POLITECNICO DI TORINO
Repository ISTITUZIONALE

Impact of Stain Normalization on Pathologist Assessment of Prostate Cancer: A Comparative Study

Original
Impact of Stain Normalization on Pathologist Assessment of Prostate Cancer: A Comparative Study / Salvi, M., Caputo, A., Balmativola, D., Scotto, M., Pennisi, O., Michielli, N., Mogetta, A., Molinari, F., Fraggetta, F.. - In: CANCERS. - ISSN 2072-6694. - ELETTRONICO. - 15:5(2023), p. 1503. [10.3390/cancers15051503]

Availability:
This version is available at: 11583/2976475 since: 2023-03-01T10:46:59Z

Publisher:
MDPI

Published
DOI:10.3390/cancers15051503

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

19 November 2026

cancers

MDPI

Article

Impact of Stain Normalization on Pathologist Assessment of Prostate Cancer: A Comparative Study

Massimo Salvi [1,*], Alessandro Caputo [2], Davide Balmativola [3], Manuela Scotto [1], Orazio Pennisi [4], Nicola Michielli [1], Alessandro Mogetta [1], Filippo Molinari [1] and Filippo Fraggetta [5]

[1] Biolab, PoliToBIOMed Lab, Department of Electronics and Telecommunications, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy
[2] Department of Medicine and Surgery, University Hospital of Salerno, 84084 Fisciano, Italy
[3] Pathology Unit, Humanitas Gradenigo Hospital, Corso Regina Margherita 8, 10153 Turin, Italy
[4] Technology Transfer and Industrial Liaison Department, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy
[5] UOC di Anatomia Patologica, ASP Catania P.O. "Gravina", 95041 Caltagirone, Italy
* Correspondence: massimo.salvi@polito.it

Simple Summary: Prostate cancer is the second most diagnosed cancer in men worldwide, with an estimated 1,276,000 new cases and 359,000 deaths in 2018. It is graded using the Gleason system into five grade groups of increasing tumor aggressiveness. However, diagnosis is hampered by a relatively high rate of inter- and intra-observer variability. Currently, the reduction of the perceived color variability is performed by physical quality controls, such as subjective assessment by visual inspection and comparison between laboratories. However, slides from different laboratories and even from different batches of the same laboratory may show significant color variations. The stain normalization procedure helps to standardize the stain color appearance of a digital image with respect to a reference image. In this study, we investigated the impact of the stain normalization process on prostate cancer biopsies from the pathologist's perspective.

Abstract: In clinical routine, the quality of whole-slide images plays a key role in the pathologist's diagnosis, and suboptimal staining may be a limiting factor. The stain normalization process helps to solve this problem through the standardization of color appearance of a source image with respect to a target image with optimal chromatic features. The analysis is focused on the evaluation of the following parameters assessed by two experts on original and normalized slides: (i) perceived color quality, (ii) diagnosis for the patient, (iii) diagnostic confidence and (iv) time required for diagnosis. Results show a statistically significant increase in color quality in the normalized images for both experts ($p < 0.0001$). Regarding prostate cancer assessment, the average times for diagnosis are significantly lower for normalized images than original ones (first expert: 69.9 s vs. 77.9 s with $p < 0.0001$; second expert: 37.4 s vs. 52.7 s with $p < 0.0001$), and at the same time, a statistically significant increase in diagnostic confidence is proven. The improvement of poor-quality images and greater clarity of diagnostically important details in normalized slides demonstrate the potential of stain normalization in the routine practice of prostate cancer assessment.

Keywords: digital pathology; prostate cancer; stain normalization; color quality; Gleason score

check for updates

Citation: Salvi, M.; Caputo, A.; Balmativola, D.; Scotto, M.; Pennisi, O.; Michielli, N.; Mogetta, A.; Molinari, F.; Fraggetta, F. Impact of Stain Normalization on Pathologist Assessment of Prostate Cancer: A Comparative Study. *Cancers* **2023**, *15*, 1503. https://doi.org/10.3390/cancers15051503

Academic Editor: Ewan Millar

Received: 24 January 2023
Revised: 17 February 2023
Accepted: 23 February 2023
Published: 27 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Prostate cancer (PCa) is the second most commonly diagnosed cancer in men worldwide, with an estimated 1,276,000 new cases and 359,000 deaths in 2018 [1]. The most recent world health organization (WHO) classification of urinary and male genital tumors [2] defines a five-tiered classification scheme [3] to better correlate the pre-existing Gleason score [4] with accumulating data on prognosis. The five resulting Grade Groups (GG), as

shown in Figure 1, range from 1 (corresponding to Gleason score 3 + 3 = 6) to 5 (corresponding to Gleason scores 4 + 5, 5 + 4, and 5 + 5) and are related to different clinical outcomes and treatment choices [5–10]. The usage of virtual microscopy and whole-slide images (WSIs) to diagnose prostate cancer has been shown by multiple groups to be adequate and non-inferior to conventional microscopic diagnosis [11–13]. However, PCa diagnosis is hampered by a relatively high rate of inter-observer and intra-observer discordance, by optical microscopy and virtual microscopy alike [14–16].

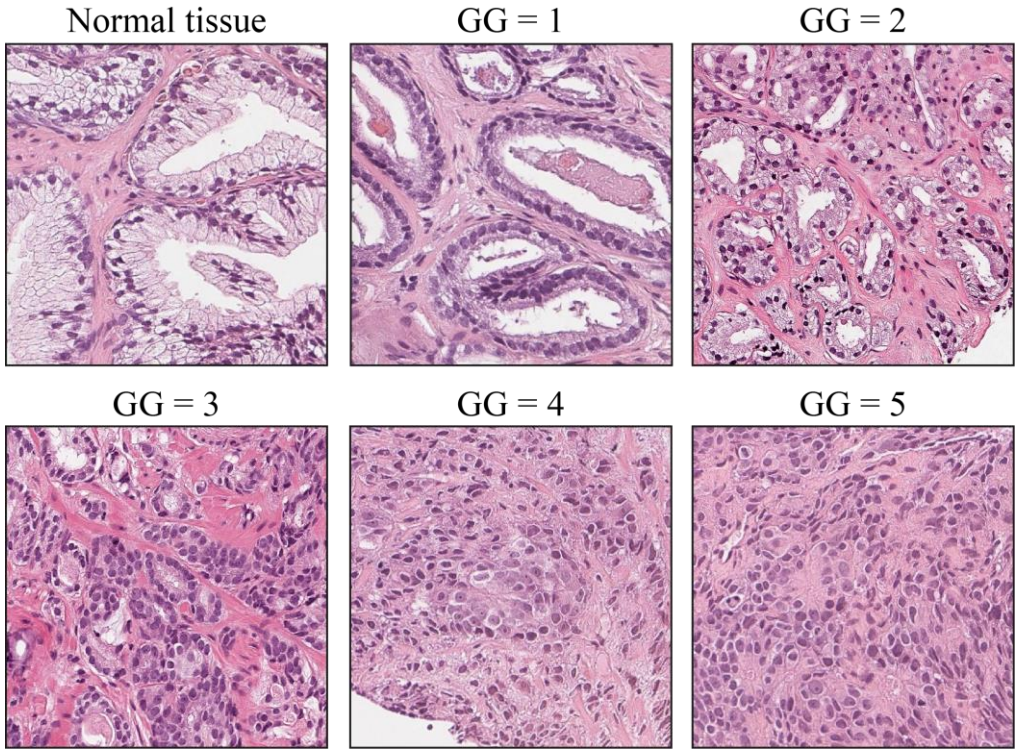


Figure 1. Prostate cancer grade groups. Normal prostatic tissue is shown for reference. Grade Group 1 (GG1) is a low-grade tumor with well-formed individual glands. With increasing GG, glands become fused, disorganized, solid, cribriform, or they are not formed at all (GG4, GG5).

Digital pathology is a rapidly growing field of pathology based on the adoption of new technologies such as WSIs, digital workflow and computer-aided diagnosis (CAD) [17]. Artificial intelligence (AI)-based tools are being used to assist pathologists and clinicians at large in the diagnosis, prognostication and treatment of patients [18,19]. The importance for a standardization of both procedures and reagents in clinical practice is emphasized in the study of Lyon [20]. However, current technology does not allow for complete standardization due to stain fading over time and the variability of the manual sectioning process. Currently, the reduction of perceived color variability and its impact on diagnosis is performed by procedural and physical quality controls, such as subjective quality assessment by visual inspection and comparison between laboratories. However, slides from different laboratories and even from staining batches of the same laboratory may be subjected to significant color variations [21,22]. These staining variations in tissue appearance make the quantitative analysis of histological tissue a complex process [23].

The quality of the histological slide plays a key role in the pathologist's diagnosis [24,25]. A staining that is too weak or too intense can hide important diagnostic details, reducing confidence, increasing the time for analysis, and possibly causing misdiagnosis. In recent years, the stain normalization procedure has been proposed to address this issue in digital slides. The stain normalization procedure helps to standardize the stain color appearance of a digital image (also denoted as source image) with respect to a reference image (also

denoted as target image). The target image is a single image chosen by the pathologist with the most optimal visual appearance and tissue staining. Stain normalization changes the chromatic information of the source image so that the color appearance matches the target image. In this way, the stain variability of histological images can be reduced. Several algorithms for color normalization have been proposed [21,26], and different studies have shown that normalized images enable improved performances of various algorithms for digital pathology image analysis [27–30]. However, to the best of our knowledge, no study has investigated the possible impact of the color normalization process on the PCa diagnostic routine on biopsies or the qualitative improvement due to this process from the pathologist's perspective.

In this study, we assess the impact of the normalization process within the diagnostic routine and specifically during the assessment of PCa. The analysis is focused on the evaluation of several parameters, such as the perceived color quality, the diagnosis for the patient, the confidence and the time for diagnosis, for both original and normalized WSIs. This paper is organized as follows: in Section 2, the description of materials and methods is provided, in Section 3, the experimental results are reported, and they are discussed in Section 4. In Section 5, the conclusions are summarized.

2. Materials and Methods

2.1. Dataset

In this study, WSIs of prostate tissue, derived from formalin-fixed and paraffin-embedded samples stained with hematoxylin and eosin (H&E) were selected from our previous works [29,31]. The selection was carried out with the aim of simulating the high stain variability of histopathological slides, ranging from lighter to darker stain color images with some minors suboptimal staining such as reddish, greyish color and weak staining. A total of 93 WSIs scanned from the same department of oncology and from different patients were collected. Bad-quality glass slides and images for which pathologists requested a second opinion with immunohistochemical markers were removed from the analysis. WSIs were digitized with three digital slide scanners: 15 slides were acquired with Aperio AT2 at 200× magnification (0.467 µm/pixel), 36 slides were scanned with Aperio GT450 at 400× magnification (0.233 µm/pixel), and 42 slides were captured with Hamamatsu NanoZoomer S210 at 200× magnification (0.467 µm/pixel).

2.2. Clinical Study Description

Stain normalization was performed for each WSI in our dataset, resulting in 2 subsets: (i) the original dataset (i.e., the 93 scanned WSIs) and (ii) the normalized dataset (i.e., the corresponding 93 normalized WSIs). We performed the normalization process with STAINS—STAndardIzation & Normalization of histological Slides—tool (AEQUIP S.r.l., Turin, Italy), an improved version of our previously published algorithm [26]. STAINS tool is based on a fully automated approach of stain normalization of H&E histopathological WSIs and consists of the following steps: (i) tissue detection to distinguish relevant information from background regions; (ii) illuminant correction based on white balancing; (iii) separation of hematoxylin and eosin color channels and (iv) normalization of both channels to adapt chromatic profile of target image. This process preserves local structures and enhances the contrast between histological tissue and background. The output of STAINS, regardless of the input image format, is a 200× tiled pyramidal TIFF image with JPEG lossy compression (quality factor = 80). For most tissue types and diagnostic tasks, optical magnification of 200× is considered sufficient to identify relevant biological features. The JPEG compression was used to save memory since the size of a single uncompressed WSI is usually over 2.5 GB [32]. The normalization of a single WSI using the STAINS tool took on average 4 min. An example of a prostate WSI, processed with the stain normalization tool, is shown in Figure 2.

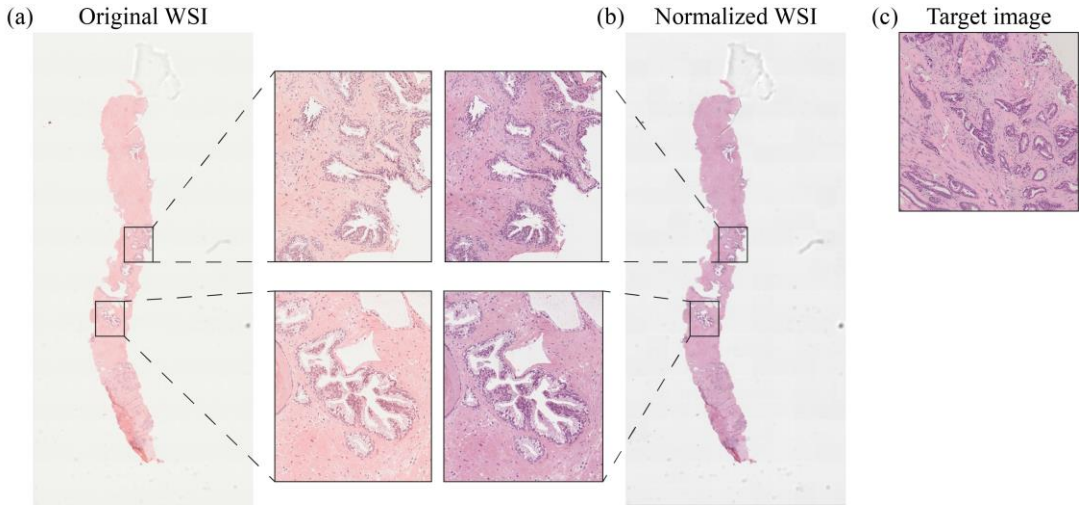


Figure 2. Stain normalization process for a prostate WSI. Original slide (**a**) and normalized slide (**b**) with zoomed-in views showing the effects at high magnification of color normalization with respect to target image (**c**).

Two different pathologists denoted as Pathologist 1 (P1) and Pathologist 2 (P2) from two medical centers and with different years of experience in pathology (P1 with 12 years and P2 with 6 years of experience) were involved in this study. Both pathologists analyzed original and normalized images using QuPath open-source software [33]. Firstly, P1 observed and evaluated the original dataset, and P2 observed and evaluated the normalized one. After 3 months, the process was repeated in the reverse direction, i.e., P1 was asked to evaluate the normalized images while P2 was asked to assess the original ones. The identification name of each WSI was randomly generated to prevent each pathologist from rechecking the case and the diagnosis previously performed. Each pathologist, for each slide, provided the following evaluations:

- Assessment of perceived stain color quality: it is quantified with a numerical scale from 1 to 10, where 1 indicates a low-quality image and 10 a high-quality image.
- Diagnosis of the given slide: if tumor is present, it is graded according to the Gleason Grade Groups [3,4] from 1 to 5. If no tumor is present on the given slide, a score of 0 is assigned.
- Confidence in the given diagnosis: rated subjectively from 1 to 10, where 1 indicates a low degree of reliability in diagnosis and 10 denotes a high degree of confidence. Operatively, a high-confidence diagnosis occurs when the pathologist thinks that the given slide is sufficient to perform a diagnosis, whereas low confidence indicates that the pathologist is not fully convinced by the appearance of the examined slide and would resort to recuts or immunohistochemical analysis.
- Time required for diagnosis: it is expressed in seconds and indicates the time taken by the pathologist to examine the image in order to decide the diagnostic classification. It was measured from the time when the image is opened (i.e., when the pathologist starts examining the image) to the time when the diagnosis is formulated; after that, the pathologist stops examining the image and writes down the diagnosis. The time required for image loading, thus, was not considered.

The workflow followed in this study is summarized in Figure 3.

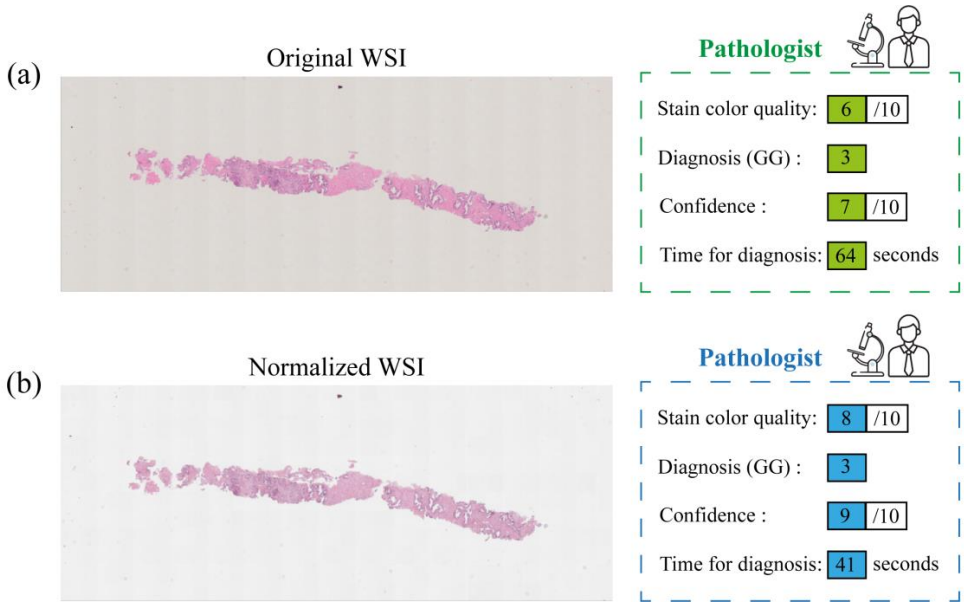


Figure 3. Workflow adopted in this study: an example of original WSI (**a**) and the corresponding normalized slide (**b**) processed by STAINS tool. Both pathologists (P1 and P2) provided evaluations on stain color quality, diagnosis, confidence, and time for diagnosis, for both datasets, i.e., the original images (an example of pathologist's evaluation is reported in green) and the corresponding normalized ones (an example of pathologist's evaluation is reported in blue).

The results of this study are divided into two main sections: color quality analysis and PCa evaluation. In the former, the pathologists' evaluation of color quality is analyzed, while in the latter, the evaluations of diagnosis, confidence and time for diagnosis are reported. In this way, we aim to assess the impact of the normalization process on the pathologist's perceived quality and diagnosis process. For the assessment of stain color quality, confidence and time for diagnosis, the values of the original and normalized WSIs are compared for each pathologist using a paired *t*-test with a 5% significance level. In addition, we assessed the agreement between pathologists in the evaluation of diagnosis using the quadratic weighted Cohen's kappa coefficient.

3. Results

3.1. Evaluation of the Color Quality

In this section, we compare color quality data of the 93 original WSIs and the corresponding normalized WSIs. Both pathologists, denoted as P1 and P2 in the following, evaluated the original and normalized dataset. We analyzed the distribution of color quality given by each pathologist. Figure 4 shows color quality values with boxplots for overall distributions. The median values for normalized image distributions (8/10 for P1 and 7/10 for P2) are greater than the median values of the original image distributions (6/10 for P1 and 5/10 for P2). In addition, for both overall distributions, 25th and 75th percentiles related to normalized images are higher than corresponding percentiles of original images. The increase in color quality in the normalized images is statistically significant (paired-sample *t* test, $p < 0.0001$ for both P1 and P2). This result shows that the perceived stain color quality is higher in normalized images for both pathologists.

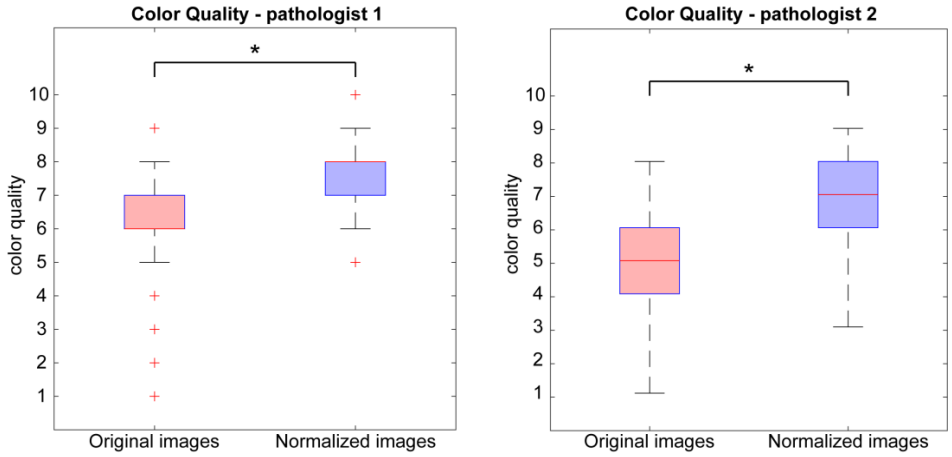


Figure 4. Boxplots of distributions of color quality values of original images (red) and normalized images (blue). The evaluation of the first pathologist (P1) is shown on the left, and the evaluation of the second pathologist (P2) is represented on the right. The plus symbol (+) indicates outliers in the distribution, while the asterisk (*) denotes a statistically significant difference between original and normalized images.

Examples of the normalization process performed on original images of suboptimal quality (score < 4/10) are shown in Figure 5. In particular, the color normalization process improves the quality of images showing reddish (Figure 5—sample #1) or grayish (Figure 5—sample #2) staining or too weak staining with poor contrast between cellular structures (Figure 5—sample #3). In addition to showing a better staining, normalized images have a more repeatable and standardized color appearance.

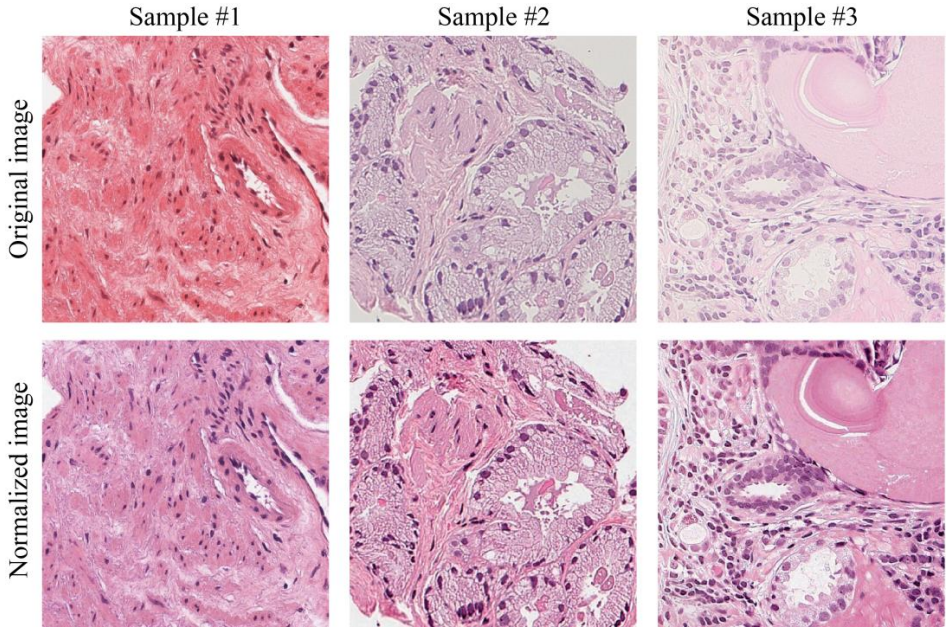


Figure 5. Visual performance of normalization process. In the first row, there are examples of tiles with suboptimal staining: reddish (sample #1), greyish color (sample #2) and weak staining (sample #3). The second row shows the corresponding normalized tiles.

3.2. Assessment of Prostate Cancer: Diagnosis, Time and Confidence

This section describes the results related to PCa evaluation: pathologists were asked to provide the diagnosis expressed as Grade Group (GG), confidence and time for diagnosis on the original 93 WSIs and the corresponding normalized ones.

Firstly, we analyzed the agreement between pathologists' diagnoses for both original and normalized images. The two experts' intra- and inter-rater agreements of GG values were investigated through the quadratic weighted Cohen's kappa coefficient. The intra-rater agreement (Cohen's kappa) for P1 (original vs. normalized) is 0.8908 and for P2 (original vs. normalized) is 0.8959. The inter-rater agreement (P1 vs. P2) is 0.8507 for original images and 0.7946 for normalized ones. Furthermore, Figure 6 shows the confusion matrices on GG scores to evaluate the consensus of both pathologists on original and normalized images.

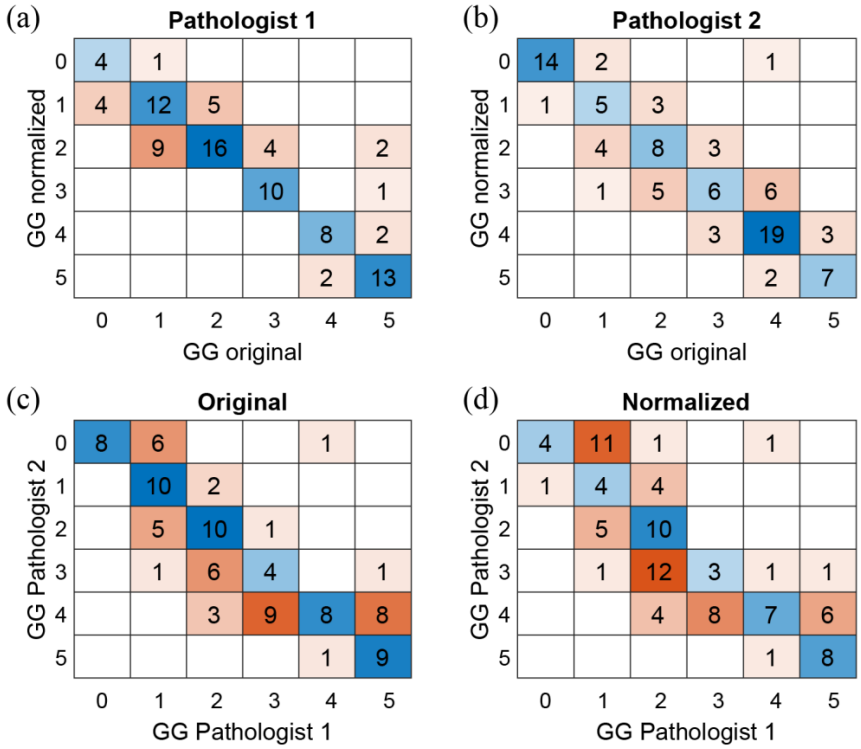


Figure 6. Confusion matrices on Gleason Grade Group (GG). First row shows the intra-agreement of P1 (**a**) and P2 (**b**) on the GG values of the original and normalized images. Second row shows the inter-agreement (P1 vs. P2) on the original (**c**) and normalized (**d**) WSIs.

The clinical process is also evaluated in terms of time employed to formulate a diagnosis. In Figure 7, the overall distributions of time for diagnosis for original WSIs and normalized ones are represented for both pathologists. Specifically, for P1, the median values are 72.0 s (original) vs. 65.0 s (normalized), and for P2, the values are 39.6 s (original) vs. 30.6 s (normalized). In addition, for P1, 33% of original images have a time for diagnosis higher than 90.0 s, while only 11% of normalized images have a time for diagnosis higher than 90.0 s. For P2, 14% of original images have a time for diagnosis higher than 90.0 s, while only 2% of normalized images have a time for diagnosis higher than 90.0 s. The decrease in diagnosis time in the normalized images is statistically significant (paired-sample t test, $p = 0.00015$ for P1 and $p < 0.0001$ for P2).

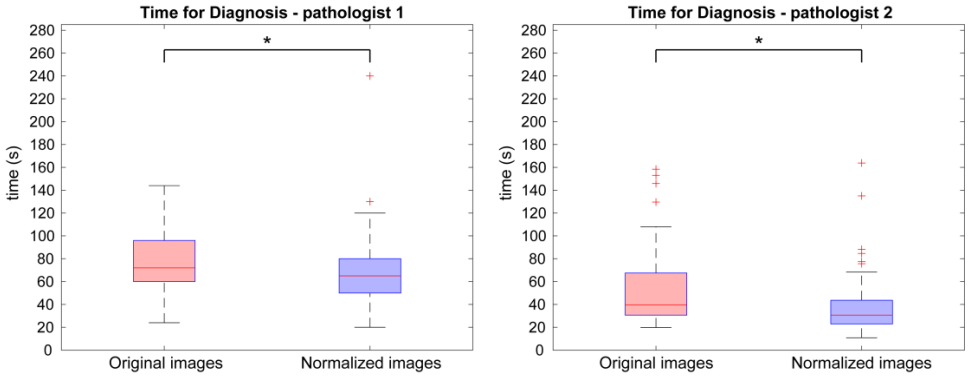


Figure 7. Comparison between time for diagnosis of the original and normalized WSIs. Boxplot of diagnosis time values of original (red) and normalized images (blue) according to the evaluations of the first pathologist (P1) on the left and second one (P2) on the right. The plus symbol (+) indicates outliers in the distribution, while the asterisk (*) denotes a statistically significant difference between original and normalized images.

Finally, we compared the diagnostic confidence related to original and normalized WSIs. We evaluated the overall distributions of the confidence data in Figure 8. Specifically, for P1, median value, 25th and 75th percentiles related to normalized images are higher than the ones related to original images, while for P2, median value and 75th percentile of normalized images are equal to the ones of the original image, while the 25th percentile is higher for normalized images. In addition, the minimum values of confidence related to original images (i.e., 3/10 for P1 and 1/10 for P2) are lower than minimum values related to normalized ones (i.e., 7/10 for P1 and 6/10 for P2). The increase in confidence in the normalized images is statistically significant (paired-sample t test, $p < 0.0001$ for P1 and $p = 0.00028$ for P2).

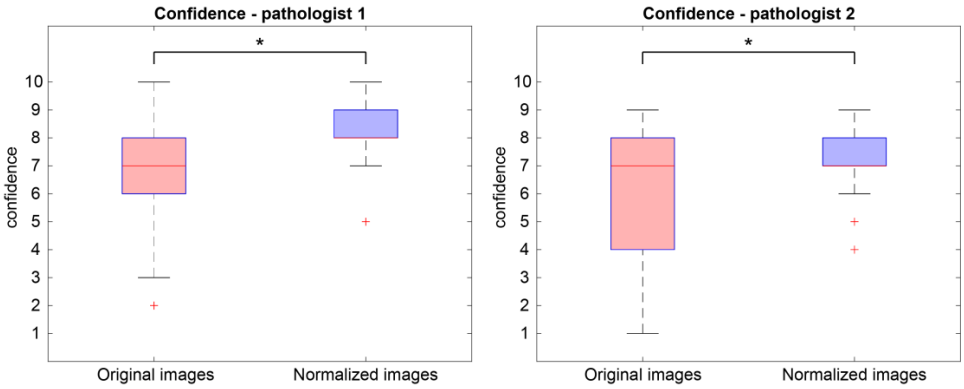


Figure 8. Boxplots of confidence values. Overall distribution of confidence values of original (red) and normalized images (blue) according to the evaluations of the first pathologist (P1) on the left and second one (P2) on the right. The plus symbol (+) indicates outliers in the distribution, while the asterisk (*) denotes a statistically significant difference between original and normalized images.

A summary of data related to the assessment of PCa is presented. According to the first pathologist (P1), mean and standard deviation values of time for diagnosis are 77.9 $\pm$ 26.6 and 69.9 $\pm$ 27.6 s for original and normalized images, respectively, while for the second expert (P2) are 52.7 $\pm$ 32.2 and 37.4 $\pm$ 24.4 s, respectively. Regarding confidence score, the original images have shown 7.0 $\pm$ 1.8 and 6.2 $\pm$ 2.4 for P1 and P2, respectively, while the corresponding normalized ones reported a confidence value of 8.2 $\pm$ 1.2 and 7.2 $\pm$ 1.2, respectively. No relevant influence about two different input magnifications

(i.e., 200× and 400×) was observed on the evaluation parameters: color quality, time for diagnosis and confidence in diagnosis.

4. Discussion

Staining variability is one of the problems affecting the quality of WSIs and can negatively influence the diagnostic process. These problems can be solved by using a stain-normalization procedure. This technique can change the color information of a digital image so that its color appearance matches the color appearance of the reference image chosen by the pathologist [21]. In our previous study, we proved the superiority of normalized images with respect to the corresponding original ones in terms of perceived image quality and absence of clinically significant artifacts in a multi-center and multi-tissue context [25].

In this study, we analyzed the impact of the normalization process on the clinical routine related to the diagnosis of PCa. For this purpose, we analyzed data collected from two pathologists: one pathologist (denoted as P1) performed the evaluation on 93 original scanned WSIs, while the second expert (denoted as P2) assessed Pca on the corresponding normalized WSIs. After 3 months, the process was repeated in the reverse direction, i.e., P1 evaluated the normalized images, and P2 analyzed the original ones. The specialists were asked to evaluate the images in terms of stain color quality, diagnosis, confidence, and time for diagnosis. To the best of our knowledge, no study analyzed the normalization process from a clinical perspective.

The experimental results are divided into two sections: perceived color quality and PCa evaluation. Regarding the color quality analysis, Figure 4 shows that the quality of normalized images is on average better than the quality of the corresponding original ones for both pathologists (6/10 vs. 8/10 for P1 and 5/10 vs. 7/10 for P2). The normalization process has a positive impact on the quality perceived by pathologists, and it is useful when the quality of starting original images is not satisfactory. Color normalization is also evaluated in terms of its impact on the PCa diagnostic process by analyzing tumor classification, confidence, and time for diagnosis. When comparing the diagnoses formulated by each pathologist on the original vs. normalized images, most are completely concordant or only slightly discordant (difference of a single class). This is shown in Figure 6 and is demonstrated by the weighted Cohen's kappa test (0.8908 and 0.8959 for P1 and P2, respectively). The inter-rater agreement is also very high (weighted Cohen's kappa: 0.8507 for original images and 0.7946 for normalized ones). The diagnosis process is also evaluated in terms of time. Results show that pathologists take longer to classify the tumor on original images with respect to the normalized ones (77.9 s vs. 69.9 s for P1 and 52.7 s vs. 37.4 s for P2). From these data, we can state that the time for diagnosis is directly related to the quality of the tissue slide, and the normalization process, by improving color quality, enables faster diagnosis. Finally, pathologists were asked to assess their confidence on diagnosis. For both pathologists, mean values of confidence increase for normalized slides with respect to original ones (i.e., 8.2/10 vs. 7.0/10 for P1 and 7.2/10 vs. 6.2/10 for P2). We can therefore assume that confidence in diagnosis is subjected to less variability due to the stain normalization process and that confidence is strongly related to stain quality. Paired analysis (original vs. normalized, for each pathologist) of the time required for diagnosis and the confidence of diagnosis similarly reveal interesting data. The time required for diagnosis is significantly shorter on normalized images compared to original images, for both pathologists ($p < 0.001$). This can be explained by the greater clarity of diagnostically important details in normalized images. Examination of low-quality original images proved difficult for both pathologists, who reported difficulties in assessing details such as presence and size of nucleoli, chromatin quality and cytoplasm shape and color. In some cases, the architecture proved too difficult to examine, leading on one side to overestimation of benign mimickers of cancer, and on the other side to underestimation of minor foci of small/malformed Gleason pattern 4 glands, which were missed in the original images but were correctly identified in normalized ones, as shown in Figure 6.

For the same reasons, the confidence in diagnosis was found to be significantly higher in normalized images compared to original images, for both pathologists ($p < 0.001$).

This study has different limitations related to the number of WSIs analyzed (n = 93) and pathologists involved (n = 2), which used a bioimage software (QuPath [33]) for research purposes, not for clinical use. Another limitation consists in the dataset composition, which originated from the same medical center and was selected to cover the high variability of color staining for the evaluation of the normalization process, instead of reporting consecutive cases belonging to same batches to represent typical clinical practice. In addition, the choice of target image for the normalization process can be investigated in a future work. We plan to employ further efforts to collect a larger number of WSIs and involve more experts with different years of experience in digital pathology with the possibility to select a specific target image for each of them. The choice of target image may depend on the pathologist's opinion according to the clinical expertise and with the objective of improving the diagnosis. We also intend to extend the study to more tissues/pathologies (e.g., breast cancer; colon cancer; etc.) and expand the use of the normalization process on different hematoxylin (e.g., Gill 1,2,3; Harris; Mayer; etc.) and eosin (e.g., watery vs. alcoholic) solutions and other histological staining techniques, e.g., periodic acid Schiff (PAS) and trichrome staining.

5. Conclusions

In this study, we analyzed the impact of the stain normalization process to the diagnosis of prostate cancer from a clinical perspective. In summary, we have shown how a digital color normalization procedure can support the physician during diagnosis formulation, reducing analysis time and increasing the confidence of diagnosis.

Author Contributions: Conceptualization, M.S. (Massimo Salvi), O.P., F.M. and F.F.; methodology, M.S. (Massimo Salvi), M.S. (Manuela Scotto), N.M. and A.M.; data curation and validation, A.C., D.B. and F.F.; statistical analysis, M.S. (Manuela Scotto); technical support, A.M.; writing—original draft preparation, M.S. (Manuela Scotto) and N.M.; writing—review and editing, M.S. (Massimo Salvi), A.C., D.B., O.P. and N.M.; supervision, F.M. and F.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: M. Salvi, O. Pennisi and F. Molinari are equity holders in AEQUIP S.r.l., Turin, Italy. The remaining authors declare no conflict of interest regarding the publication of this article.

References

1. Ferlay, J.; Ervik, M.; Lam, F.; Colombet, M.; Mery, L.; Piñeros, M.; Znaor, A.; Soerjomataram, I.; Bray, F. Global Cancer Observatory: Cancer Today. *Lyon. Fr. Int. Agency Res. Cancer* **2018**, *3*, 2019.
2. WHO. Classification of Tumours Editorial Board. In *Urinary and Male Genital Tumours*, 5th ed.; World Health Organization: Geneva, Switzerland, 2022; Volume 8, ISBN 978-92-832-4512-4.
3. Epstein, J.I.; Egevad, L.; Amin, M.B.; Delahunt, B.; Srigley, J.R.; Humphrey, P.A. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *Am. J. Surg. Pathol.* **2016**, *40*, 244–252. [CrossRef] [PubMed]
4. Gleason, D.F. Classification of Prostatic Carcinomas. *Cancer Chemother. Rep.* **1966**, *50*, 125–128. [PubMed]
5. Kirmiz, S.; Qi, J.; Babitz, S.K.; Linsell, S.; Denton, B.; Singh, K.; Auffenberg, G.; Montie, J.E.; Lane, B.R. Grade Groups Provide Improved Predictions of Pathological and Early Oncologic Outcomes Compared with Gleason Score Risk Groups. *J. Urol.* **2019**, *201*, 278–283. [CrossRef] [PubMed]
6. Knipper, S.; Karakiewicz, P.I.; Heinze, A.; Preisser, F.; Steuber, T.; Huland, H.; Graefen, M.; Tilki, D. Definition of High-Risk Prostate Cancer Impacts Oncological Outcomes after Radical Prostatectomy. *Urol. Oncol. Semin. Orig. Investig.* **2020**, *38*, 184–190. [CrossRef]

7. Leong, J.Y.; Herrera-Caceres, J.O.; Goldberg, H.; Tham, E.; Teplitsky, S.; Gomella, L.G.; Trabulsi, E.J.; Lallas, C.D.; Fleshner, N.E.; Tilki, D.; et al. Questioning the Status Quo: Should Gleason Grade Group 1 Prostate Cancer Be Considered a "Negative Core" in Pre-Radical Prostatectomy Risk Nomograms? An International Multicenter Analysis. *Urology* **2020**, *137*, 102–107. [CrossRef]
8. Pukl, M.; Keyes, S.; Keyes, M.; Guillaud, M.; Volavšek, M. Multi-Scale Tissue Architecture Analysis of Favorable-Risk Prostate Cancer: Correlation with Biochemical Recurrence. *Investig. Clin. Urol.* **2020**, *61*, 482–490. [CrossRef]
9. Wenzel, M.; Würnschimmel, C.; Chierigo, F.; Mori, K.; Tian, Z.; Terrone, C.; Shariat, S.F.; Saad, F.; Tilki, D.; Graefen, M.; et al. Pattern of Biopsy Gleason Grade Group 5 (4 + 5 vs. 5 + 4 vs. 5 + 5) Predicts Survival After Radical Prostatectomy or External Beam Radiation Therapy. *Eur. Urol. Focus* **2022**, *8*, 710–717. [CrossRef]
10. Preisser, F.; Wang, N.; Abrams-Pompe, R.S.; Chun, F.K.-H.; Graefen, M.; Huland, H.; Tilki, D. Oncologic Outcomes of Organ-Confined Gleason Grade Group 4-5 Prostate Cancer after Radical Prostatectomy. *Urol. Oncol.* **2022**, *40*, 161.e9–161.e14. [CrossRef]
11. Borowsky, A.D.; Glassy, E.F.; Wallace, W.D.; Kallichanda, N.S.; Behling, C.A.; Miller, D.V.; Oswal, H.N.; Feddersen, R.M.; Bakhtar, O.R.; Mendoza, A.E.; et al. Digital Whole Slide Imaging Compared With Light Microscopy for Primary Diagnosis in Surgical Pathology. *Arch. Pathol. Lab. Med.* **2020**, *144*, 1245–1253. [CrossRef]
12. Camparo, P.; Egevad, L.; Algaba, F.; Berney, D.M.; Boccon-Gibod, L.; Compérat, E.; Evans, A.J.; Grobholz, R.; Kristiansen, G.; Langner, C.; et al. Utility of Whole Slide Imaging and Virtual Microscopy in Prostate Pathology. *APMIS* **2012**, *120*, 298–304. [CrossRef] [PubMed]
13. Mukhopadhyay, S.; Feldman, M.D.; Abels, E.; Ashfaq, R.; Beltaifa, S.; Cacciabeve, N.G.; Cathro, H.P.; Cheng, L.; Cooper, K.; Dickey, G.E.; et al. Whole Slide Imaging Versus Microscopy for Primary Diagnosis in Surgical Pathology: A Multicenter Blinded Randomized Noninferiority Study of 1992 Cases (Pivotal Study). *Am. J. Surg. Pathol.* **2018**, *42*, 39–52. [CrossRef] [PubMed]
14. Oyama, T.; Allsbrook, W.C.J.; Kurokawa, K.; Matsuda, H.; Segawa, A.; Sano, T.; Suzuki, K.; Epstein, J.I. A Comparison of Interobserver Reproducibility of Gleason Grading of Prostatic Carcinoma in Japan and the United States. *Arch. Pathol. Lab. Med.* **2005**, *129*, 1004–1010. [CrossRef] [PubMed]
15. Allsbrook, W.C.; Mangold, K.A.; Johnson, M.H.; Lane, R.B.; Lane, C.G.; Amin, M.B.; Bostwick, D.G.; Humphrey, P.A.; Jones, E.C.; Reuter, V.E.; et al. Interobserver Reproducibility of Gleason Grading of Prostatic Carcinoma: Urologic Pathologists. *Hum. Pathol.* **2001**, *32*, 74–80. [CrossRef] [PubMed]
16. Ozkan, T.A.; Eruyar, A.T.; Cebeci, O.O.; Memik, O.; Ozcan, L.; Kuskonmaz, I. Interobserver Variability in Gleason Histological Grading of Prostate Cancer. *Scand. J. Urol.* **2016**, *50*, 420–424. [CrossRef]
17. Fraggetta, F.; L'Imperio, V.; Ameisen, D.; Carvalho, R.; Leh, S.; Kiehl, T.-R.; Serbanescu, M.; Racoceanu, D.; Della Mea, V.; Polonia, A.; et al. Best Practice Recommendations for the Implementation of a Digital Pathology Workflow in the Anatomic Pathology Laboratory by the European Society of Digital and Integrative Pathology (ESDIP). *Diagnostics* **2021**, *11*, 2167. [CrossRef]
18. Rathore, S.; Iftikhar, M.A.; Chaddad, A.; Niazi, T.; Karasic, T.; Bilello, M. Segmentation and Grade Prediction of Colon Cancer Digital Pathology Images Across Multiple Institutions. *Cancers* **2019**, *11*, 1700. [CrossRef]
19. Jang, H.-J.; Song, I.-H.; Lee, S.-H. Deep Learning for Automatic Subclassification of Gastric Carcinoma Using Whole-Slide Histopathology Images. *Cancers* **2021**, *13*, 1811. [CrossRef]
20. Lyon, H.O.; De Leenheer, A.P.; Horobin, R.W.; Lambert, W.E.; Schulte, E.K.; Van Liedekerke, B.; Wittekind, D.H. Standardization of Reagents and Methods Used in Cytological and Histological Practice with Emphasis on Dyes, Stains and Chromogenic Reagents. *Histochem. J.* **1994**, *26*, 533–544. [CrossRef]
21. Roy, S.; kumar Jain, A.; Lal, S.; Kini, J. A Study about Color Normalization Methods for Histopathology Images. *Micron* **2018**, *114*, 42–61. [CrossRef]
22. Salehi, P.; Chalechale, A. Pix2pix-Based Stain-to-Stain Translation: A Solution for Robust Stain Normalization in Histopathology Images Analysis. In Proceedings of the 2020 International Conference on Machine Vision and Image Processing (MVIP), Qom, Iran, 18–20 February 2020; pp. 1–7.
23. Gurcan, M.N.; Boucheron, L.E.; Can, A.; Madabhushi, A.; Rajpoot, N.M.; Yener, B. Histopathological Image Analysis: A Review. *IEEE Rev. Biomed. Eng.* **2009**, *2*, 147–171. [CrossRef] [PubMed]
24. Taylor, C.R. Immunohistochemistry in Surgical Pathology: Principles and Practice. In *Histopathology: Methods and Protocols*; Day, C.E., Ed.; Springer: New York, NY, USA, 2014; pp. 81–109. ISBN 978-1-4939-1050-2.
25. Michielli, N.; Caputo, A.; Scotto, M.; Mogetta, A.; Pennisi, O.A.M.; Molinari, F.; Balmativola, D.; Bosco, M.; Gambella, A.; Metovic, J.; et al. Stain Normalization in Digital Pathology: Clinical Multi-Center Evaluation of Image Quality. *J. Pathol. Inform.* **2022**, *13*, 100145. [CrossRef] [PubMed]
26. Salvi, M.; Michielli, N.; Molinari, F. Stain Color Adaptive Normalization (SCAN) Algorithm: Separation and Standardization of Histological Stains in Digital Pathology. *Comput. Methods Programs Biomed.* **2020**, *193*, 105506. [CrossRef] [PubMed]
27. Pontalba, J.T.; Gwynne-Timothy, T.; David, E.; Jakate, K.; Androutsos, D.; Khademi, A. Assessing the Impact of Color Normalization in Convolutional Neural Network-Based Nuclei Segmentation Frameworks. *Front. Bioeng. Biotechnol.* **2019**, *7*, 300. [CrossRef]
28. Swiderska-Chadaj, Z.; de Bel, T.; Blanchet, L.; Baidoshvili, A.; Vossen, D.; van der Laak, J.; Litjens, G. Impact of Rescanning and Normalization on Convolutional Neural Network Performance in Multi-Center, Whole-Slide Classification of Prostate Cancer. *Sci. Rep.* **2020**, *10*, 14398. [CrossRef]
29. Salvi, M.; Bosco, M.; Molinaro, L.; Gambella, A.; Papotti, M.; Acharya, U.R.; Molinari, F. A Hybrid Deep Learning Approach for Gland Segmentation in Prostate Histopathological Images. *Artif. Intell. Med.* **2021**, *115*, 102076. [CrossRef]

30. Janowczyk, A.; Madabhushi, A. Deep Learning for Digital Pathology Image Analysis: A Comprehensive Tutorial with Selected Use Cases. *J. Pathol. Inform.* **2016**, *7*, 29. [CrossRef]
31. Salvi, M.; Molinari, F.; Acharya, U.R.; Molinaro, L.; Meiburger, K.M. Impact of Stain Normalization and Patch Selection on the Performance of Convolutional Neural Networks in Histological Breast and Prostate Cancer Classification. *Comput. Methods Progr. Biomed. Updat.* **2021**, *1*, 100004. [CrossRef]
32. Liu, F.; Hernandez-Cabronero, M.; Sanchez, V.; Marcellin, M.W.; Bilgin, A. The Current Role of Image Compression Standards in Medical Imaging. *Information* **2017**, *8*, 131. [CrossRef]
33. Bankhead, P.; Loughrey, M.B.; Fernández, J.A.; Dombrowski, Y.; McArt, D.G.; Dunne, P.D.; McQuaid, S.; Gray, R.T.; Murray, L.J.; Coleman, H.G.; et al. QuPath: Open Source Software for Digital Pathology Image Analysis. *Sci. Rep.* **2017**, *7*, 16878. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.