Surgery and Bioengineering Department, University of Siena, [4]Department of Human Pathology and Oncology, Division of Pathological Anatomy and Histopathology, University of Siena, [5]Department of Human Pathology and Oncology, Division of Pathological Anatomy and Histopathology, University of Siena, [6]Gastroenterology and Gastrointestinal Endoscopy Unit, University Hospital, Siena, [7]Gastrointestinal Endoscopy Unit, National Cancer Center Hospital, Tokyo, Japan [8]Clinical Laboratory Division, National Cancer Center Hospital, Tokyo, [9]Department of Human Pathology and Oncology, Division of Pathological Anatomy and Histopathology, University of Siena

TABLE OF CONTENTS

1. ABSTRACT

The agreement between Italian and Japanese endoscopists and pathologists on endoscopic and histopathological diagnoses of superficial gastric lesions is verified with the use of Paris and Vienna classifications. The correlations between Paris endoscopic types and Vienna histopathological categories is high in both the independent Italian and Japanese evaluations. However, the agreement between Italian and Japanese endoscopists is moderate due to the difficult evaluation of the height of the lesions, in particular when they are mixed. The agreement on the size of the lesions is fairly good. The probability of the same allocation to the Vienna categories of a single case is 87%, disagreements remaining in dysplasia grading, between dysplasia, not only high-grade but also low-grade, and *in situ* carcinoma, and on cancer invasion of the lamina propria. The results indicate that use of the Paris and Vienna classifications has reduced the discrepancies between Western and Japanese endoscopists and pathologists in the diagnosis of these lesions.

2. INTRODUCTION

Between 2000 and 2002 two milestones were reached in the diagnosis of superficial gastric lesions: the Paris endoscopic classification (1) and the Vienna histopathological classification (2). Both aimed to reduce the interobserver discrepancies and in particular those between Western and Japanese endoscopists and pathologists in the diagnosis of lesions ranging from reactive epithelium to low-grade and high-grade dysplasia, and to early cancer in its non invasive or invasive patterns (3-10).

The two classifications, on which well- known and influential endoscopists and pathologists reached agreement, were expected to increase the reproducibility of diagnoses also among less well-known and influential endoscopists and pathologists. The aim of this study was to contribute to verifying whether use of the Paris and Vienna classifications have helped to reduce the diagnostic discrepancies and enhance interobserver reproducibility between Western (in this case Italian) and Japanese

Superficial gastric lesions

Table 1. Paris endoscopic classification

Protruding	
Peduncolated	0-Ip
Sessile	0-Is
Non-protruding and non-excavated	
Slightly elevated	0-IIa
Completely flat	0-IIb
Slightly depressed	0-IIc
Elevated and depressed types	0-IIc+IIa; 0-IIa+IIc
Excavated	
Ulcer	0-III
Excavated and depressed types	0-IIc+III; 0-III+IIc

Table 2. Vienna histopathological classification

Category	Feature
Category **1**	Negative for neoplasia/dysplasia
Category **2**	Indefinite for neoplasia/dysplasia
Category **3**	Non invasive low-grade neoplasia (low grade adenoma/dysplasia)
Category **4**	Non invasive high-grade neoplasia
Category **4.1**	High grade adenoma/dysplasia
Category **4.2**	Non invasive carcinoma (carcinoma *in situ*)
Category **4.3**	Suspicion of invasive carcinoma
Category **5**	Invasive neoplasia
Category **5.1**	Intramucosal carcinoma
Category **5.2**	Submucosal carcinoma or beyond

endoscopists and pathologists. With this aim in mind, we attempt to: 1) check on the agreement on the Paris endoscopic classification of superficial gastric lesions independently diagnosed by Italian and Japanese endoscopists; 2) correlate Paris endoscopic types with Vienna histopathological categories, as they resulted from diagnoses made independently by Italian and Japanese endoscopists and pathologists; 3) establish the interobserver agreement on histopathological diagnoses made independently by Italian and Japanese pathologists.

The cases studied were 48 and include reactive lesions, low- and high-grade dysplasias and early cancers.

3. MATERIALS AND METHODS

3.1. Endoscopy

Endoscopic examination was performed at Gastroenterology and Gastrointestinal Endoscopy Unit of Siena University Hospital. In all cases chromoendoscopy with indigo carmine solution (0.5-1.0 %) was used to obtain a precise analysis and classification of the lesions. The elevation or depression of superficial lesions was evaluated with a millimeter scale for the columnar epithelium according to guidelines of Japanese Gastric Cancer Association (11). All endoscopic images, at least 4 per case, were stored electronically for later evaluation using Image Bank Software (United Medical Software-Italy). Endoscopic images were reviewed and the diagnoses confirmed in photographs by the same endoscopists who had made the original diagnoses, according to Paris classification (Table 1). These photographs, without any diagnostic indications, were then sent to Japanese endoscopists working at the National Cancer Center in Tokyo, who subtyped them and sent their diagnoses back to the endoscopists in Siena.

A complete intra- and interobserver agreement was reached among Italian endoscopists and among Japanese endoscopists.

3.2. Histopathology

A similar procedure was followed for the histopathological diagnoses: previous diagnoses made by the Italian pathologists at Deparment of Human Pathology and Oncology, University of Siena, according to Vienna classification (Table 2), on at least 4 samples per case, were confirmed by the same pathologists on the same slides. These slides were then sent to Japanese pathologists at National Cancer Center in Tokyo, without any comments; the Japanese pathologists sent their diagnoses back to the pathologists in Siena. Reproducibility was checked and totally reached among histopathologists in Italy and in Japan respectively.

Endoscopic diagnoses from Italy and Japan were evaluated for agreement-disagreement. Japanese endoscopic diagnoses were matched with Japanese histopathological diagnoses, and Italian endoscopic diagnoses were matched with Italian histopathological diagnoses. Finally, Japanese histopathological diagnoses were matched with Italian histopathological diagnoses. Data were submitted to statistical analysis.

3.3. Statistics

Descriptive statistics was computed, including frequency count, minimum, maximum, mean and standard deviation for numerical variables (age and size), frequency count and percentage for qualitative variables (site, sex, Paris type, Vienna category). The Kolmogorov-Smirnov (KS) test was applied to verify the normality of distribution of the quantitative variables: once normality had been assessed, parametric tests were used to compare different data groups, as the Student-t test and the analysis of variance (ANOVA) with Bonferroni pairwise comparison. For non-normal data, the non-parametric tests of Mann-Whitney and Kruskal-Wallis were used. The chi-square test applied to contingency tables was performed to compare frequency distribution of Paris and Vienna classification for gender and site. The same test was used to analyze the correlations between Paris and Vienna

Table 3. Site of the lesions in the stomach

Site	Number	%
Body	25	52
Antrum	20	42
Cardias	1	2
Stump	2	4
Total cases	**48**	**100**

Table 4. Size of the lesions according to Italian and Japanese endoscopists

	Italian data		Japanese data	
SIZE	**NUMBER**	**%**	**NUMBER**	**%**
<1	4	8	10	21
1-2	32	67	24	50
>2	12	25	14	29
Total cases	**48**	**100**	**48**	**100**

Table 5. Distribution of Paris types and Vienna categories according to Italian and Japanese endoscopists

a. Distribution of Paris types				
	Italian data		Japanese data	
PARIS TYPE	**NUMBER**	**%**	**NUMBER**	**%**
I (Ip,Is)	15	31	9	19
IIa	9	19	17	35
IIc; IIa+IIc;IIb+IIc	21	44	21	44
III	3	6	1	2
Total cases	**48**	**100**	**48**	**100**
b. Distribution of Vienna categories				
	Italian data		Japanese data	
VIENNA CATEGORY	**NUMBER**	**%**	**NUMBER**	**%**
1	10	21	11	23
2	1	2	0	0
3	16	33	11	23
4.1	3	6	0	0
4.2	0	0	11	23
4.3	1	2	0	0
5.1	10	21	7	15
5.2	7	15	8	16
Total cases	**48**	**100**	**48**	**100**

classification. In order to make this correlation, the Paris endoscopic types were grouped as follows: O-I (Is, Ip); O-IIa; O-IIc (IIc, IIa+IIc, IIb+IIc); O-III.

All these tests were separately applied to the Italian and Japanese data (12). The agreement between them was analyzed using the Cohen kappa coefficient. The kappa coefficient was interpreted using the Landis and Koch scale: less than < 0.20 slight agreement; kappa 0.21-0.40 fair agreement; kappa 0.41-0.60 moderate agreement; kappa 0.61-0.80 good agreement; kappa 0.81-1-00 very good agreement (13). For ordinal classification values, the Kendall tau coefficient was also evaluated. The Kendall tau was directly interpreted as the probability of obtaining concordant or discordant pairs (14). Pearson (negative KS test) or Spearman (positive KS test) correlation coefficients (r) were used to analyze the quantitative variable size (12).

A significance level of 95% (p= 0.05) was chosen for all statistical analyses (computed using the SPSS 10 software). The inter- and intraobserver variation was assessed in all endoscopic and histopathological diagnoses. All diagnoses were performed blind by two independent observers. The Student's T test was applied to determine any difference.

4. RESULTS AND DISCUSSION

4.1. Results

The mean age of the patients was 71 (range 46-86); the males were 26 and the females 22. Data on the site of the lesions in the stomach are summarized in Table 3. The size of the lesions and endoscopic types and histopathological categories according to Italian and Japanese endoscopists and pathologists are reported respectively in Table 4 and 5 (a, b).

No correlations were found between allocation of cases to the Paris types and Vienna categories and age and sex of the patients, site in the stomach and size of the lesions.

There was a statistically significant correlation between Paris types and Vienna categories: p= 0.019 for Italian results, p= 0.011 for Japanese results. The majority (75% for the Italians, 56% for the Japaneses) of Vienna category 1 cases were endoscopically type I; 35% for the Italians and 27% for the Japaneses of Vienna category 3 cases were type I; 35% for the Italians and 54% for the Japaneses of Vienna category 3 were IIa; 100% for the Italians and 46% for the Japaneses of Vienna category 4 were type IIc; 76% for the Italians and 67% for the Japaneses of category 5 cases were type IIc.

The agreement between Italian and Japanese endoscopists was moderate on the basis of Landis and

Superficial gastric lesions

Table 6. Agreement between Italian and Japanese endoscopists on Paris types (kappa=0.54)

		Italian endoscopists				
		I	IIa	IIc	III	Total cases
Japanese endoscopists	I	9	0	0	0	9
	IIa	5	7	5	0	17
	IIc	1	2	16	2	21
	III	0	0	0	1	1
	Total cases	15	9	21	3	48

Table 7. Agreement between Italian and Japanese pathologists on Vienna categories, including subcategories (Kappa = 0.57; tau = 0.87)

		Italian pathologists							
		1	2	3	4.1	4.2	4.3	5.1	5.2	Total cases
Japanese pathologists	1	10	1	0	0	0	0	0	0	11
	2	0	0	0	0	0	0	0	0	0
	3	0	0	11	0	0	0	0	0	11
	4.1	0	0	0	0	0	0	0	0	0
	4.2	0	0	5	2	0	0	4	0	11
	4.3	0	0	0	0	0	0	0	0	0
	5.1	0	0	0	1	0	1	4	1	7
	5.2	0	0	0	0	0	0	2	6	8
	Total cases	10	1	16	3	0	1	10	7	48

Table 8. Agreement between Italian and Japanese pathologists on Vienna categories, excluding subcategories (kappa = 0.69; tau = 0.88)

		Italian pathologists					
		1	2	3	4	5	Total cases
Japanese pathologists	1	10	1	0	0	0	11
	2	0	0	0	0	0	0
	3	0	0	11	0	0	11
	4	0	0	5	2	4	11
	5	0	0	0	2	13	15
	Total cases	10	1	16	4	17	48

Koch scale (kappa= 0.54): for type I it was 60% because 5 out of 15 cases (33%) classified by the Italian endoscopists as type I were classified by the Japanese endoscopists as type IIa, 1 type I (7%) became IIc; 7 out of 9 cases (78%) were classified as type IIa both by the Italians and the Japaneses, 2 cases (22%) IIa for the Italians became IIc; 2 out of the 3 (67%) type III cases for the Italians were classified as IIc by the Japaneses; there was agreement on 16 cases out of the 21 (76%) classified as IIc by the Italians since 5 cases (24%) classified as IIc by the Italians became IIa (Table 6).

Regarding the size, a statistically significant agreement has been found (r = 0.74).

The agreement between the Italian and the Japanese pathologists was moderate on the basis of Landis and Koch scale (kappa= 0.57) or good according to the Kendall tau coefficient (tau= 0.87) (Table 7). They agreed almost completely on negative (category 1) cases, with the exception of 1 case classified as category 2 (indefinite for dysplasia/neoplasia) by the Italians which the Japanese pathologists defined as negative. There was an agreement of 65% in the classification of category 3 (low-grade dysplasia/adenoma) cases, while 5 of the 16 (29%) cases allocated to this category by the Italians were 4.2 (*in situ* carcinomas) for the Japaneses; 2 cases (67%) that were category 4.1 (high-grade dysplasia/adenoma) according to the Italians were 4.2 for the Japaneses, 1 case (33%) 4.1 for the Italians became 5.1 (intramucosal carcinoma) for the Japaneses; 1 case allocated to category 4.3 (suspicious for invasive carcinoma) by the Italians was allocated to

category 5.1 by the Japaneses; of the 10 cases classified as category 5.1 by the Italians there was agreement in 4 cases (40%), while 4 cases (40%) became 4.2 and 2 cases (20%) were placed in category 5.2 (submucosal carcinoma) by the Japanese pathologists. In Table 8 the agreement (kappa = 0.69; tau= 0.88) has been evaluated once the subcategories of category 4 and 5 had been eliminated.

4.2. Discussion

There has been much disagreement between Western and Japanese pathologists over the years regarding the diagnosis of superficial gastric lesions, with a lack of interobserver reproducibility in the differential diagnosis between reactive and dysplastic changes, between high-grade dysplasia and intramucosal carcinoma, and in dysplasia grading (6-10). Nevertheless, it is imperative that clinicians obtain a definite diagnosis of low-grade or high-grade dysplasia, as low-grade implies endoscopic surveillance, while high-grade implies resection of the lesion. The Vienna consensus conference offered a satisfactory histopathological agreement among experts, and, even if it admitted the categories "indefinite" and "suspicious", it increased the chance of agreement among pathologists (2). The same can be said for the Paris agreement on the endoscopic classification of superficial neoplastic lesions designed to explore the clinical relevance of the Japanese endoscopic classification (1). The term "superficial" at endoscopy has been clarified: the depth of penetration is not more than into the submucosa. The agreement reached by this group represents the general framework for the endoscopic classification of superficial

lesions and the lesions of the Vienna classification are allocated to the endoscopic subtypes of the Paris classification.

The Vienna and Paris consensus conferences were undoubtedly two important steps towards improving the reproducibility of the diagnosis of gastric lesions which are superficial at endoscopy. It remains to be verified whether substantial interobserver reproducibility has been achieved in daily practice, in particular between endoscopists and pathologists working in a Western country and their counterparts in Japan. In fact, before Vienna it had been assessed that Japanese pathologists, basing their diagnosis on cytological and architectural changes, interpreted as carcinoma lesions that Western pathologists diagnosed as dysplasia due to the absence of invasion into the lamina propria (6). Other points of disagreement regarded borderline lesions between reactive and dysplastic changes and also in the Vienna classification a category is "indefinite for dysplasia".

In the present study, we have analyzed the level of agreement/disagreement between Italian and Japanese endoscopists and pathologists in a series of 48 superficial gastric lesions ranging from reactive (negative for dysplasia) changes to low- and high-grade dysplasia and early cancer, using the Paris and Vienna classifications. The correlation between Paris types and Vienna categories was very high both in the Italian and Japanese evaluations, and this is encouraging if one considers how important such correlation is for patient management. However, the agreement between the endoscopic classification of these cases made independently by the Japanese and Italian endoscopists was moderate (kappa= 0.54). The most frequent discrepancies were between type I and IIa , due to the difficulty of the evaluation of the height of the lesions. Other discrepancies consisted in the different allocation of some cases to type IIa or IIc, and of some other cases to IIc or III. This may be due to the fact that these cases were mixed and one or another of their components could be considered prevalent.

The agreement on the size of the lesions was fairly good (r=0.74), since no statistically significant differences were found (p=0.078).

The total agreement between Italian and Japanese pathologists was moderate (kappa= 0.57). However, the probability of the same allocation by Italian and Japanese pathologists of a single case was 87% (tau= 0.87). In fact, the entity of disagreement depends on the distance between the differently diagnosed categories. A complete agreement was registered in the diagnosis of negative (reactive) lesions and the only case diagnosed as "indefinite" by the Italian pathologists was included in the negative cases by the Japanese pathologists. Since "indefinite" changes have been included in the Vienna classification due to the recognized inability of even experienced pathologists to assess the presence or absence of dysplasia in some lesions with certainty, we can say that after Vienna borderline lesions between reactive and dysplastic changes are more easily classified as reactive or

dysplastic. Major disagreements were registered in the allocation of dysplasia to Vienna category 3 or 4. This is not new or surprising since in dysplasia grading Western pathologists tend to give importance to nuclear stratification while Japanese pathologists give more importance to nuclear pleomorphism and size and nucleolar prominence (2, 15).

Discrepancies remain between Italian and Japanese pathologists, at least in the series of cases we dealt with, regarding the differential diagnosis between dysplasia and carcinoma, not only between high-grade dysplasia and *in situ* carcinoma, but also between low-grade dysplasia and *in situ* carcinoma. *In situ* carcinoma was diagnosed by the Japanese pathologists in a significant number of cases that were diagnosed as high-grade or low-grade dysplasia by the Italian pathologists. This is in agreement with the study of Lauwers *et al* (8). A few disagreements were registered regarding the presence or absence of invasion of the lamina propria due to the difficulty of seeing the epithelial basal membrane in routine sections. Invasion of the lamina propria was seen more frequently by the Italian pathologists (intramucosal carcinoma instead of *in situ* carcinoma). In the revised version of Vienna classification (16), intramucosal carcinoma is allocated to category 4 (4.4) due to use of endoscopic resection for all the mucosal lesions. For this reason, two tables are here presented, the one (table 7) including the subcategories of category 4 and 5 of Vienna classification, the other one (table 8) without these subcategories.

In table 8 one can see that the diagnostic discrepancies are considerably less numerous than in table 7 and kappa coefficient is 0.69 instead of 0.57.

Moreover, the Italian pathologists diagnosed as intramucosal carcinomas 2 cases that the Japanese pathologists defined as submucosal carcinomas: in both these cases it remains disputable whether the cancer was confined to the mucosa and muscolaris mucosae or not, since the submucosa was hardly visible in the biopsies.

5. CONCLUSIONS

This study, although limited by the small number of cases, confirms that use of the Paris and Vienna classifications increases the interobserver reproducibility of endoscopic and histopathological diagnoses of superficial gastric lesions, including the inter-country reproducibility between Japanese and Italian endoscopists and pathologists. The diagnostic discrepancies that remain are at least in part due to the fact that the changes dealt with do not form discrete patterns, but are characterized by the gradual transition typical of continuous lesions, (for instance from high grade dysplasia to *in situ* carcinoma) with an inevitable degree of overlap. Moreover, discrepancies are due to filling in and overlooking, which are common phenomena in the human observation of optical objects: the epithelial basal membrane may be interrupted or not visible due to technical inadequacy; the

height of slightly elevated or slightly depressed lesions may be extremely difficult to define exactly.

6. REFERENCES

1. The Paris Endoscopic Classification of Superficial Neoplastic Lesions: esophagus, stomach, and colon. *Gastrointest Endosc*58 (Suppl 6), S3-S43 (2003)

2. RJ Schlemper, R Riddell, Y Kato, F Borchard, H Cooper, S Dawsey, M Dixon, C Fenoglio-Preiser, J Flejou, K Geboes, T Hattori, T Hirota, M Itabashi, M Iwafuchi, A Iwashita, Y Kim, T Kirchner, M Klimpfinger, M Koike, G Lauwers, K Lewin, G Oberhuber, F Offner, A Price, C Rubio, M Shimizu, T Shimoda, P Sipponen, E Solcia, M Stolte, H Watanabe, and H Yamabe: The Vienna Classification of gastrointestinal epithelial neoplasia. *Gut*47, 251-255 (2000)

3. P Tosi, P Luzi, JPA Baak, C Miracco, C Vindigni, R Lio, P Barbini: Gastric dysplasia: a stereological and morphometrical assessment. *J Pathol*152, 83-94 (1987)

4. P Tosi, JPA Baak, P Luzi, C Miracco, R Lio, P Barbini: Morphometric distinction of low- and high-grade dysplasias in gastric biopsies. *Hum Pathol*20, 839-844 (1989)

5. P Tosi, MI Filipe, P Luzi, C Miracco, R Santopietro, R Lio, V Sforza, P Barbini: Gastric intestinal metaplasia type III cases are classified as low-grade dysplasia on the basis of morphometry. *J Pathol*169, 73-78 (1993)

6. RJ Schlemper, M Itabashi, Y Kato, KJ Lewin, RH Riddell, T Shimoda, P Sipponen, M Stolte, H Watanabe, H Takahashi, R Fujita: Differences in diagnostic criteria for gastric carcinoma between Japanese and Western pathologists. *Lancet*349, 1725-1729 (1997)

7. RH Riddell, M Iwafuchi: Problems arising from Eastern and Western classification systems of gastrointestinal dysplasia and carcinoma: are they resolvable? *Histopathology*33, 197-202 (1998)

8. GY Lauwers, M Shimizu, P Correa, RH Riddell, Y Kato, KJ Lewin, H Yamabe, DG Sheahan, D Lewin, P Sipponen, PS Kubilis, H Watanabe: Evaluation of gastric biopsies for neoplasia: differences between Japanese and Western pathologists. *Am J Surg Pathol*23, 511-518 (1999)

9. RS Jass: Discrepancies between East and West. *Cancer* 88: 969-970 (2000)

10. M Rugge, P Correa, MF Dixon, T Hattori, G Leandro, K Lewin, RH Riddell, P Sipponen, H Watanabe: Gastric dysplasia: the Padova international classification. *Am J Surg Pathol*24: 167-176 (2000)

11. Japanese Gastric Cancer Association: Japanese classification of gastric carcinoma: 2nd English edition. *Gastric Cancer*1: 10-24 (1998)

12. P Armitage, G Berry: Statistical methods in medical research. *Blackwell Scientific Publications Limited*, Oxford (1994)

13. JR Landis, GG Koch: The measurement of observer agreement for categorical data. *Biometrics*33: 159-174 (1977)

14. WJ Conover: Practical Non-Parametric Statistics, 2nd edn. *John Wiley and Sons*, New York (1980)

15. RJ Schlemper, M Itabashi, Y Kato, KJ Lewin, RH Riddell, T Shimoda, P Sipponen, M Stolte, H Watanabe: Differences in the diagnostic criteria used by Japanese and Western pathologists to diagnose colorectal carcinoma. *Cancer*82: 60-69 (1998)

16. RJ Schlemper, Y Kato, M Stolte: Review of histological classifications of gastrointestinal epithelial neoplasia: differences in diagnosis of early carcinomas between Japanese and Western pathologists. *J Gastroenterol*36: 445-456 (2001)

Send correspondence to: Piero Tosi, Department of Human Pathology and Oncology, Division of Pathological Anatomy and Histopathology, University of Siena, Italy, Via delle Scotte, 6, 53100 Siena, Italy, Tel: 39-0577-233233, Fax: 39-0577-233235, E-mail: tosi@unisi.it