Giunchi, Francesca; Jordahl, Kristina; Bollito, Enrico; Colecchia, Maurizio; Patriarca, Carlo; D'Errico, Antonietta; Vasuri, Francesco; Malvi, Deborah; Fornari, Alessandro; Bonetti, Luca Reggiani; Corti, Barbara; Papotti, Mauro; Degiuli, Paolo; Loda, Massimo; Montironi, Rodolfo; Fiorentino, Michelangelo; Rider, Jennifer R.. Interpathologist concordance in the histological diagnosis of focal prostatic atrophy lesions, acute and chronic prostatitis, PIN, and prostate cancer. VIRCHOWS ARCHIV. 470 (6) pp: 1-5. DOI: 10.1007/s00428-017-2123-1

Virchows Archiv

Concordance in the histological diagnosis of Focal Prostatic Atrophy Lesions, Acute and Chronic Prostatitis, PIN and Prostate Cancer Among 15 European Pathologists

--Manuscript Draft--

Manuscript Number:	VIAR-D-17-00009R1
Full Title:	Concordance in the histological diagnosis of Focal Prostatic Atrophy Lesions, Acute and Chronic Prostatitis, PIN and Prostate Cancer Among 15 European Pathologists
Article Type:	Brief Report
Corresponding Author:	Michelangelo; Fiorentino, MD, PhD S.Orsola-Malpighi Hospital Bologna, ITALY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	S.Orsola-Malpighi Hospital
Corresponding Author's Secondary Institution:	
First Author:	Francesca Giunchi
First Author Secondary Information:	
Order of Authors:	Francesca Giunchi
	Kristina Jordahl, PhD
	Enrico Bollito, MD
	Maurizio Colecchia, MD
	Carlo Patriarca, MD
	Antonietta D'Errico, MD
	Francesco Vasuri, MD, PhD
	Deborah Malvi, MD
	Alessandro Fornari, MD
	Luca Reggiani Bonetti, MD
	Barbara Corti, MD
	Mauro Papotti, MD
	Paolo DeGiuli, MD
	Massimo Loda, MD
	Rodolfo Montironi, MD
	Michelangelo; Fiorentino, MD, PhD
	Jennifer Rider
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	Epidemiological and biological evidence indicate a causal relationship between the presence of proliferative atrophic lesions (PAH), the development of PIN and prostate cancer. Inflammatory and atrophic lesions of the prostate are widely underestimated and not generally mentioned in pathology reports. We performed a histopathological concordance study among 15 dedicated and non-dedicated genito-urinary pathologists on 116 histological slides containing prostate atrophic lesions, PIN and cancer.

	We found an overall percent agreement between all possible pairs of reviewers of 80.2% for prostate cancer, 67% for PIN, and 48.7% for any atrophic changes. When we designated a genitourinary pathologist single gold standard, the mean percent agreement increased: 96.6% for prostate cancer, 91.7% for PIN, 71.9% for PAH. According to the raising relevance of PAH in prostate cancer our results on histopathological concordance support the inclusion of at least PAH in the routine pathology reporting of pathological prostate specimens.
Response to Reviewers:	Reviewer #1: Giunchi et al describe the interobserver variability of the diagnosis of atrophic lesions of the prostate, PIN and prostate cancer. Compared to previous studies, the current study uses real glass slides and includes both malignant and benign lesions. This may be of some interest but there are several problems with the paper. We thank the reviewer for the appreciation of our work. 1. Two uropathologists selected the slides but the gold standard seems to be based on the assessment of one of them. Who was used as gold standard? Was his gold standard diagnosis unbiased by the other colleague? Or did the two expert pathologists discuss the cases in an open discussion? Did he first decide the gold standard and then a selection was done by the two of them together or vice versa? If the selection was done first it must be difficult to avoid discussing the diagnosis. This would influence their diagnoses and affect the results of the reproducibility study. As stated in the Methods section the cases were selected by two dedicated uro-pathologists (EB and MF) from two different institutions, Each pathologist was blinded to the other's evaluation and they did not discuss the cases. The gold standard was another GU pathologist (RM) from a third institution who did not participate in case selection. For comparison, we also present data for the overall percentage agreement, which would not be as heavily dependent on the evaluation by a single reviewer. 2. 120 slides were selected including biopsies, RPs, TURPs. How many of each? Why are there only 115 and 116 cases, respectively in Tab 1? Were all cases circulated among 30 pathologists? Did you use recuts? How many sets? The two pathologists who selected the cases asked for the re-cutting of three copies of each selected block. Then they drew circles around 121 areas of interests with a bullpen in 61 histological slides. The three sets of slides were then circulated among the 15 pathologists who accepted to join the study. One pathologist did not review 5 cases leaving 116 cases for the analysis of prostate cancer. Another pathologist did not review 1 case for some types of inflammation/atrophy, and two cases for other types of inflammation/atrophy, leaving either 114 or 115 cases for those analyses. We have now added this information in the Methods section. 3. Which pathologists participated? The panel had 15 members but there are 11 institutions in the author list. This indicates that some work in the same department. Should be specified as it may affect the results. The 15 pathologists belonged to 9 institutions since other centers participated just for the statistical analyses. Slides were circulated blindly and each pathologist provided separate review files to the statisticians who completed the statistical analyses. This was made possible since 6 and 2 pathologists respectively belonged to the two institutions that were in charge of case selection. 4. There should be a Table where the gold standard diagnoses are compared to the results of the other pathologists. I.e. for each diagnostic entity the number of gold standard diagnoses (i.e. the total number of cases) should be given and then the summarized results of the other panelists. 5. There is no information about how the panelists voted, only percentages of agreement and kappas. 6. How was the performance of individual pathologists? The 15 panel members could be summarized in a Table.

We attempted to compile all of the individual reviews in one table but ultimately decided that it would be too complicated to be readily interpretatble. Therefore we have decided to add to the existing table another two columns for the agreement among the GU and non GU dedicated pathologist with the gold standard.

7. How was the histopath characteristics of the cancers? Grade? Extent? It may be very easy to recognize cancer vs atrophy. Were cancers selected with an emphasis on difficult cases, e.g. cancer of atrophic type.

The purpose of the study was not the inter-rater agreement on prostate cancer. Apart from the circled atrophic lesions we have asked to the reviewers to check and report the presence of cancer in the 61 slides and to provide Gleason's score according to the 2005 criteria that were the up-to-date classification at that time. Cancer areas were not circled and there were random tumors with variable Gleason score but all with conventional acinar morphology.

8. Were any of the patients treated with hormones or radiotherapy?

None of the patients was treated with hormone ablation or radiation therapy. This information is now included in the Methods section.

9. The first statement in the Introduction is that long term inflammation is an important aetiology factor in PCa. This is still hypothetical. The link between inflam and cancer is theoretically interesting but so far it has little practical relevance. Chronic inflammation and atrophy is almost ubiquitous in middle aged and elderly men. There is no evidence that it would have clinical relevance to report these lesions. No recommendations have been issued that these patients should be followed or treated in a different way that any other man with a PSA elevation and non-symptomatic benign biopsies. It would change the game considerably if such recommendations were issued and convincing studies would be needed. The authors are therefore recommended to express their views in a much more cautious way.
10. I disagree with the main conclusion that the results support the inclusion of PAH in the diagnosis. Just because a diagnosis is reproducible this does not mean that it is clinically relevant. Different sorts of atrophy may be seen in a large number of elderly men and has no clinical importance as far as we know today. It may be relevant to mention e.g. PAH in a report just to make sure that it is documented and a pathologist in a referral center would see that the lesion has been observed and interpreted as benign. The statement in the abstract that atrophy and inflammation is underestimated is misleading. They may very well be observed but many pathologists consider that they are not relevant to report.

We agree with the reviewer that the link between histological inflammatory lesions and prostate cancer is not yet demonstrated. However, we are enforced to disagree with the reviewer on the direct link between inflammation and prostate cancer that is supported by mounting evidence in the literature. A search on Pubmed for "inflammation and prostate cancer" yields 2104 papers and most of them are related to infections, dietary habits and metabolic diseases inducing higher inflammation regardless of age. Histological inflammatory changes are not generally described in pathology reports except for acute or granulomatous inflammation. This is peculiar compared to many other organs were chronic inflammation is consistently reported and it is related to various risks (i.e. in the stomach or in the liver). The main goal of the present paper is to sensitize the pathology audience to the relevance of inflammation in prostate cancer and the potential correlations to histology that could be demonstrated only after widespread implementation of this nomenclature in the pathology reports.

11. Some linguistic revision is needed

The manuscript has been proof-read by the native English speaking co-authors

Reviewer #2: In this sudy, Giunchi and co-workers evaluated the diagnostic concordance of several bening and malignant prostatic lesions, with emphasis on

	atrophic lesions, among 15 Pathologists, some of then especially dedicated to GU tract Pathology. They found that overall agreement was "fair" for PAH and partial atrophy, whereas it was strong to almost perfect for PIN and prostate carcinoma. Results improved when a single gold standard GU tract Pathologist was designated. In this setting, the concordance rating did not significantly differed between GU Pathologists and Non-GU Pathologists. Based on these findings, the authors propose that a statement on PAH/PIA should be included in standard pathology reporting of all prostate specimens. Globally, this is na interesting study that adds to the published literature on diagnostic concordance on prostate pathology. Nevertheless, some issues require clarification. We thank the reviewer for the appreciation of our work Major issues: 1. The diagnostic problems in prostatic pathology (as well as in other organs) are different according to specimen under analysis, i.e., diagnosing prostate cancer, PIN or atrophic lesions in biopsy is a quite different challenge than in prostatectomy specimens. As the study lumped together diagnosis in biopsies, radical prostatectomies and TURPs, it is not clear whether specimen type influenced the results. Thus, not only the number of cases but also the concordance statistics for each type of specimen should be provided. The type of sample actually could not effect the analysis. In fact, the two pathologists who selected the cases drew circles around 120 areas of interests with atrophic lesions with a bullpen in 61 histological slides. We requested to the 15 pathologists who accepted to join the study to score just the areas of interest except for checking the entire slide in case of prostate cancer and PIN. We have now added this information in the Methods section. 2. Apparently, diagnoses were based on HE stained slides, only. Nevertheless, it is widely accepted that adequate immunostains are a valuable ancillary technique for differential diagnosis among prostate cancer, PIN and atrophic lesions in challenging cases. Please indicate why immunostaning was not included. Since the areas of interest with atrophic lesions were already circled in the slides we have decided to send for review just the H&E. 3. It is somewhat confusing that concordance increased when a "gold standard" was set for comparison (Method 2). This seems counterintuitive. Could the authors provide an explanation for the source of variation in Method 1 compared to Method 2? The first method considers all possible pairs of reviewers and then computes the average agreement between all of those possible pairs. This method produces lower percent agreement because all of the reviewers, whether they correctly or incorrectly identified a particular lesion type, are compared to each other. The misclassification of lesion types is compounded in this analysis, but it has the advantage of not relying heavily on the evaluation of a single reviewer. In contrast, the second method compares the scoring to a single rater. This rater was selected because of his expertise in the field and is likely to misclassify lesions less frequently than the other reviewers. Therefore, this analysis more closely reflects the percent agreement with the “true” lesion type. 4. The similar results observed for concordance among GU dedicated and non-GU dedicated Pathologists is also surprising, considering previous publications on the diagnostic concordance between expert Uropathologists vs. community-based Pathologists. How do the authors explain this homogeneity of results? We agree with the reviewer on this issue. A possible explanation is that the pathologists selected for the study were recruited in medical centers with considerable workload of GU pathology in the routine practice. Another possible reason could be related to circling of the areas of interest that might have guided the reviewer. We have now added a sentence about this in the study limitations part of the Discussion. 5. It should be made clearer why reporting PAH/PIA could improve patient care and in

	which setting (prostate biopsy for prostare cancer suspects?). How would this impact patient management? We agree with the reviewer that currently there is no direct evidence of the predictive/prognostic role of PIA/PAH. However, the direct link between inflammation and prostate cancer is supported by mounting evidence in the literature. A search on Pubmed for “inflammation and prostate cancer” yields 2104 papers and most of them are related to infections, dietary habits and metabolic diseases inducing higher inflammation. Histological inflammatory changes are not generally described in pathology reports except for acute or granulomatous inflammation. This is peculiar compared to many other organs were chronic inflammation is consistently reported and it is related to various risks (i.e. in the stomach or in the liver). The main goal of the present paper is to sensitize the pathology audience to the relevance of inflammation in prostate cancer and the potential correlations to histology that could be demonstrated only after widespread implementation of this nomenclature in the pathology reports. Minor issues: 1. Abstract, line 3: should read "...widely underestimated and generally not mentioned..." 2. Page 4, line 7: should read "...1990s..." 3. Page 5, line 2: should read "(SACF)" instead of "(PACF)" 4. Page 5, line 7: should read "...found that 83% of atrophic lesions..." 5. Page 5, line 59: please add the designation of the kappa category 0.61-0.80 Thank you. We made the corrections.

Reviewer #1: Giunchi et al describe the interobserver variability of the diagnosis of atrophic lesions of the prostate, PIN and prostate cancer. Compared to previous studies, the current study uses real glass slides and includes both malignant and benign lesions. This may be of some interest but there are several problems with the paper.

We thank the reviewer for the appreciation of our work.

1. Two uropathologists selected the slides but the gold standard seems to be based on the assessment of one of them. Who was used as gold standard? Was his gold standard diagnosis unbiased by the other colleague? Or did the two expert pathologists discuss the cases in an open discussion? Did he first decide the gold standard and then a selection was done by the two of them together or vice versa? If the selection was done first it must be difficult to avoid discussing the diagnosis. This would influence their diagnoses and affect the results of the reproducibility study.

As stated in the Methods section the cases were selected by two dedicated uro-pathologists (EB and MF) from two different institutions, Each pathologist was blinded to the other's evaluation and they did not discuss the cases. The gold standard was another GU pathologist (RM) from a third institution who did not participate in case selection. For comparison, we also present data for the overall percentage agreement, which would not be as heavily dependent on the evaluation by a single reviewer.

2. 120 slides were selected including biopsies, RPs, TURPs. How many of each? Why are there only 115 and 116 cases, respectively in Tab 1? Were all cases circulated among 30 pathologists? Did you use recuts? How many sets?

The two pathologists who selected the cases asked for the re-cutting of three copies of each selected block. Then they drew circles around 121 areas of interests with a bullpen in 61 histological slides. The three sets of slides were then circulated among the 15 pathologists who accepted to join the study. One pathologist did not review 5 cases leaving 116 cases for the analysis of prostate cancer. Another pathologist did not review 1 case for some types of inflammation/atrophy, and two cases for other types of inflammation/atrophy, leaving either 114 or 115 cases for those analyses. We have now added this information in the Methods section.

3. Which pathologists participated? The panel had 15 members but there are 11 institutions in the author list. This indicates that some work in the same department. Should be specified as it may affect the results.
The 15 pathologists belonged to 9 institutions since other centers participated just for the statistical analyses. Slides were circulated blindly and each pathologist provided separate review files to the statisticians who completed the statistical analyses. This was made possible since 6 and 2 pathologists respectively belonged to the two institutions that were in charge of case selection.

4. There should be a Table where the gold standard diagnoses are compared to the results of the other pathologists. I.e. for each diagnostic entity the number of gold standard diagnoses (i.e. the total number of cases) should be given and then the summarized results of the other panelists.

5. There is no information about how the panelists voted, only percentages of agreement and kappas.

6. How was the performance of individual pathologists? The 15 panel members could be summarized in a Table.

We attempted to compile all of the individual reviews in one table but ultimately decided that it would be too complicated to be readily interpretatble. Therefore we have decided to add to the existing table another two columns for the agreement among the GU and non GU dedicated pathologist with the gold standard.

7. How was the histopath characteristics of the cancers? Grade? Extent? It may be very easy to recognize cancer vs atrophy. Were cancers selected with an emphasis on difficult cases, e.g. cancer of atrophic type.

The purpose of the study was not the inter-rater agreement on prostate cancer. Apart from the circled atrophic lesions we have asked to the reviewers to check and report the presence of cancer in the 61 slides and to provide Gleason's score according to the 2005 criteria that were the up-to-date classification at that time. Cancer areas were not circled and there were random tumors with variable Gleason score but all with conventional acinar morphology.

8. Were any of the patients treated with hormones or radiotherapy?

None of the patients was treated with hormone ablation or radiation therapy. This information is now included in the Methods section.

9*. The first statement in the Introduction is that long term inflammation is an important aetiology factor in PCa. This is still hypothetical. The link between inflam and cancer is theoretically interesting but so far it has little practical relevance. Chronic inflammation and atrophy is almost ubiquitous in middle aged and elderly men. There is no evidence that it would have clinical relevance to report these lesions. No recommendations have been issued that these patients should be followed or treated in a different way that any other man with a PSA elevation and non-symptomatic benign biopsies. It would change the game considerably if such recommendations were issued and convincing studies would be needed. The authors are therefore recommended to express their views in a much more cautious way.*

10. I disagree with the main conclusion that the results support the inclusion of PAH in the diagnosis. Just because a diagnosis is reproducible this does not mean that it is clinically relevant. Different sorts of atrophy may be seen in a large number of elderly men and has no clinical importance as far as we know today. It may be relevant to mention e.g. PAH in a report just to make sure that it is documented and a pathologist in a referral center would see that the lesion has been observed and interpreted as benign. The statement in the abstract that atrophy and inflammation is underestimated is misleading. They may very well be observed but many pathologists consider that they are not relevant to report.

We agree with the reviewer that the link between histological inflammatory lesions and prostate cancer is not yet demonstrated. However, we are enforced to disagree with the reviewer on the direct link between inflammation and prostate cancer that is supported by mounting evidence in the literature. A search on

Pubmed for “inflammation and prostate cancer” yields 2104 papers and most of them are related to infections, dietary habits and metabolic diseases inducing higher inflammation regardless of age. Histological inflammatory changes are not generally described in pathology reports except for acute or granulomatous inflammation. This is peculiar compared to many other organs were chronic inflammation is consistently reported and it is related to various risks (i.e. in the stomach or in the liver). The main goal of the present paper is to sensitize the pathology audience to the relevance of inflammation in prostate cancer and the potential correlations to histology that could be demonstrated only after widespread implementation of this nomenclature in the pathology reports.

11. Some linguistic revision is needed

The manuscript has been proof-read by the native English speaking co-authors

Reviewer #2: In this sudy, Giunchi and co-workers evaluated the diagnostic concordance of several bening and malignant prostatic lesions, with emphasis on atrophic lesions, among 15 Pathologists, some of then especially dedicated to GU tract Pathology. They found that overall agreement was "fair" for PAH and partial atrophy, whereas it was strong to almost perfect for PIN and prostate carcinoma. Results improved when a single gold standard GU tract Pathologist was designated. In this setting, the concordance rating did not significantly differed between GU Pathologists and Non-GU Pathologists. Based on these findings, the authors propose that a statement on PAH/PIA should be included in standard pathology reporting of all prostate specimens.
Globally, this is na interesting study that adds to the published literature on diagnostic concordance on prostate pathology. Nevertheless, some issues require clarification.

We thank the reviewer for the appreciation of our work

Major issues:

1. The diagnostic problems in prostatic pathology (as well as in other organs) are different according to specimen under analysis, i.e., diagnosing prostate cancer, PIN or atrophic lesions in biopsy is a quite different challenge than in prostatectomy specimens. As the study lumped together diagnosis in biopsies, radical prostatectomies and TURPs, it is not clear whether specimen type influenced the results. Thus, not only the number of cases but also the concordance statistics for each type of specimen should be provided.

The type of sample actually could not effect the analysis. In fact, the two pathologists who selected the cases drew circles around 120 areas of interests with atrophic lesions with a bullpen in 61 histological slides. We requested to the 15 pathologists who accepted to join the study to score just the areas of interest except for checking the entire slide in case of prostate cancer and PIN. We have now added this information in the Methods section.

2. Apparently, diagnoses were based on HE stained slides, only. Nevertheless, it is widely accepted that adequate immunostains are a valuable ancillary technique for differential diagnosis among prostate cancer, PIN and atrophic lesions in challenging cases. Please indicate why immunostaning was not included.

Since the areas of interest with atrophic lesions were already circled in the slides we have decided to send for review just the H&E.

3. It is somewhat confusing that concordance increased when a "gold standard" was set for comparison (Method 2). This seems counterintuitive. Could the authors provide an explanation for the source of variation in Method 1 compared to Method 2?

The first method considers all possible pairs of reviewers and then computes the average agreement between all of those possible pairs. This method produces lower percent agreement because all of the reviewers, whether they correctly or incorrectly identified a particular lesion type, are compared to each other. The misclassification of lesion types is compounded in this analysis, but it has the advantage of not relying heavily on the evaluation of a single reviewer. In contrast, the second method compares the scoring to a single rater. This rater was selected because of his expertise in the field and is likely to misclassify lesions less frequently than the other reviewers. Therefore, this analysis more closely reflects the percent agreement with the “true” lesion type.

4. The similar results observed for concordance among GU dedicated and non-GU dedicated Pathologists is also surprising, considering previous publications on the diagnostic concordance between expert Uropathologists vs. community-based Pathologists. How do the authors explain this homogeneity of results?

We agree with the reviewer on this issue. A possible explanation is that the pathologists selected for the study were recruited in medical centers with considerable workload of GU pathology in the routine practice. Another possible reason could be related to circling of the areas of interest that might have guided the reviewer. We have now added a sentence about this in the study limitations part of the Discussion.

5. It should be made clearer why reporting PAH/PIA could improve patient care and in which setting (prostate biopsy for prostare cancer suspects?). How would this impact patient management?

We agree with the reviewer that currently there is no direct evidence of the predictive/prognostic role of PIA/PAH. However, the direct link between inflammation and prostate cancer is supported by mounting evidence in the literature. A search on Pubmed for “inflammation and prostate cancer” yields 2104 papers and most of them are related to infections, dietary habits and metabolic diseases inducing higher inflammation. Histological inflammatory changes are not generally described in pathology reports except for acute or granulomatous inflammation. This is peculiar compared to many other organs were chronic inflammation is consistently reported and it is related to various risks (i.e. in the stomach or in the liver). The main goal of the present paper is to sensitize the pathology audience to the relevance of inflammation in prostate cancer and the potential correlations to histology that could be demonstrated only after widespread implementation of this nomenclature in the pathology reports.

Minor issues:

1. Abstract, line 3: should read "...widely underestimated and generally not mentioned..."

2. Page 4, line 7: should read "...1990s..."

3. Page 5, line 2: should read "(SACF)" instead of "(PACF)"

4. Page 5, line 7: should read "...found that 83% of atrophic lesions..."

5. Page 5, line 59: please add the designation of the kappa category 0.61-0.80

Thank you. We made the corrections.

Concordance in the histological diagnosis of Focal Prostatic Atrophy Lesions, Acute and Chronic Prostatitis, PIN and Prostate Cancer Among 15 European Pathologists

Francesca Giunchi MD[1], Kristina Jordahl PhD[2], Enrico Bollito MD[3], Maurizio Colecchia MD[4], Carlo Patriarca MD[5], Antonietta D'Errico MD[1], Francesco Vasuri MD, PhD[1] , Deborah Malvi MD[1], Alessandro Fornari MD[3], Luca Reggiani Bonetti MD[6], Barbara Corti MD[1], Mauro Papotti MD[7], Paolo DeGiuli MD[8], Massimo Loda MD[9], Rodolfo Montironi MD[10], Michelangelo Fiorentino MD, PhD[1] and, Jennifer R. Rider ScD[2, 11]

MF and JR co-shared senior authorship.

[1] *Pathology Department, S. Orsola-Malpighi Hospital, University of Bologna, Viale Ercolani 4/2, 40138 Bologna, Italy*

[2] *Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, 677 Huntington Avenue, MA 02115, USA.*

[3] *Department of Oncology, Division of Pathology, San Luigi Gonzaga Hospital, Orbassano, Turin area, Regione Gonzole, 10, 10043 Italy.*

[4] *Pathology Department, Fondazione IRCCS Istituto Nazionale dei Tumori, Milano, Via Giacomo Venezian 1, 20133, Italy.*

[5] *Pathology Department, St Anna Hospital, Como, Via Ravona, 20, 22020, Italy.*

[6] *Pathology Department, University of Modena and Reggio Emilia, Via Università, 4, 41121, Italy.*

[7] *Department of Oncology and Pathology, University of Turin, Turin, Via Giuseppe Verdi, 8, 10124, Italy.*

[8] *Pathology Department, S.Lazzaro Hospital, Alba, Via Pietrino Belli, 26, 12051, Italy.*

[9] *Medical Oncology Department, Dana-Farber Cancer Institute, Boston , 450 Brookline Ave, Boston, MA 02215, USA*

[10] *Pathology Department, Polytechnic University of the Marche Region, School of Medicine, United Hospitals, Ancona via Conca 71, 60126, Italy.*

[11]*Department of Epidemiology, Boston University School of Public Health, Boston, 715 Albany Street, T317E, MA 02118, USA.*

Corresponding author:

Michelangelo, Fiorentino MD, PhD.

Addarii Institute of Oncology,

Viale Ercolani 4/2, 40138, Bologna, Italy.

Phone: +36 051 636 4556

michelangelo.fiorentino@aosp.bo.it

Abstract:

Epidemiological and biological evidence indicate a causal relationship between the presence of proliferative atrophic lesions (PAH), the development of PIN and prostate cancer. Inflammatory and atrophic lesions of the prostate are widely underestimated and not generally mentioned in pathology reports.

We performed a histopathological concordance study among 15 dedicated and non-dedicated genito-urinary pathologists on 116 histological slides containing prostate atrophic lesions, PIN and cancer.

We found an overall percent agreement between all possible pairs of reviewers of 80.2% for prostate cancer, 67% for PIN, and 48.7% for any atrophic changes. When we designated a genitourinary pathologist single gold standard, the mean percent agreement increased: 96.6% for prostate cancer, 91.7% for PIN, 71.9% for PAH.

According to the raising relevance of PAH in prostate cancer our results on histopathological concordance support the inclusion of at least PAH in the routine pathology reporting of pathological prostate specimens.

Key words: Atrophic lesions, Inflammation, PAH, Prostate.

Introduction:

Long-term chronic inflammation is linked to the development of carcinoma in several organ systems and it is also an important aetiology factor in prostate cancer (PCa). [1] Based on observations in other organs such as the stomach, liver and large bowel in the 1990s, the interest in inflammatory and atrophic lesions in prostate cancer was further cultivated. Inflammatory cells could produce cellular or genomic irreversible damage in prostate cells; cause loss of tolerance to normal prostate antigens; and induce an autoimmune self-perpetuating reaction leading in turn to a "field effect" for the development of PCa [2]**.** The major events potentially leading to prostate inflammation are infections (virus, fungi, mycobacteria and parasite, and rarely bacteria), hormonal alteration, physical trauma, urine reflux and dietary habits. These exposures can result in injury to the luminal cell layer, which in turn induces reactive (defensive) hyperplasia of basal cells called "proliferative inflammatory atrophy" (PIA), thereby initiating genetic instability. Cytokines released by the inflammatory cells slowly induce epithelial proliferation and angiogenesis with the accumulation of genomic changes, eventually resulting in neoplastic transformation through PIN (prostatic intraepithelial neoplasia). [1] Several studies have referred to a morphological transition between PIA and PIN, as well as PIA and prostate cancer. These observations are further supported by evidence of an elevated proliferative fraction in atrophic areas and the closely proximity of these regions to PIN and PCa.[3, 4] In addition somatic genomic alterations detectable in PIN and PCa have been found in cells in PIA. In particular, these cells show molecular effects of inflammatory stress, such as high levels of glutathione S-transferase P1 (GSTP1), GSTA1 and cyclooxygenase-2 (COX-2). [2]

Prostatic inflammation and atrophy are not routinely collected as part of a standard histopathologic review, yet there is considerable interest in the potential role of atrophy and inflammation in the development and progression of prostate cancer. Recent studies have related the presence of prostatic atrophic lesions and inflammation to prostate cancer risk [5, 6, 7, 8] and survival [9], with mixed results. One possible explanation for the inconsistent findings is misclassification of the various lesion types due to subjectivity in grading.

The increasing correlation among inflammation, atrophic lesions, PIN and PCa led De Marzo et al. to propose a histological classification of the different atrophic lesions in 4 morphological patterns: simple atrophy (SA), partial atrophy (PA), post atrophic hyperplasia (PAH) and simple athropy with cyst formation (SACF). PAH can be considered a surrogate of PIA. While other studies have shown that PIN is frequently overdiagnosed in pathological specimens [10], the reliability of the assessment of inflammation and atrophic changes is unclear. A concordance study found that 83% of atrophic lesions were classified as the correct subtype, but there was substantial variability in accuracy across specific lesion types [11]. To specifically address this gap in the literature, we undertook a histopathological concordance study that focused on inflammation, atrophy, PIN and cancer among 15 European pathologists, including both those dedicated and non-dedicated to genitourinary pathology.

Materials and Methods:

Two dedicated uropathologists (MF, EB) selected 61 slides of prostate tissue and identified 121 areas of interests (ROIs) with inflammation, atrophic lesions, PIN and PCa on the basis of the original pathology report from biopsies, radical prostatectomy and TURP specimens. The slides were selected from the Pathology archives at the S. Orsola-Malpighi Hospital in Bologna and the S. Luigi Gonzaga Hospital in Turin. None of the included patients was treated with hormone ablation or radiation therapy. Specifically, the two pathologists who selected the cases asked for the re-cutting of three copies of each selected block. Then they drew circles around the area of interests with a bullpen. The three sets of slides were anonymized and circulated among the 15 pathologists who accepted to join the study for histological revision of prostate atrophic lesions according to the De Marzo et al. classification, [11]. Twenty-one of the 61 slides included at least a single focus of PCa together with atrophic lesions. Of the 30 pathologists originally invited to the study, 15 professionals from 9 centers accepted to review the slides and completed the evaluation form. Eight of these 15 pathologists were dedicated experienced uropathologists while 7 were general pathologists. Participating pathologists were asked to independently record the presence of the following histological features: prostate cancer, PIN, SA, SACF, PAH/PIA, PA, acute prostatitis, and chronic prostatitis. Two of the uropathologists did not review all of slides for all lesion types, leaving 116 slides for analysis of

prostate cancer, 115 for most types of inflammation and atrophy, and 114 for simple atrophy with cyst formation. . The entire review process took about two years.

Statistical Methods

We used two methods to evaluate the interpreter reliability of each histological feature. First we considered the overall percentage agreement, which is the average percent agreement across all possible rater pairs. We also estimated kappa statistics, which range from -1 to 1, where 1 indicates perfect agreement, 0 represents the agreement expected by chance, and values <0 indicate less agreement than is expected by chance. A suggested interpretation of kappa is that 0.21-0.40 is fair agreement, 0.41-0.60 is moderate agreement, and 0.61-0.99 is almost perfect agreement. The exact kappa coefficient was developed because it reduces to Cohen's un-weighted Kappa for two raters. Second, we used the rating of one particularly experienced dedicated genitourinary pathologist (RM) as the "gold standard" by which to compare all other raters. We calculated the mean percentage agreement and mean Cohen's Kappa across the 15 pairs. Comparisons were made among all 15 pathologists, and the subset of dedicated GU pathologists (N=8).

Results:

Table 1 shows that the overall percent agreement between all possible pairs of reviewers was 80.2% for prostate cancer, 67% for PIN, and 48.7% for any atrophic changes. While the kappa statistics for prostate cancer indicated nearly perfect agreement, agreement for PIN and atrophic changes were in the fair to moderate range. When specific types of focal prostatic atrophy were considered, the overall percentage agreement ranged from 5.2% for SA to 43% for SACF, all with kappa statistics indicating modest precision. In particular the agreement for PAH and partial atrophy was similar and low: 26.1% and 22.6% respectively. The average agreement between pairs of raters was 20% for chronic inflammation and 53% for acute inflammation.

When a single gold standard genitourinary pathologist was designated, the mean percent agreement increased: 96.6% for prostate cancer (kappa = 0.88) and 91.7% for PIN (kappa=0.55). There was

agreement with the gold standard by 92.5% of the raters regarding the presence of any atrophic changes, but the kappa statistic indicated that this agreement was often due to chance (kappa=0.47). Simple atrophy had the lowest agreement (65.2%), followed by PAH (71.9%), partial atrophy (84.7%), and simple atrophy with cyst formation (88.4%). Pathologists agreed with the gold standard rater for chronic and acute inflammation in 71.9% and 69.4% of ratings, respectively. The dedicated genitourinary pathologists consistently had higher agreement for all lesion types, which was especially pronounced for atrophic changes and specific types of focal atrophy. The agreement for atrophic lesions (92.2%) was stratified as follows: 86,9% for SACF, 84,6% for PA, 77.4% for PAH and 64.9% for SA. The agreement among non-genitourinary pathologists with the gold standard was similar to the results for dedicated genitourinary pathologists.

Discussion:

Our current understanding of PIA has raised questions about the role of this lesion in the development of prostate cancer. Moreover, the association between chronic inflammation/PAH and prostate cancer appears to be stronger compared to the other atrophic lesions such as SA. Nevertheless few studies have tested the diagnostic agreement among pathologists for atrophic lesions, which will impact the results of studies of this topic and inform the need to conduct centralized pathological reviews. In this study we evaluated the ability of general and genitourinary-dedicated pathologists to recognize atrophic lesions according to the classification proposed by De Marzo in order to determine whether there is utility in introducing the characterization of these lesions into standard pathology reports. Our results highlighted a moderate agreement for atrophic lesions among both genitourinary and general pathologists, but with patterns that varied by atrophy type. The agreement was favorable for PAH and simple atrophy with cyst formation among all the pathologists. However, for PA and SA the concordance was suboptimal. Further training on recognizing these lesions types may be warranted. With respect to the diagnosis of prostate cancer and PIN, our results confirmed several previous reports that found excellent agreement for cancer and very good agreement for PIN. Our concordance data support the standard reporting of PIN.

Our study is affected by several limitations. The number of raters and reviewed cases were both modest. However, we estimated that the 121 slides included in the study would allow us to identify differences in lesion attribution among the 15 participating pathologists. Other limitations derive from the binary categorization of the variables as presence or absence of each histological feature. Had we included in the review the extent of inflammation, the concordance findings may have changed. Similarly, we provided the reviewers with H&E slides that were already circled for areas of interest to be scored. If areas of interest were not indicated, as in real-world clinical and research settings, non GU-dedicated pathologists may have underestimated the atrophic lesions. Despite the reference pathologist being the most experienced, the reference gold standard was likely imperfect. However, if we selected as the gold standard, for instance, the majority opinion of the raters we may have introduced more substantial bias.

This study did not confirm results from other pathology review studies where dedicated genitourinary pathologists displayed higher diagnostic concordance for prostate cancer, PIN and atrophic lesions compared to general pathologists [12]. In our data, differences in the agreement with the gold standard for dedicated and general European pathologist were minimal.

Given the growing interest in inflammation in prostate cancer risk assessment, our results support the inclusion of at least PAH/PIA in the standard pathology reporting of all pathological prostate specimens, but that some specific types of atrophy and for inflammation, additional training may improve concordance.

Compliance with Ethical Standards:

Funding

No external funds were obtained for this work**.**

Conflict of Interest

The authors declare no conflicts of interest.

References

1. De Marzo AM, Platz EA, Sutcliffe S, et al. Inflammation in prostate carcinogenesis. Nat Rev Cancer. 2007 Apr;7(4):256-69.

2. Nelson WG, De Marzo AM, Isaacs WB. Prostate cancer. N Engl J Med. 2003 Jul 24;349(4):366-81.

3. Sfanos KS, De Marzo AM. Prostate cancer and inflammation: the evidence. Histopathology. 2012 Jan;60(1):199-215.

4. De Marzo AM, Marchi VL, Epstein JI, Nelson WG. Proliferative inflammatory atrophy of the prostate: implications for prostatic carcinogenesis. Am J Pathol. 1999 Dec;155(6):1985-92.

5. Moreira DM, Nickel JC, Gerber L, et al. Baseline prostate inflammation is associated with a reduced risk of prostate cancer in men undergoing repeat prostate biopsy: results from the REDUCE study.Cancer. 2014 Jan 15;120(2):190-6.

6. Porcaro AB, Rubilotta E, Petrozziello A, et al. Chronic inflammation of the prostate type IV with respect to risk of prostate cancer. Arch Ital Urol Androl. 2014 Sep 30;86(3):208-11.

7. Yli-Hemminki TH, Laurila M, Auvinen A, et al. Histological inflammation and risk of subsequent prostate cancer among men with initially elevated serum prostate-specific antigen (PSA) concentration in the Finnish prostate cancer screening trial. BJU Int. 2013 Oct;112(6):735-41.

8. Gurel B, Lucia MS, Thompson IM Jr, et al. Chronic inflammation in benign prostate tissue is associated with high-grade prostate cancer in the placebo arm of the prostate cancer prevention trial. Cancer Epidemiol Biomarkers Prev. 2014 May;23(5):847-56.

9. Davidsson S, Fiorentino M, Andrén O, et al. Inflammation, focal atrophic lesions, and prostatic intraepithelial neoplasia with respect to risk of lethal prostate cancer. Cancer Epidemiol Biomarkers Prev. 2011 Oct;20(10):2280-7.

10. Bostwick DG, Ma J. Over-diagnosis of high-grade prostatic intraepithelial neoplasia: a prospective study of 251 cases. BJU Int. 2007 Nov;100(5):1036-9.

11. De Marzo AM, Platz EA, Epstein JI, et al. A working group classification of focal prostate atrophy lesions. Am J Surg Pathol. 2006 Oct; 30(10):1281-91.

12. Egevad L, Ahmad S.A, Algaba F, et al. Standardization of Gleason grading among 337 European pathologists. Histopathology 2013, 62, 247-256.

Table 1. Concordance among 15 pathologists in histological assessment of inflammation, atrophy, PIN and cancer

	Subjects	All pathologists N=15		GU pathologists (N=7)		Non-GU pathologists (N=8)		GU pathologists (N=7) Vs gold standard		Non-GU pathologists (N=8) Vs gold standard		All pathologists (N=15) Vs gold standard	
		Overall % agreement	Exact kappa	Average % agreement	Average Cohen's Kappa	Average % agreement	Average Cohen's Kappa	Average % agreement	Average Cohen's Kappa	Average % agreement	Average Cohen's Kappa	Average % agreement	Average Cohen's Kappa
Acute inflammation	115	53.0%	0.34	70.0%	0.31	68.8%	0.28	70,0%	0,305	68,8%	0,276	69.4%	0.42
Chronic inflammation	115	20.0%	0.36	72.5%	0.46	71.5%	0.45	72,5%	0,456	71,5%	0,447	71.9%	0.46
Atrophic changes	115	48.7%	0.32	92.2%	0.30	89.1%	0.36	92,2%	0,298	89,1%	0,362	90.6%	0.47
PAH	115	26.1%	0.42	77.4%	0.44	79.6%	0.42	77,4%	0,442	79,6%	0,420	71.9%	0.55
Partial atrophy	115	22.6%	0.21	84.6%	0.23	84.7%	0.15	84,6%	0,231	84,7%	0,148	84.7%	0.39
Simple atrophy	115	5.22%	0.22	64.9%	0.22	65.5%	0.25	64,9%	0,221	65,5%	0,251	65.2%	0.38
Simple atrophy with cyst formation	114	43.0%	0.42	86.9%	0.57	89.8%	0.58	86,9%	0,567	89,8%	0,576	88.4%	0.59
PIN	115	67.0%	0.42	92.9%	0.44	90.7%	0.42	92,9%	0,435	90,7%	0,422	91.7%	0.55
Prostate cancer	116	80.2%	0.84	96.9%	0.89	96.3%	0.87	96,9%	0,888	96,3%	0,866	96.6%	0.88

Figure Legend: Figure A) Morphological patterns of the focal prostate atrophic lesions**:** partial atrophy (PA), simple atrophy with cyst formation (SACF), post-atrophic hyperplasia (PAH) and simple atrophy (SA) (H&E, 100x magnification). **B)** Morphological pathway of progression of prostate cancer development through simple atrophy, PAH and PIN (H&E, 100x magnification).

Table Legend: **Table 1.** Condorance among 15 pathologists in histological assessment of inflammation, atrophy, PIN and cancer

Table

Table 1. Concordance among 15 pathologists in histological assessment of inflammation, atrophy, PIN and cancer

	Subjects	All pathologists N=15		GU pathologists (N=7)		Non-GU pathologists (N=8)		GU pathologists (N=7) Vs gold standard		Non-GU pathologists (N=8) Vs gold standard		All pathologists (N=15) Vs gold standard	
		Overall % agreement	Exact kappa	Average % agreement	Average Cohen's Kappa	Average % agreement	Average Cohen's Kappa	Average % agreement	Average Cohen's Kappa	Average % agreement	Average Cohen's Kappa	Average % agreement	Average Cohen's Kappa
Acute inflammation	115	53.0%	0.34	70.0%	0.31	68.8%	0.28	70,0%	0,305	68,8%	0,276	69.4%	0.42
Chronic inflammation	115	20.0%	0.36	72.5%	0.46	71.5%	0.45	72,5%	0,456	71,5%	0,447	71.9%	0.46
Atrophic changes	115	48.7%	0.32	92.2%	0.30	89.1%	0.36	92,2%	0,298	89,1%	0,362	90.6%	0.47
PAH	115	26.1%	0.42	77.4%	0.44	79.6%	0.42	77,4%	0,442	79,6%	0,420	71.9%	0.55
Partial atrophy	115	22.6%	0.21	84.6%	0.23	84.7%	0.15	84,6%	0,231	84,7%	0,148	84.7%	0.39
Simple atrophy	115	5.22%	0.22	64.9%	0.22	65.5%	0.25	64,9%	0,221	65,5%	0,251	65.2%	0.38
Simple atrophy with cyst formation	114	43.0%	0.42	86.9%	0.57	89.8%	0.58	86,9%	0,567	89,8%	0,576	88.4%	0.59
PIN	115	67.0%	0.42	92.9%	0.44	90.7%	0.42	92,9%	0,435	90,7%	0,422	91.7%	0.55
Prostate cancer	116	80.2%	0.84	96.9%	0.89	96.3%	0.87	96,9%	0,888	96,3%	0,866	96.6%	0.88

Figure