POLITECNICO DI TORINO
Repository ISTITUZIONALE

Not all stains are made equal: impact of stain normalization on prostate cancer diagnosis

Original
Not all stains are made equal: impact of stain normalization on prostate cancer diagnosis / Cazzaniga, G., Caputo, A., L'Imperio, V., Gibilisco, F., Scotto, M., Pennisi, O.M.A., Michielli, N., Mogetta, A., Molinari, F., Fraggetta, F., Salvi, M.. - In: PATHOLOGICA. - ISSN 1591-951X. - 116:5(2024), pp. 285-294. [10.32074/1591-951X-1008]

Availability:
This version is available at: 11583/2996184 since: 2025-01-03T15:47:49Z

Publisher:
Pacini Editore

Published
DOI:10.32074/1591-951X-1008

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)

18 November 2026

PATHOLOGICA 2024;116:285-294;
DOI: 10.32074/1591-951X-1008

Original article

Not all stains are made equal: impact of stain normalization on prostate cancer diagnosis

Giorgio Cazzaniga[1*], Alessandro Caputo[2*], Vincenzo L'Imperio[1], Fabio Gibilisco[3], Manuela Scotto[4], Orazio Maria Antonino Pennisi[5], Nicola Michielli[4], Alessandro Mogetta[4], Filippo Molinari[4], Filippo Fraggetta[3], Massimo Salvi[4]

[1] Department of Medicine and Surgery, Pathology, IRCCS Fondazione San Gerardo dei Tintori, University of Milano-Bicocca, Italy; [2] Department of Medicine and Surgery, University Hospital of Salerno, Baronissi, Italy; [3]UOC Anatomia Patologica, ASP Catania P.O. "Gravina", Caltagirone, Italy; [4] Biolab, PoliTo[BIO]Med Lab, Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy; [5] Technology Transfer and Industrial Liaison Department, Politecnico di Torino, Turin, Italy
**Equally contributing author*

Summary

Objective. Stain normalization is a technique used to standardize the color appearance of digital whole slide images (WSIs). This study aimed to assess the impact of digital stain normalization on prostate cancer diagnosis by pathologists.

Methods. A multi-institutional board of four pathologists evaluated 407 hematoxylin and eosin (H&E) prostate WSIs before and after stain normalization. The presence/absence of prostate adenocarcinoma, the Grade Groups as well as color quality perception and time required for diagnosis were recorded.

Results. After normalization, color quality improved significantly for all pathologists (median scores increased from 4-6 to 7-8/10). Average diagnosis time decreased from 50s to 35s ($p < 0.001$). Inter-pathologist reproducibility for Gleason risk group showed a fair to good level of agreement, with an improvement after normalization.

Conclusions. Stain normalization enhanced pathologists' diagnosis of prostate cancer by improving color standardization, reducing diagnosis time, and increasing inter-observer reproducibility. These findings highlight the potential of stain normalization to improve accuracy and efficiency in digital pathology.

Key words: digital pathology, prostate cancer, stain normalization, gleason score, H&E staining

Received: April 8, 2024
Accepted: June 28, 2024

Correspondence
Massimo Salvi
E-mail: massimo.salvi@polito.it

How to cite this article: Cazzaniga G, Caputo A, L'Imperio V, et al. Not all Stains are Made Equal: Impact of Stain Normalization on Prostate Cancer Diagnosis. Pathologica 2024;116:285-294. https://doi.org/10.32074/1591-951X-1008

© Copyright by Società Italiana di Anatomia Patologica e Citopatologia Diagnostica, Divisione Italiana della International Academy of Pathology

OPEN ACCESS

This is an open access journal distributed in accordance with the CC-BY-NC-ND (Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International) license: the work can be used by mentioning the author and the license, but only for non-commercial purposes and only in the original version. For further information: https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en

Introduction

Digital pathology (DP) adoption is transforming pathology practice as laboratories transition to digitized workflows [1]. Realizing the full potential of this transition requires optimizing each step, including automated tracking, streamlined workflows [2], and proper validation of whole slide images (WSIs) [3]. However, variability in tissue preparation across labs can impede this shift by affecting WSI quality [4]. This variability poses challenges for telepathology consultations and digitizing archived slides with faded stains [5].

Digital stain normalization tools have been developed to address variable staining quality by standardizing appearance across digital images to enable consistent interpretation [6]. Using normalization during preprocessing has shown positive impacts by improving model generalization for various AI algorithms [7,8]. In our recent review [9], we identify data

augmentation and stain normalization as powerful techniques to significantly improve the performance and consistency of deep learning models in digital pathology. A preliminary study on stain normalization also suggested benefits for pathologists' perception of WSIs [10]. However, the impact of normalization in clinical diagnosis settings remains unclear. One area requiring further analysis is interobserver concordance in Gleason grading, which exhibits notable variability [11] with implications for prostate cancer management. Therefore, this study aims to comprehensively evaluate the impact of stain normalization on a large set of WSIs for prostate cancer diagnosis. Specifically, this study assesses how normalization could influence pathologists' confidence, reproducibility, grading, and annotations. A multi-institutional panel of pathologists examined images with and without normalization to evaluate its potential to improve accuracy, efficiency, and reliability in digital pathology interpretation.

Materials and methods

Datasets

The dataset consisted of 407 hematoxylin and eosin (H&E) stained prostate WSIs obtained from The Cancer Genome Atlas (TCGA) program. The WSIs in this study were heterogeneous in terms of magnification and scanning devices. 253 WSIs were captured at 20x magnification (0.467 µm/pixel) and 154 WSIs were captured at 40x magnification (0.233 µm/pixel). In terms of scanning devices, 203 WSIs were captured using Leica scanners in.svs format, 151 WSIs were captured using Hamamatsu scanners in ndpi format, and 53 WSIs were captured using 3DHISTECH scanners in.mrxs format. To address this variability, the dataset was partitioned into four balanced batches considering magnification, scanning device and stain variability. Each batch was designed to contain a similar distribution of images from different centers and scanning devices, as well as a comparable range of staining quality. This approach aimed to create 'consistent' datasets, meaning that each batch had approximately the same level of variability in terms of image sources and staining characteristics. The batches were then combined to create the final datasets assigned to each pathologist for evaluation. This strategy was employed to mitigate potential bias and ensure that no single batch was significantly easier or more challenging to analyze than the others. The design of the entire study is reported in Figure 1. The study was conducted in accordance with the Declaration of Helsinki, the WSI used were obtained from the publicly available dataset of TCGA not requiring additional Local Ethical Board approval.

Stain Normalization

Stain normalization is a technique used to adjust and align the color appearance of histological slides with

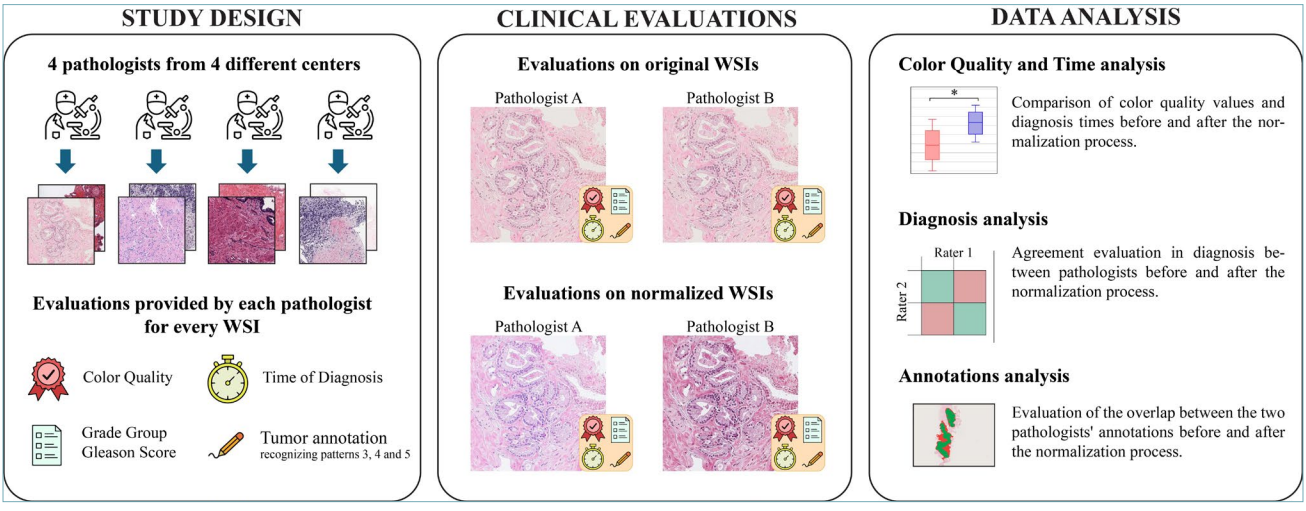


Figure 1. Study design showing division of multicentric WSI collection into batches presented to pathologists from different institutions. Pathologists evaluated tumor annotations, Gleason risk group, diagnosis time, and color quality for original and normalized images after a washout period. Each WSI received assessments from two pathologists before and after normalization. Analysis focused on color, diagnosis time, agreement between pathologists regarding diagnosis and tumor localization.

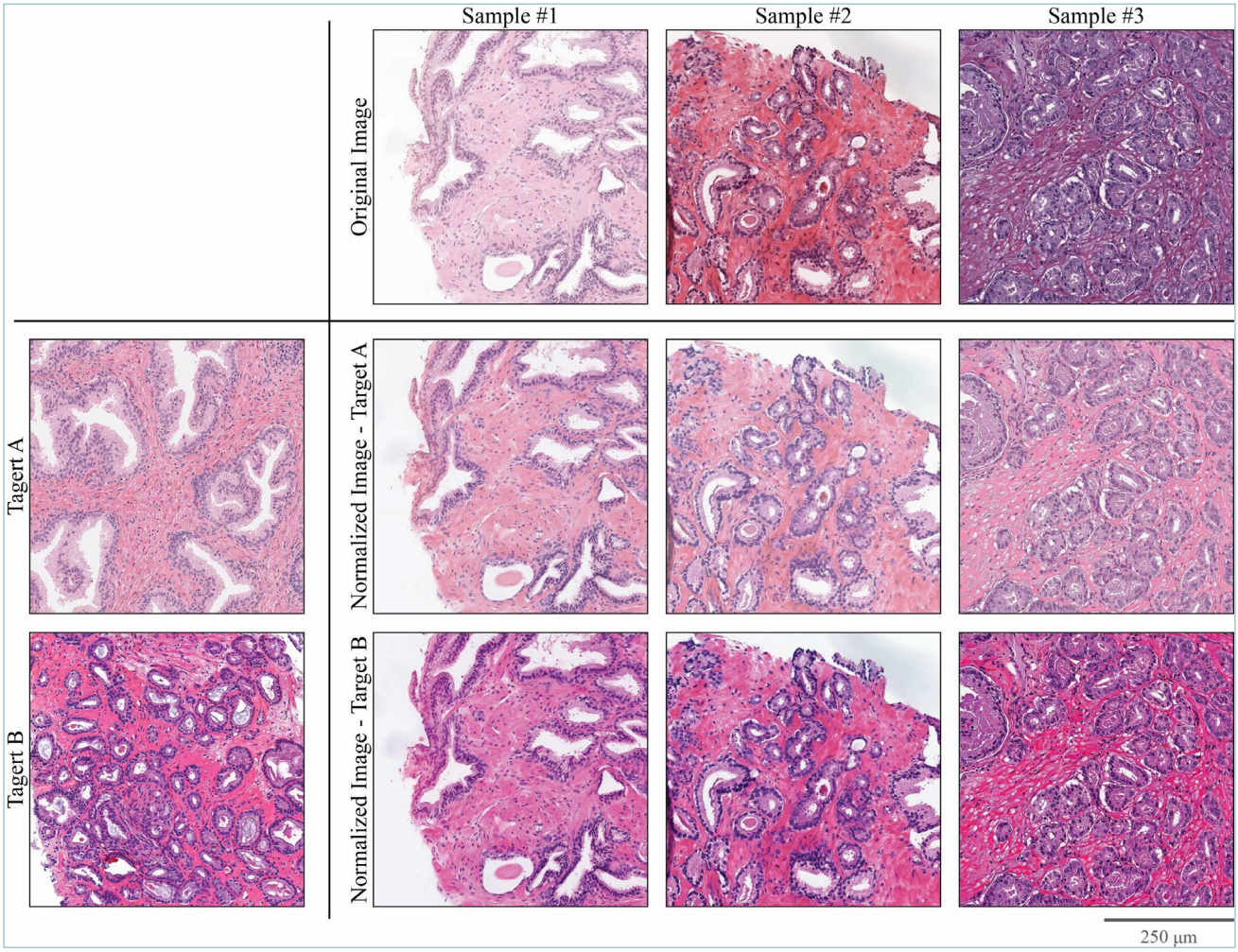


Figure 2. Example of stain normalization using different target images chosen by pathologists. The algorithm starts by estimating the H&E stain colors in each image, then adjusting the image appearance to align with target color values, while preserving local tissue contrast. The normalization process alters the color characteristics of the image while maintaining the structural integrity of the original content. Normalization took approximately 3 minutes per WSI.

respect to a target image, thus reducing the impact of staining variations. To carry out the stain normalization process, an improved version of the STAINS (STAndardIzation & Normalization of histological Slides) tool [12] was utilized. Briefly, the digital stain normalization process involves standardizing the appearance of histology images by removing inter-image variability in stain conditions while preserving biologically relevant features. Before STAINS normalization, tissue detection is performed to identify relevant tissue. Then, color parameters for the transformation are computed on the entire WSI. Finally, normalization is performed, applying the same transformation to every tile.

Each pathologist was asked to provide a reference WSI that they subjectively judged as optimally stained. These target images were selected at a resolution of 20x. Stain normalization was then conducted for each pathologist using their chosen reference slide as the target. This approach allowed the pathologists to evaluate the normalized image according to their individual chromatic preferences. Figure 2 presents an example of the normalization process utilizing two different target images. During this process, the color characteristics of the image are modified while ensuring the preservation of structural integrity of the original content. In this regard, a previous study was conducted, showing that this normalization method does not introduce clinically significant artifacts [10].

The stain normalization process was performed on a workstation with 8 cores and 128 GB of RAM. The normalization time is consistent across pathologists, with mean computational times ranging from 3.22 to 3.47 minutes per WSI, depending on the pathologist. These processing times are based on an average WSI area of around 6.10 × 109 pixels.

Pathologists WSIs evaluation

A panel consisting of four pathologists (referred to as path #1, #2, #3, and #4) from different institutions and with different years of experience participated in the evaluation of the WSIs. The evaluation process initially consisted of assessing the original images, followed by re-evaluating the normalized versions of the same slides after a 3-month washout period. To ensure a comprehensive evaluation, two batches of WSIs were randomly assigned to each pathologist. Furthermore, each batch was evaluated by two different pathologists, as indicated in Table SI. To maintain anonymity and prevent tracing back to the original WSIs, the normalized slides were anonymized during the second round of evaluation. For each WSI, the pathologists were provided with an Excel spreadsheet (Microsoft, Redmont, USA) containing the following data (Fig. S1):

1 Diagnosis: presence/absence of invasive acinar prostate adenocarcinoma, with specification of the Gleason Score and Grade Group, as per the most recent WHO classification [13];
2 Time required for diagnosis, from the slide opening to the diagnosis formulation (in seconds);
3 Perceived color quality, on a scale from 1 (lowest) to 10 (highest quality)
4 QuPath [14] (version 0.4.3) was used to visualize the WSIs (Fig. S2). After the visual evaluation, pathologists performed annotations of the tumor area, with color-coded discrimination of Gleason patterns 3, 4 and 5, then extracted as geojson format.

Data analysis

Normally distributed continuous variables are expressed as mean ± standard deviation (SD), while non-normally distributed variables are expressed as median and interquartile range (IQR). Intra-pathologist variability and inter-pathologist reproducibility between original and normalized slides for final diagnosis and Gleason risk group were measured using Cohen's kappa coefficient. Perceived image quality and time required for diagnosis were compared between original and normalized WSIs using a paired t-test. Additionally, comparability of tumor area annotations and prevalent Gleason pattern before and after normalization was assessed using the Dice Similarity Coefficient (DSC) at a magnification of 5x using Python (version 3.7.11). Statistical analysis was performed in Matlab (MathWorks, USA) with a significance level set at 0.001.

Results

Color quality and time to diagnosis after normalization

Table I shows the color quality scores and diagnosis times for each pathologist on the original and normalized WSIs. Across all image batches, normalization led to a noticeable decrease in average diagnosis time of approximately 13 seconds per WSI. Additionally, normalization improved the perceived color quality scores for all pathologists, as indicated by increased median scores (Fig. 3). This trend was confirmed in a sub-analysis of images categorized by original low, medium, and high color quality (Fig. S3).
Having two pathologists score each image enabled direct comparison of color quality before and after normalization (Tab. II). Median scores increased significantly across all batches after normalization ($p < 0.001$). Analyzing the distributions of original ver-

Table I. Color quality and diagnosis time metrics for each individual pathologist on original and normalized WSIs. Mean, standard deviation, and median values are shown. Asterisks indicate statistically significant increases in color quality and decreases in diagnosis time after normalization (paired t-test, $p < 0.001$).

Pathologist	Images	Color Quality		Time to diagnosis	
		Mean ± Std. Dev.	*Median*	*Mean ± Std. Dev.*	*Median*
Path #1	Original	5.28 ± 2.25*	6	111.72 s ± 65.44 s*	100.00 s
	Normalized	7.77 ± 1.61*	8	75.58 s ± 36.98 s*	70.00 s
Path #2	Original	4.05 ± 2.05*	4	27.76 s ± 16.99 s*	23.00 s
	Normalized	6.98 ± 1.59*	7	17.86 s ± 12.13 s*	15.00 s
Path #3	Original	5.27 ± 2.39*	6	36.47 s ± 20.83 s*	30.00 s
	Normalized	6.84 ± 1.74*	7	31.14 s ± 12.87s*	30.00 s
Path #4	Original	4.84 ± 1.89*	5	27.83 s ± 16.27 s*	23.00 s
	Normalized	7.28 ± 2.40*	7.5	23.83 s ± 12.82 s*	20.00 s

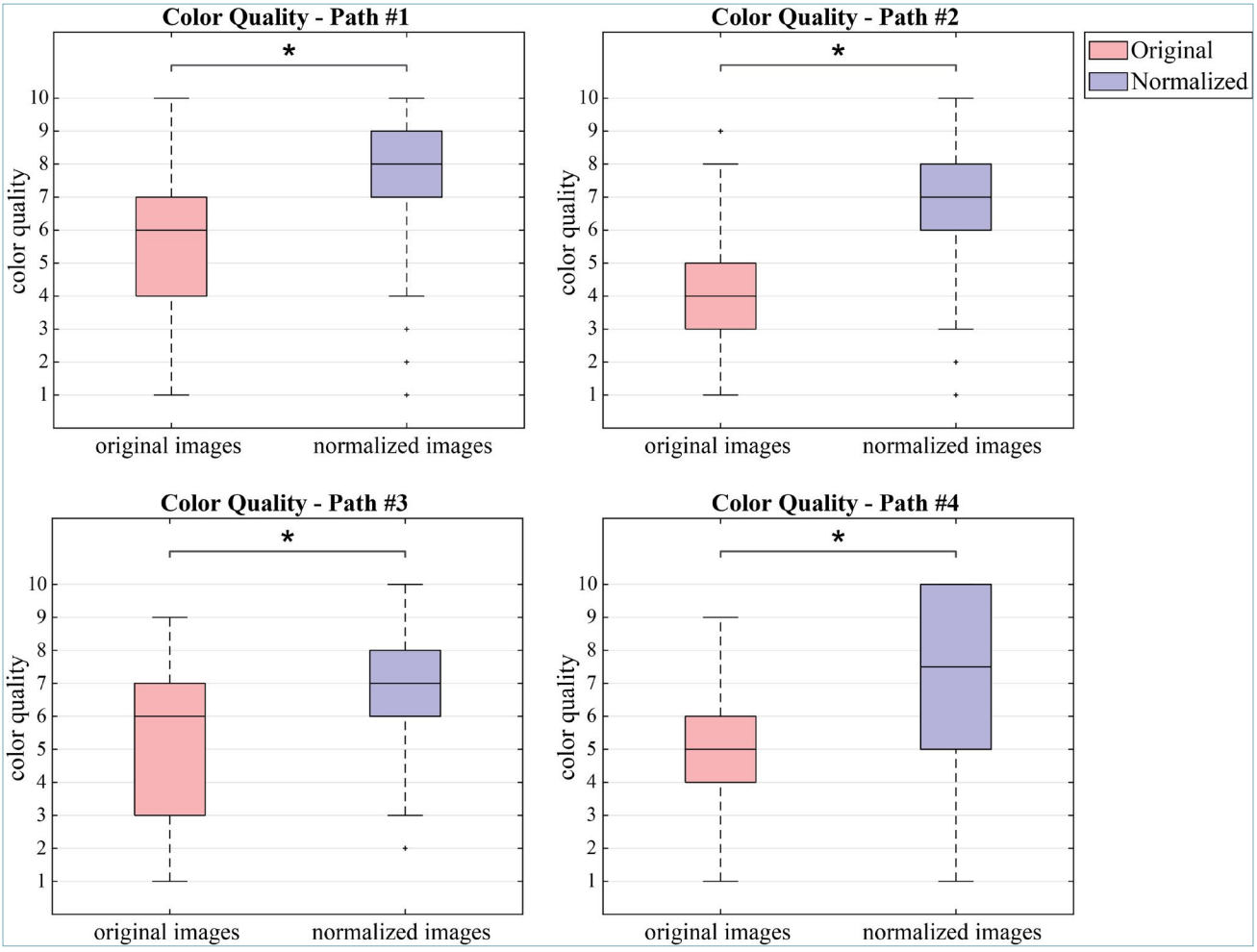


Figure 3. Distributions of color quality scores by each pathologist for original (red) and normalized (blue) WSIs. The asterisk indicates a statistically significant increase in color quality after normalization ($p < 0.001$).

Table II. Inter-pathologist agreement on Gleason risk groups and similarity of tumor annotations, before and after normalization. Cohen's kappa was calculated for Gleason risk groups, defined by grouping grade groups into 4 categories: no tumor, grade group 1, grade groups 2-3, and grade groups 4-5. The mean Dice similarity coefficient (DSC) compared tumor regions and individual Gleason pattern areas annotated before and after normalization. The asterisk indicates statistically significant increase in agreement after normalization ($p < 0.001$).

Batch	Images	Color Quality Mean ± SD	Cohen's kappa for Risk Group	DSC on tumor annotations Mean (No. WSIs)			
				Tumor	*Pattern 3*	*Pattern 4*	*Pattern 5*
Batch 1	Original	4.95 ± 1.90*	0.6782	0.702 (82)	0.537 (51)	0.452 (52)	0.246 (14)
	Normalized	7.48 ± 1.61*	0.6609	0.733 (82)	0.589 (51)	0.431 (52)	0.165 (14)
Batch 2	Original	4.77 ± 1.83*	0.5174	0.676 (87)	0.286 (48)	0.467 (61)	0.125 (11)
	Normalized	7.40 ± 1.49*	0.5539	0.725 (87)	0.295 (48)	0.477 (61)	0.170 (11)
Batch 3	Original	4.61 ± 1.94*	0.4762	0.665 (85)	0.231 (61)	0.384 (61)	0.036 (6)
	Normalized	6.96 ± 1.45*	0.4964	0.696 (85)	0.229 (61)	0.475 (61)	0.049 (6)
Batch 4	Original	5.11 ± 1.92*	0.6363	0.661 (80)	0.409 (59)	0.376 (56)	0.070 (9)
	Normalized	7.03 ± 1.91*	0.6853	0.749 (80)	0.506 (59)	0.397 (56)	0.019 (9)

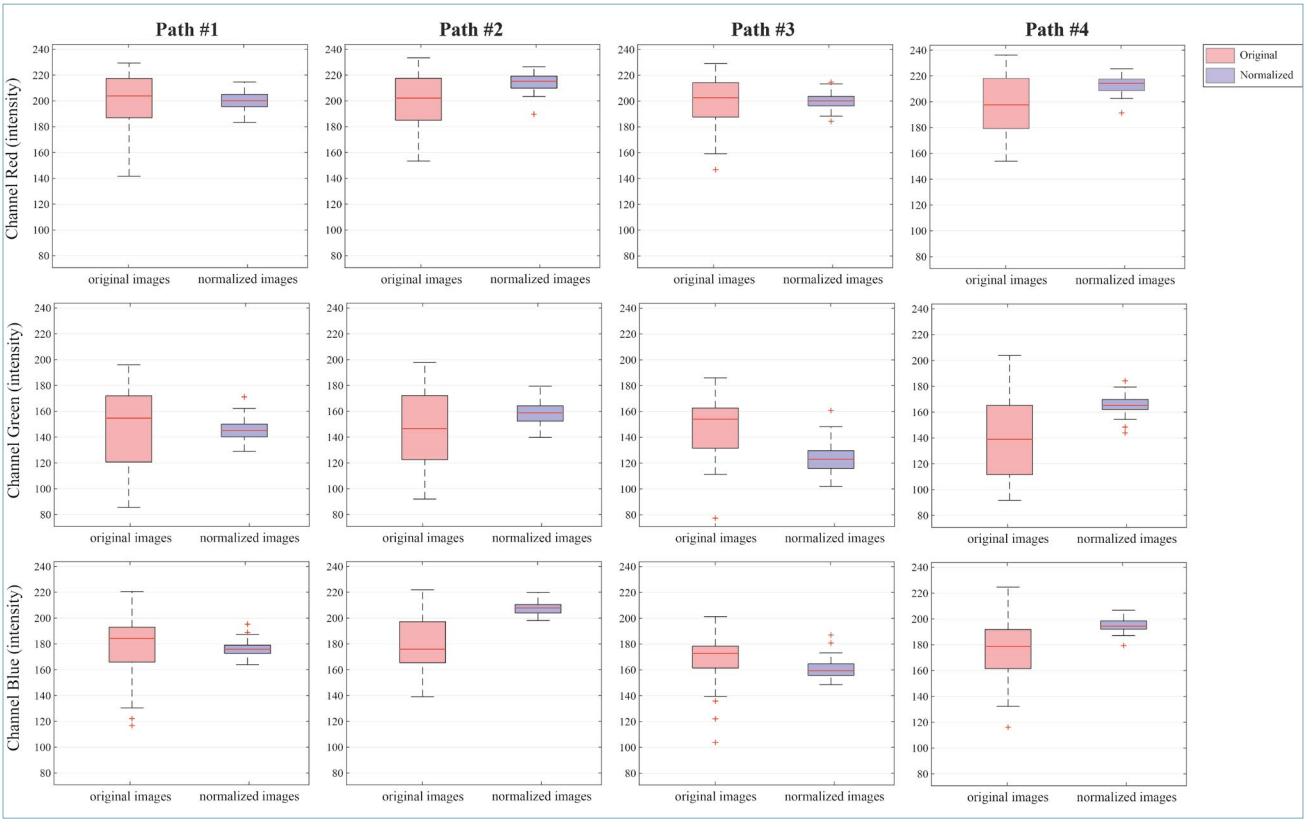


Figure 4. Boxplots comparing the distribution of RGB brightness features for original and normalized WSIs across each pathologist. The normalized WSIs show reduced variability and improved clustering around optimal brightness levels compared to the original WSIs.

sus normalized quality scores in increasing quality groups reinforced this trend (Fig. S4). The improved image quality from normalization reduced the average diagnosis time significantly ($p < 0.001$).

Quantitative analysis of stain normalization

To quantitatively assess the impact of stain normalization on the WSI dataset, we used HistoQC tool (version 2.1) to extract brightness features from the red, green, and blue color channels of the original and normalized WSIs. The distribution of these brightness features was compared before and after normalization using boxplots (Fig. 4).

The boxplots demonstrate that stain normalization effectively reduces the variability in brightness across the WSIs. The original WSIs exhibit a wider spread of feature values, indicating high variability in staining. In contrast, the normalized WSIs have a tighter distribution of brightness values clustered around the optimal levels for each color channel.

Normalization impact on cancer diagnosis

We evaluated if normalization affected the pathologists' evaluation of cancer presence and assignment of risk groups. Risk groups were stratified into 4 levels: no tumor, grade group 1, grade groups 2-3, and grade groups 4-5 [15]. Cancer detection concordance between the original histology slides and normalized images was almost identical (Fig. S5). Meanwhile, inter-pathologist agreement for assigning the final risk group within each batch improved slightly after normalization (Tab. II). Agreement ranged from fair to good (0.48-0.68 for original WSIs, and 0.50-0.69 after normalization). In three out of four batches, normalization led to a slight increase in agreement between pathologists. Only one batch showed a minimal decrease from 0.68 to 0.66 originally. Figure 5 exemplifies how normalization improved inter-pathologist consensus for one slide. In the original WSIs, Gleason pattern assignment greatly differed between pathologists. However, after normalization, one pathologist changed the assessment and a higher consensus was achieved.

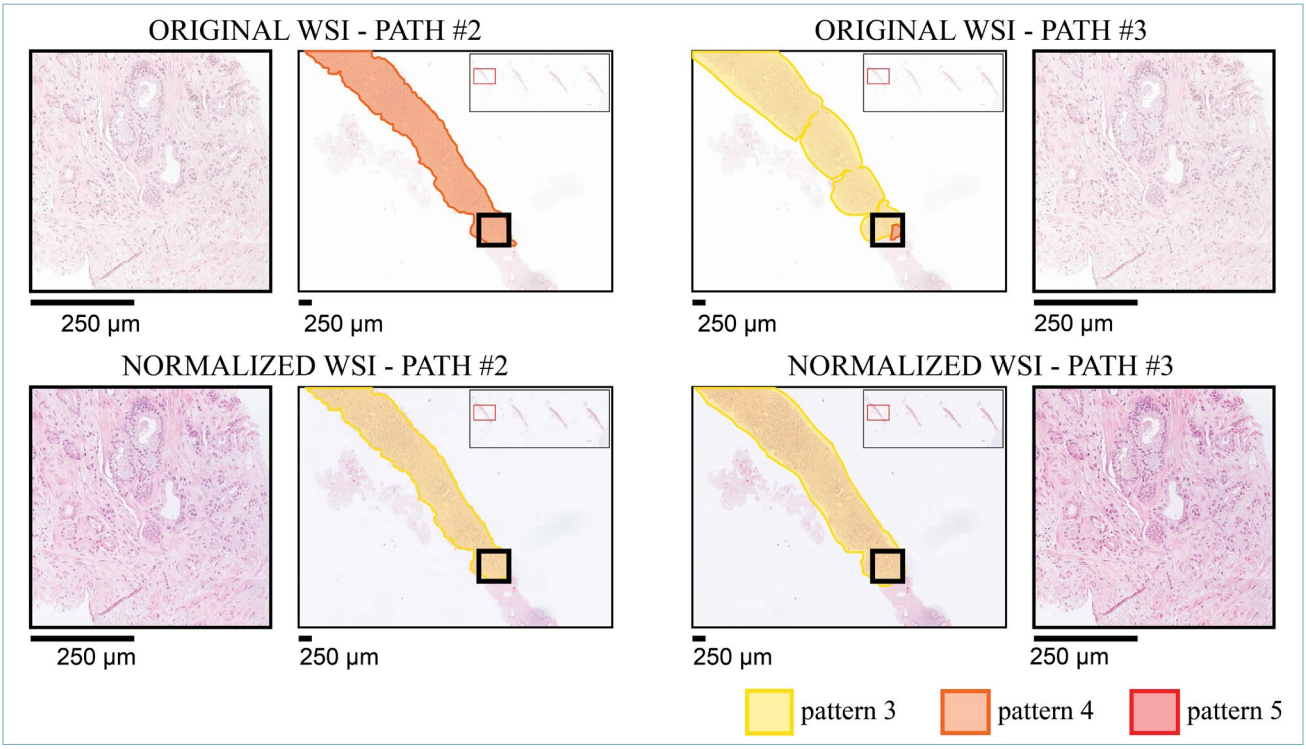


Figure 5. Example case where stain normalization improved inter-pathologist agreement on Gleason patterns (20x magnification). In the original image, there was disagreement between pathologists. Pathologist 1 detected Gleason pattern 4 (orange) while Pathologist 2 identified only pattern 3 (yellow). After normalization, both agreed on final pattern 3.

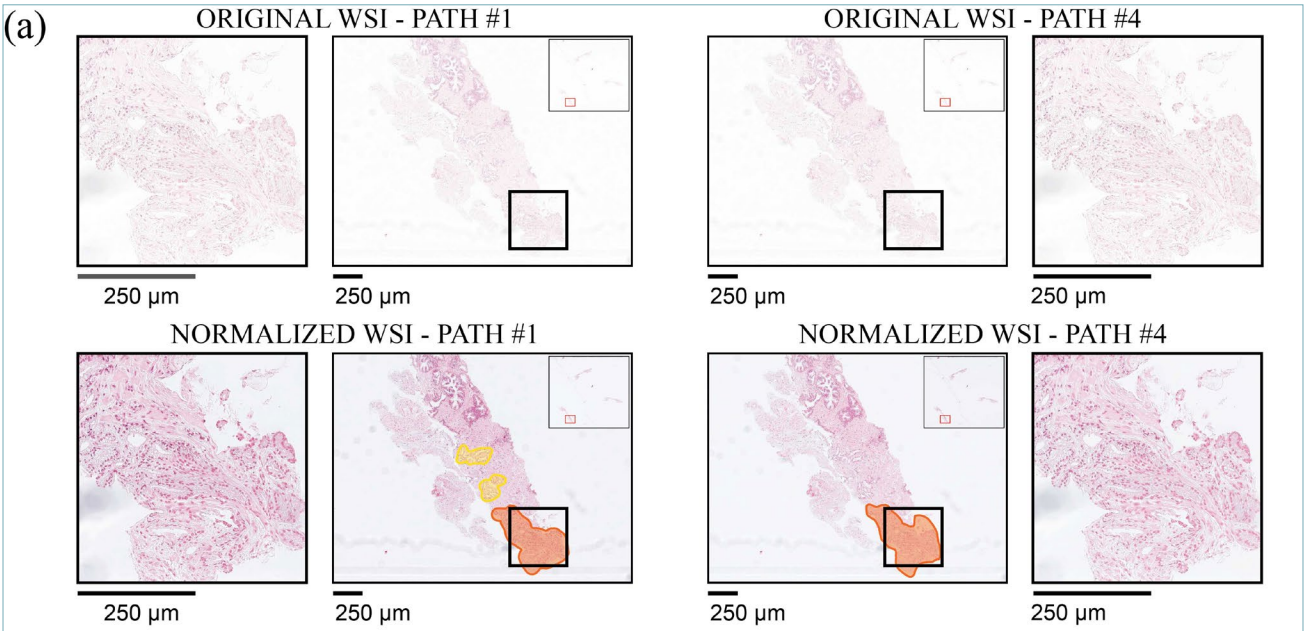


Figure 6. Representative cases showing the effect of normalization on tumor detection and color quality ratings. (a) In the original WSI, both pathologists failed to detect any tumors (top row) and rated color quality as low (1/10). After normalization (bottom row), they agreed on detecting cancerous foci and rated color quality higher (5/10 and 3/10). Normalization also partially revealed an overlapping Gleason pattern 4 region (orange). (b) Originally, Pathologist #1 did not detect the cancerous tissue identified by Pathologist #2 and rated color quality as 1/10. Following normalization, both pathologists detected tumors in the same areas. They also rated perceived color quality much higher (5/10 and 7/10, respectively).

Tumor detection and pattern recognition

We assessed the effect of normalization on inter-pathologist agreement for localizing tumors and annotating their extent (Tab. II). Overall, tumor localization agreement, as measured by DSC, improved from the original to normalized slides (0.66-0.70 originally vs. 0.70-0.75 after normalization). The boxplots in Figure S6 illustrate the DSC distributions for each batch. A statistically significant increase in DSC was observed for three of the four batches (paired t-test, $p < 0.001$). Figure 6A shows examples where both pathologists only detected cancer in the normalized images, with increased color quality scores. Detection and assignment of Gleason pattern 4 also improved across all batches after normalization. Agreement for pattern 3 improved in three of four batches. This suggests normalization enhanced detection and classification of cancerous regions. Figure 6B exemplifies a case where one pathologist missed cancer in the original but identified it in the same location after normalization. No significant improvement occurred for pattern 5, likely due to the small number of cases available.

Discussion

The staining of H&E tissue sections is crucial for accurate pathology interpretation but can be affected by variables including fixation, processing, section thickness, and staining methods [16]. Operator- and laboratory-specific processes, along with factors impacting slide quality, such as fading over time or overly thick sections, require pathologists to adapt to color variations when consulting across institutions [17]. Stain normalization approaches aim to address variability and provide consistency for AI models. However, the benefits for pathologists in clinical practice are less studied, especially with the transition to digital pathology introducing scanning variables [18].

Our study evaluated prostate samples from multiple institutions assessed by four pathologists before and after digital stain normalization. The aim was to document subjective and objective reliability variables and record significant changes in diagnostic concordance using a routine sample in a standard pathology laboratory [19]. Customizing the normalization protocol based on each pathologist's reference slide respected their individual routine and preferences. This avoided imposing a one-size-fits-all standardized approach, which can be highly subjective given pathology's visual nature and variability in individual experience levels.

Perceived color quality improved for all pathologists across batches, with lower quality slides showing the most pronounced enhancement from near unreadable to sufficiently acceptable (Figs. S3-S4). This suggests normalization can mitigate challenges from poor staining and aging, which is important when material is limited. Subjectively improved suitable staining eases the diagnostic burden on pathologists, potentially improving working conditions [20]. With rising examination requests and a pathologist shortage, reduced diagnosis time and improved workflow from normalization provides pivotal advancements [21]. Faster diagnosis addresses time constraints in pathology and enables a more synchronized, consensus-driven environment. Enhanced images improved concordance on Grade risk group and tumor quantification, which are crucial for prostate cancer risk stratification and treatment decisions [22]. Significant gains occurred for patterns 3 and 4, key for stratification, potentially reducing uncertainty [23]. However, no improvement was seen for pattern 5, likely due to limited cases. Better tumor detection provides a more accurate extent estimate and ensures prognostic details are not missed [24].

These effects can have a dramatic impact on clinical practice and patient outcomes, especially when focal biopsy changes lead to substantial alterations in therapy [25]. In a sample like a prostate biopsy, the likelihood of missing a focus of tumor significantly increases with suboptimal staining techniques [26]. Figure 5A shows a compelling case where normalization enabled identifying occult cancer. Initially no tumor was seen due to poor staining, but after normalization, foci of acinar carcinoma became visible.

While our study provides insights into the utility of normalization, there are limitations that should be acknowledged. As there is no single ground truth for histopathological evaluation, improved concordance between pathologists may not necessarily correlate with diagnostic accuracy. Additionally, some original WSIs had poor quality and weak staining, which normalization only partially addressed. In addition, the study did not perfectly replicate real-world clinical practice, as pathologists diagnosed based on a single WSI without supporting information (e.g., additional sections or immunohistochemistry). While normalization addresses chromatic variation, technical defects like out-of-focus areas or tissue folds remain challenging to assess. Future work should emphasize an optimized diagnostic workflow incorporating high-quality WSIs and comprehensive diagnostic approaches to fully evaluate clinical impact.

Our findings indicate the potential value of stain normalization in daily pathological diagnosis. Reducing color variability and enabling pathologists to view their preferred settings can significantly improve interpretability and reliability while preserving the original image. Integrating normalization into digital pathology systems could enhance accuracy, efficiency, and ultimately patient outcomes.

Conclusion

In conclusion, this study demonstrates the positive impact of stain normalization during prostate cancer diagnosis. By reducing staining variability, normalization enhanced color quality, reduced diagnosis time, and improved concordance among pathologists on cancer detection and grading. These benefits can lead to more precise risk assessment and treatment decisions for patients. The integration of stain normalization into digital pathology workflows holds promise for advancing the field by improving the consistency, efficiency, and reliability of pathological evaluations. Overall, these findings indicate the potential of color standardization tools to significantly enhance patient care as digital pathology is increasingly adopted.

Conflict of interest statement

V. L'Imperio received personal fees (as consultant and/or speaker bureau) from Eli Lilly, Roche, Novartis unrelated to the current work. A. Caputo received personal fees (as consultant and/or speaker bureau) from Roche unrelated to the current work. O. A. M. Pennisi, F. Molinari, and M. Salvi are equity holders in AEQUIP S.r.l., Turin, Italy. M. Scotto, N. Michielli, and A. Mogetta also have working relationships with AEQUIP S.r.l., Turin, Italy. Remaining authors declare no competing interests regarding the publication of this article.

Funding

The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Authors' contributions

F.M., O.P. and M.S. performed study concept and design; M.Sc. gathered the data; A.C., G.C., F.G., V.L. provided analysis and interpretation of the data; M.Sc. performed development of methodology, N.M. and A.M. provided technical support; V.L., M.Sc., M.S. wrote the original draft of the paper; A.C., G.C., F.G., F.F. reviewed and revised the paper. All authors read and approved the final paper.

Data availability statement

The Excel spreadsheets compiled by the pathologists and the normalized images are available to the corresponding authors upon reasonable request.

Ethical consideration

The study was conducted in accordance with the Declaration of Helsinki, the WSI used were obtained from the publicly available dataset of TCGA not requiring additional Local Ethical Board approval.

References

[5] Caputo A, L'Imperio V, Merolla F et al. The slow-paced digital evolution of pathology: lights and shadows from a multifaceted board. Pathologica. 2023;115(3):127-136. https://doi.org/10.32074/1591-951X-868.

[6] Fraggetta F, L'Imperio V, Ameisen D, 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 (Basel). 2021;11(11):2167. https://doi.org/10.3390/diagnostics11112167.

[7] Evans AJ, Brown RW, Bui MM et al. Validating whole slide imaging systems for diagnostic purposes in pathology: guideline update from the College of American Pathologists in collaboration with the American Society for Clinical Pathology and the Association for Pathology Informatics. Arch Pathol Lab Med. 2022;146(4):440-450. https://doi.org/10.5858/arpa.2020-0723-CP.

[8] Kanwal N, Pérez-Bueno F, Schmidt, A, et al. The devil is in the details: Whole slide image acquisition and processing for artifacts detection, color variation, and data augmentation: A review. IEEE Access 2022;10:58821-58844. https://doi.org/10.1109/ACCESS.2022.3176091.

[9] Eccher, A. et al. Cost analysis of archives in the pathology laboratories: from safety to management. J Clin Pathol. 2023;76(10):659-663. https://doi.org/10.1136/jcp-2023-209035.

[10] 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. https://doi.org/10.1016/j.cmpb.2020.105506.

[11] Salvi M, Molinari F, Acharya UR, et al. Impact of stain normalization and patch selection on the performance of convolutional neural networks in histological breast and prostate cancer classification. Comput. Methods Programs Biomed Update. 2021;1:100004. https://doi.org/10.1016/j.cmpbup.2021.100004.

[12] Salvi M, Mogetta A, KM Meiburger, et al. Karpinski score under digital investigation: A fully automated segmentation algorithm to identify vascular and stromal injury of donors' kidneys. Electronics (Basel). 2020;9(10):1644. https://doi.org/10.3390/electronics9101644.

[13] Seoni S, Shahini A, Meiburger KM et al. All you need is data preparation: A systematic review of image harmonization techniques in Multi-center/device studies for medical support systems. Comput Methods Programs Biomed. 2024;250:108200. https://doi.org/10.1016/j.cmpb.2024.108200.

[14] Michielli N, Caputo A, Scotto M, et al. Stain normalization in digital pathology: clinical multi-center evaluation of image quality. J Pathol Inform. 2022;13:100145. https://doi.org/10.1016/j.jpi.2022.100145.

[15] Ozkan TA, Eruyar AT, Cebeci OO, et al. Interobserver variability in Gleason histological grading of prostate cancer. Scand J Urol. 2016;50(6):420-424. https://doi.org/10.1080/21681805.2016.1206619.

[16] Salvi M, Caputo A, Balmativola D, et al. Impact of stain normalization on pathologist assessment of prostate cancer: a comparative study. Cancers (Basel). 2023;15(5):1503. https://doi.org/10.3390/cancers15051503.

[17] Netto GJ, Amin MB, Berney DM, et al. The 2022 World Health Organization classification of tumors of the urinary system and male genital organs – part B: prostate and urinary tract tumors. Eur Urol. 2022;82(5):469-482. https://doi.org/10.1016/j.eururo.2022.07.002.

[18] Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep. 2017;7(1):16878. https://doi.org/10.1038/s41598-017-17204-5.

[19] Iczkowski KA, Van Leenders GJLH, Van der Kwast TH The 2019 International Society of Urological Pathology (ISUP) consensus

conference on grading of prostatic carcinoma. Am J Surg Pathol. 2021;45(7):1007. https://doi.org/10.1097/PAS.0000000000001678.

[20] Bejnordi BE, Timofeeva N, Otte-Höller I, et al. Quantitative analysis of stain variability in histology slides and an algorithm for standardization. SPIE Medical Imaging: Digital Pathology. 2014;9041:45-51. https://doi.org/10.1117/12.2043683.

[21] Avanaki, A. R. N. Espig KS, Sawhney S, et al. Aging display's effect on interpretation of digital pathology slide. SPIE Medical Imaging: Digital Pathology. 2015;9420:22-33. https://doi.org/10.1117/12.2082315.

[22] Khan A, Janowczyk A, Müller F, et al. Impact of scanner variability on lymph node segmentation in computational pathology. J Pathol Inform. 2022;13:100127. https://doi.org/10.1016/j.jpi.2022.100127.

[23] Egevad L, Allsbrook WC Jr, Epstein JI. Interobserver reproducibility of Gleason grading of prostatic carcinoma: general pathologists. Hum Pathol. 2006;37(3):292-7. https://doi.org/10.1016/j.humpath.2005.10.011.

[24] Chetty R, Johnson JE. The management of implementing a digital pathology workflow. Diagn Histopathol. 2023;29. https://doi.org/10.1016/j.mpdhp.2023.06.010.

[25] D'Abbronzo G, Lucà S, Carraturo E, et al. Shortage of pathologists in Italy: survey of students and residents. Pathologica. 2023;115(3):172-180. https://doi.org/10.32074/1591-951X-852.

[26] Yeong J, Sultana R, Teo J, et al. Gleason grade grouping of prostate cancer is of prognostic value in Asian men. J Clin Pathol. 2017;70(9):745-753. https://doi.org/10.1136/jclinpath-2016-204276.

[27] Madueke I, Abern MR. Counterpoint: Should Active Surveillance Be Used for Gleason 3+ 4 Prostate Cancer? Oncology (Williston Park). 2019;33(6):235-242.

[28] Sharma M, Miyamoto H. Percent Gleason pattern 4 in stratifying the prognosis of patients with intermediate-risk prostate cancer. Transl Androl Urol. 2018;7(Suppl 4):S484-S489. https://doi.org/10.21037/tau.2018.03.20.

[29] Koca O, Calışkan S, Oztürk M, et al. Significance of atypical small acinar proliferation and high-grade prostatic intraepithelial neoplasia in prostate biopsy. Korean J Urol. 2011;52(11):736-40. https://doi.org/10.4111/kju.2011.52.11.736.

[30] Rao S, Masilamani S, Sundaram S et al. Quality measures in preanalytical phase of tissue processing: Understanding its value in histopathology. J Clin Diagn Res. 2016;10(1):EC07-11. https://doi.org/10.7860/JCDR/2016/14546.7087.