WHITEPAPER

Best practice guide to ensure **quality and long-term success** using artificial intelligence to analyze digital pathology slides

Authors:
Richard Fox, BVetMed, MRCVS, Dipl ECVP, Veterinary Pathologist at Aiforia
Anniina Oksman, Product Manager at Aiforia

aiforia®
AI for image analysis

TABLE OF CONTENTS

Introduction

Digital pathology, the process of digitizing pathology slides for analysis, has revolutionized the field of pathology. Artificial analysis on digital pathology slides involves the application of Machine Learning algorithms (Deep Learning) using Convolutional Neural Networks. These artificial intelligence (AI) techniques are used to assist pathologists in diagnosing diseases accurately and efficiently.

To ensure the reliability and accuracy of results, it is crucial to establish, at the outset, best practices, for both upstream and downstream workflow processes (See Fig. 1), along with stringent quality control measures. Ensuring long-term success in artificial analysis on digital pathology slides requires rigorous adherence to these measures throughout the project lifecycle.

Significant drift in the protocols, procedures and data sets is likely to affect the performance of an AI model which may require additional training to compensate for these changes if they are irreversible. Establishing these principles, and quality control measures, before initialization of AI model training, should assist institutions in maintaining good model performance with both short-term and long-term use models.

The best practices here are part of the wider 'Good Machine Learning Practice' guidelines set out by the US FDA, UK, and Canadian Governments. These 10 guiding principles are intended to lay the foundation for the development of these products. Online resources for this initiative can be found in the resources section.

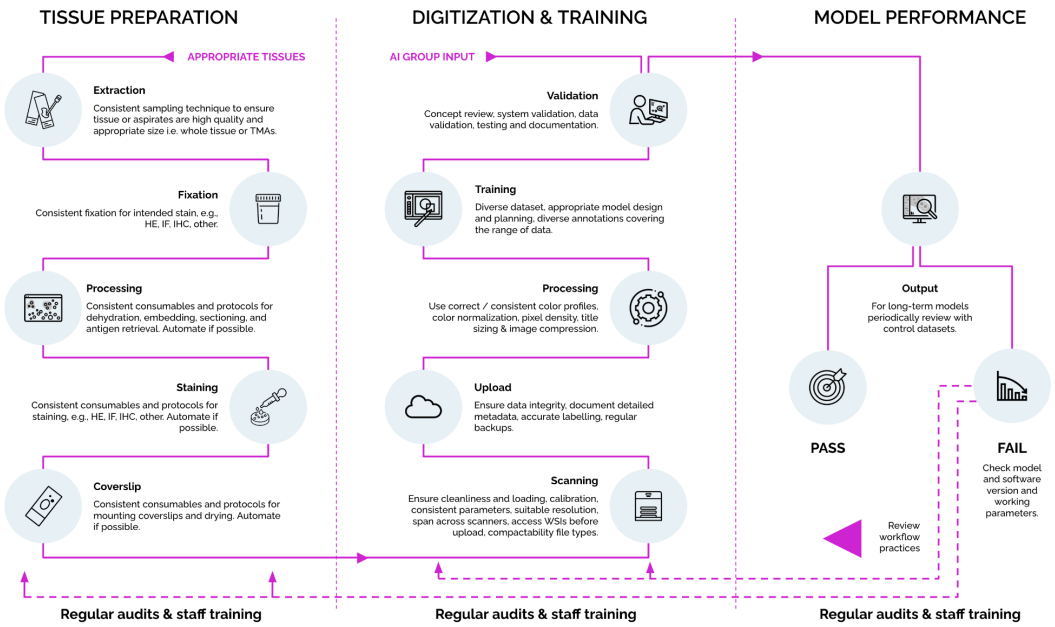

Figure 1. Overview of the control steps involved in producing and maintaining a long-term AI model

Project initialization and planning

Initial assessment

Conduct a thorough needs assessment involving scientists/pathologists, IT professionals, and stakeholders to understand the specific requirements and challenges for the particular AI project that is required.

Data collection and annotation

Assemble a diverse, representative dataset with proper metadata, ensuring it covers various pathologies, tissue types, and staining techniques that will be assessed in the project. Over time, upstream data may undergo drift, and an assessment of the material being analyzed should also be undertaken.

Engage expert pathologists / scientists for assessment of accurate annotation, establishing clear guidelines and protocols for consistent labeling (set a standard protocol for the ground truth from the outset).

Upstream workflows

Quality control measures in histopathology wet lab and digital scanning procedures are essential to ensure consistent and reliable material for artificial analysis on histopathology slides.

n.b If changes are purposely made in wet lab and scanning methodologies, or consumables any AI model should be checked for performance changes. Regular quality control (QC) checks should be made, particularly if a long term model* is used, to ensure drift is not occurring with unintentional changes to methodologies.

*A short-term model is generally used for a specific project with a finite dataset (such as a one-off pre-clinical study). A long-term model is one that is continuously used with a perpetual dataset (such as a CE-IVD model).

Sample collection and processing

- Standardize sample collection protocols to maintain consistency in tissue acquisition if this can be controlled (often, clinical specimens / retrospective tissues may not be under internal control).
- Cytologic preparations should be automated / standardized wherever possible i.e. blood smears to maintain consistency. This is often not possible in clinical situations, but a standardized collection protocol should be in place.
- Document detailed information about sample origin, processing time, and fixation methods.
- Standardize the time between tissue harvesting and fixing and also the duration of fixation.
- Standardize sampling techniques - i.e., production of whole tissue specimens or TMAs for example.

Tissue sectioning

- Use sharp and well-maintained microtomes to obtain consistently thin and uniform tissue sections, which also reduces tissue artifacts.
- Regularly calibrate and maintain tissue processing equipment to ensure proper fixation and embedding.
- Monitor section thickness to prevent variations and ensure slide uniformity.
- QC slides for sectioning, processing, and embedding artifacts (rips and holes).

Staining procedures

- Follow standardized staining protocols to ensure consistent results across samples.
- Regularly test and monitor the quality of staining reagents, such as Hematoxylin and Eosin (H&E) stains.
- Implement standardized dehydration and clearing steps to avoid artifacts in stained sections.
- Antigen retrieval protocols - consistency in antigen retrieval (heat or protease), consider the likelihood of variability with pressure cookers, microwaves, steamers or machines like PT-Link.
- Immunohistochemistry (IHC), Immunocytochemistry (ICC), or Immunofluorescence (IF) protocols: consistency in staining results, i.e., manual > semi-automated > fully automated.
- Ensure consistent quality of consumables.

Slide preparation

- Implement a standardized process for slide mounting and ensure proper drying to avoid bubbles or smudges (such as tape or automatic glass cover-slipping).
- Ensure the consumables, such as the glass slides, cover slips, and mountants, are consistent.

Quality assurance checks

- Conduct regular internal quality audits to assess the accuracy and consistency of staining results.
- Implement double-check mechanisms, where another technician reviews the stained slides for discrepancies.

Continuous training and education

- Provide continuous training for lab technicians on updated protocols, best practices, and quality control procedures.
- Conduct regular workshops and refresher courses to enhance the skills of lab staff.
- Stay informed about the latest advancements in histopathology techniques, staining methods, and digital scanning technologies.
- Regularly participate in conferences, seminars, and webinars to learn from experts and stay abreast of industry standards.

Slide digitization

Scanner calibration

- Calibrate digital scanners regularly to maintain color accuracy and sharpness in scanned images (brightfield or fluorescence).
- If possible, verify scanner performance using standardized calibration slides designed for this purpose.
- Ensure software versions are maintained, compression is consistent, and file types, pixel density, i.e., 0.25um/pixel, are consistent within AI projects.
- If more than one digital scanner will be used for dataset analysis, (including the additional benefit of fallback and redundancy, as well as robustness of any model), the datasets should cover the breadth of scanners that will be used during all phases of model development and in use. This allows for mitigation of (often subtle) variations such as objectives, sharpness, tonality, saturation, or uneven illumination, particularly between different manufacturers or even models in the same series.

Slide preparation for scanning

- Ensure slides are clean and free of debris or stains before scanning to prevent artifacts in digital images.
- Inspect slides for overhanging labels, tape, or excess mounting medium that may affect the fit of the slide in the scanning mechanism and impede operation.
- Wipe slides with a lint-free soft cloth to remove loose debris, water spots, or fingerprints from the upper and lower surfaces.
- Check for cracks or chips, particularly at the corners. Any loose glass shards could damage the scanning mechanism.
- Wait until slides are fully dry before loading them into the scanner to avoid residue getting on the scanner mechanisms or objectives especially (can also cause linear banding issues)

Image acquisition

- Set scanning parameters consistently, including resolution, focus, and illumination, for uniform image capture.
- Ensure slides are placed flat into the scanner rack or carrier, and any securing mechanism engaged. If the slide is not flush during scanning, there is a high likelihood of unfocused areas in your scan. Labels on the bottom of the slide may prevent the slide flaying flat, so ensure these are only placed on the top-facing side of the slide.
- Randomly select scanned images for manual inspection to detect any anomalies or artifacts.
- Implement a slide tracking system to ensure all slides are scanned and accounted for.

Quality control checks for digital images

- Perform automated checks for image artifacts, such as blurriness, saturation, or uneven illumination.
- Manually inspect a subset of scanned images to ensure they meet the required quality standards (focus, assess all levels to endpoint magnification, linear banding/stitching/ other scanning artifacts).
- Compare several slides side by side to check color reproducibility (especially when multiple scanners are being used to collect the dataset).
- Document and address any issues identified during quality checks.
- Ensure high-quality slide preparation to maintain slide integrity.
- Use high-resolution scanners to digitize pathology slides, capturing fine details.
- Implement color calibration techniques to standardize color representation.

Data annotation and labeling

- Expert pathologists/scientists should oversee the annotation of slides, marking regions of interest and labeling specific structures.
- Use standardized protocols and controlled vocabularies for consistent labeling.
- Implement double-check mechanisms to validate annotations for accuracy.
- Use techniques such as stain normalization to mitigate color variations between slides.

Data management and documentation

Metadata documentation

- Document detailed metadata for each sample, including patient information, sample type, and processing history.
- Ensure accurate labeling of digital images with corresponding patient and sample identifiers.

Version control

- Implement version control systems for digital images and associated metadata to track changes and updates.
- Maintain a record of any modifications made to images or metadata over time.

Data backup and security

- Regularly backup digital images and associated data to prevent data loss in case of technical failures.
- Implement strict access controls and encryption protocols to protect patient data and comply with privacy regulations.

n.b. AI model development procedures are not part of the brief of this document (more information can be found in Aiforia Community).

Downstream workflow

Data quality control

- Regularly monitor and assess the quality of digitized slides to prevent artifacts.
- Implement automated checks to detect scanning errors, such as blurriness or missing areas.

Annotation quality control

- Conduct periodic reviews of annotations to ensure consistency and accuracy.
- Organize regular training sessions for annotators to maintain annotation standards.
- Implement inter-annotator agreement checks and resolve discrepancies through consensus.

Model performance monitoring

- Establish a monitoring system to track the model's performance over time.
- Implement feedback loops with pathologists/scientists to refine the model based on real-world feedback.

Compliance and security

- Adhere to data protection regulations and standards to ensure patient data confidentiality.
- Implement secure data transmission protocols and encryption methods for data transfer and storage.

Long-term model maintenance and sustainability

Model performance monitoring

- Utilize a monitoring system/workflow to continuously assess the model performance on new data, such as a set of QC slides where data can be measured and compared over time).
- Implement feedback loops with pathologists for real-time model refinement based on clinical feedback where necessary.

Documentation and knowledge transfer

- Maintain comprehensive documentation covering data sources, pre-processing steps, model architectures, and training strategies.
- Facilitate knowledge transfer within the team and new members to ensure continuity.

Regular updates and retraining

- Keep the model up-to-date with regular updates based on new data and evolving pathology practices. This is important if there is likely to be upstream data drift in sample population or if new criteria in tissue assessment are established.
- If necessary, retrain the model periodically, incorporating the latest advancements in AI and deep learning techniques.

Compliance and ethical considerations

- Adhere to regulatory frameworks, data protection laws, and ethical guidelines to ensure patient privacy and compliance.
- Obtain necessary approvals and permissions from relevant regulatory bodies before deploying the AI system in clinical practice.

Conclusion

Implementing robust best practices for upstream and downstream workflow processes, along with stringent quality control measures, is essential for the accurate and reliable artificial analysis of digital pathology slides. We hope that by following these guidelines, institutions can enhance the efficiency of pathological diagnosis, leading to improved patient outcomes and better overall healthcare delivery.

Resources

Online

U.S. Food & Drug Administration (2021, October 27). *Good machine learning practice for medical device development: guiding principles.* https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

GOV.UK. (2014, June 13). *Software and artificial intelligence (AI) as a medical device.* https://www.gov.uk/government/publications/software-and-artificial-intelligence-ai-as-a-medical-device/software-and-artificial-intelligence-ai-as-a-medical-device

References

Abels et al. (2019, September 3). Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. *The Journal of Pathology*, *249*(3), 286-294. https://doi.org/10.1002/path.5331

Aeffner et al. (2019, March 8). Introduction to digital image analysis in whole-slide imaging: a white paper from the Digital Pathology Association. *J Pathol Inform.*, *10*(9). https://doi.org/10.4103/jpi.jpi_82_18

Louis et al .(2016, January). Computational pathology: a path ahead. *Arch Pathol Lab Med, 140*(1), 41-50. https://doi.org/10.5858/arpa.2015-0093-sa

Niazi et al. (2019, May). Digital pathology and artificial intelligence. *Lancet Oncol., 20*(5), e253-e261. https://doi.org/10.1016/s1470-2045(19)30154-8

Zarella et al. (2023, September). Artificial intelligence and digital pathology: clinical promise and deployment considerations. *J Med Imaging (Bellingham)*, *10*(5). https://doi.org/10.1117/1.jmi.10.5.051802

Zarella et al. (2019, February). A practical guide to whole slide imaging: a white paper from the Digital Pathology Association. *Arch Pathol Lab Med*, *143*(2), 222-234. https://doi.org/10.5858/arpa.2018-0343-ra

Glossary

Artificial Intelligence (AI): The development of computer systems that can perform tasks that typically require human intelligence, including image analysis in digital pathology.

Digital pathology: The practice of converting glass slides containing tissue samples into digital images for analysis, storage, and management.

Upstream workflow: Refers to the initial stages of the process, starting from the acquisition of raw data to the preparation and preprocessing of that data for further analysis.

Downstream workflow: Refers to the later stages of the process, after the data has been prepared and trained models are applied for analysis and interpretation.

Whole Slide Imaging (WSI): The process of scanning entire pathology slides into digital images, allowing for high-resolution examination and analysis.

Drift: Refers to a concept related to the changes in the distribution of data over time which includes causes like instrumentational changes, changes in patient populations, laboratory protocol changes or temporal trends.

Image analysis: The process of extracting meaningful information and patterns from digital pathology images using computational techniques.

Datasets: A dataset refers to a collection of digital images and associated information that is used to train, validate, and test machine learning models.

Model performance: Refers to the effectiveness and accuracy of a machine learning model in carrying out a specific task, such as classification, segmentation, or detection of features within digital pathology images.

Machine Learning (ML): A subset of AI that involves the development of algorithms and models that enable computers to learn from data and make predictions or decisions without explicit programming.

Deep Learning: A type of machine learning that involves neural networks with multiple layers (deep neural networks), enabling the system to learn hierarchical representations of data.

Convolutional Neural Network (CNN): A type of deep neural network that is particularly effective in image analysis tasks, using convolutional layers to automatically learn features from images.

Stakeholders: These are individuals or groups who have a vested interest in the development, deployment, and outcomes of the AI application(s).

Metadata: Metadata refers to additional information associated with the digital pathology images beyond the pixel data itself, such as patient data, slide description, annotations, unique identifiers,

staining techniques, date and time, image dimensions, pixel resolution, color space, digital signatures or authentication file formats.

Training data: The dataset used to train machine learning models, consisting of labeled examples that the algorithm learns from.

Validation data: A subset of data used to assess the performance and generalization ability of a machine learning model during training.

Testing data: Independent data used to evaluate the final performance and generalization of a machine learning model.