PUBLISHED VERSION OF THE ARTICLE:
Fries A, Pirotte M, Vanhee L, Bonnet P, Quatresooz P, Debruyne C, et al. Validating instructional design and predicting student performance in histology education: Using machine learning via virtual microscopy. Anat Sci Educ. 2024; 17: 984–997. https://doi.org/10.1002/ase.2346

# Validating Instructional Design and Predicting Student Performance in Histology Education: Using Machine Learning via Virtual Microscopy



Allyson Fries[1], Marie Pirotte[1], Laurent Vanhee[2], Pierre Bonnet[1], Pascale Quatresooz[1], Christophe Debruyne[2], Raphaël Marée[2], Valérie Defaweux[1]

1. Department of Biomedical and Preclinical Sciences, Faculty of Medicine, University of Liege, Liège, Belgium

2. Montefiore Institute of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium

Correspondence concerning manuscript should be sent to:

Valérie Defaweux valerie.defaweux@uliege.be - Quartier Hôpital, B23 Anatomie, Avenue Hippocrate 15- 4000 Liège (Belgium)

# Abstract

As a part of modern technological environments, virtual microscopy enriches histological learning, with support from large institutional investments. However, existing literature does not supply empirical evidence of its role in improving pedagogy. Virtual microscopy provides fresh opportunities for investigating user behavior during the histology learning process, through digitized histological slides. This study establishes how students' perceptions and user behavior data can be processed and analyzed using Machine Learning algorithms. These also provide predictive data called learning analytics that enable predicting students' performance and behavior favorable for academic success. This information can be interpreted and used for validating instructional designs. Data on the perceptions, performances, and user behavior of 552 students enrolled in a histology course were collected from the virtual microscope, Cytomine®. These data were analyzed using an ensemble of Machine Learning algorithms, the extra-tree regression method, and predictive statistics. The predictive algorithms identified the most pertinent histological slides and descriptive tags, alongside ten types of student behavior conducive to academic success. We used these data to validate our instructional design, and align the educational purpose, learning outcomes, and evaluation methods of digitized histological slides on Cytomine®. This model also predicts students' examination scores, with an error margin of less than 0.5 out of 20 points. The results empirically demonstrate the value of a digital learning environment for both students and teachers of histology.

## Key words

Education, Histology, Virtual Microscopy, Learning Analytics

## Introduction

Histology, also known as microscopic anatomy, is a cornerstone of healthcare professionals' pre-clinical learning (McBride & Drake, 2018). Located at the crossroads of anatomy, biochemistry, and physiology, histology studies the morphology of cells, tissues, and organs, and relates these structural elements to their biological functions. This is particularly important for future healthcare professionals, because several human disorders are cellular in nature. A detailed understanding of cellular structure, differentiation, and function is fundamental to biomedical sciences, diagnosis, and treatment. This branch of knowledge is primarily acquired through reasoning, based on the observation of microscopic structures. Traditionally, histology was comprehended through theoretical courses, that were supplemented with practical training in the laboratory, wherein students used an optical microscope to observe, identify, and interpret histological slides (Hussein *et al.*, 2015; Wu and Chiang, 2022). The use of virtual microscopy in teaching has been a major turning point in modern histological education (Husmann *et al.*, 2009; Mione *et al.*, 2013; Maity *et al.*, 2023). Thus, digitized histological slides are shared by, and for, universities worldwide (Lee *et al.*, 2018; Hortsch *et al.*, 2023).

## Virtual Microscopy and Cytomine®

Virtual microscopy is a technique of storing and sharing microscopic images (Gatumu *et al.*, 2013; Hortsch, 2023). As a more economical alternative to optical microscopy, it allows all students to view the same slide with optimal quality. Moreover, pedagogical teams can select the most representative slides for teaching a subject, and give students access to rare slides, without any risk of damage. Students can view all available histological slides anytime and anywhere, simply by connecting to a computer with an internet connection.

Recent publications have shown that virtual microscopy learning activities qualitatively promote higher-order thinking skills, have a positive effect on the perception of students and teachers, improve the learning outcomes and enhance learner collaborations and group learning (Herodotou *et al.,* 2020, 2022; Chimmalgi and Hortsch, 2022). However, the literature does not supply empirical evidence of its role in improving pedagogy.

Since 2012, the virtual microscope Cytomine®, developed by Marée *et al*. (2016, 2019), is available to students on a local Learning Management System (LMS), enabling teachers to design, centralize, structure, and create learning content and activities. Cytomine® is an open-source software (Cytomine, 2023) that stores, manipulates, and allows viewing high-dimension images up to a certain number of giga-pixels. This study used Cytomine for visualizing digitized histological slides, as it allows detailed exploration, such as by magnifying the region of interest (zooming in/out). The teacher can create marked-out paths that pin-point different structures of interest, accompanied by questions and illustrated answers. Moreover, students too have a layer at their disposal, for adding their own annotations to the slides (Marée *et al.,* 2016).

These functions present new opportunities in pedagogical engineering (Multon *et al.*, 2015; 2018) for addressing difficulties in learning histology, as reported by Garcia *et al.* (2019).

## Learning Analytics and Machine Learning

Virtual microscopy is a digital learning platform that generates a substantial amount of user-related data. Big data are vast sets of heterogeneous data, that can be analyzed, synthesized, and understood through specific statistical and algorithmic tools of Machine Learning (Shmueli, 2010). The analysis of such big data is called "data analysis." However, in the case of educational data, it has been adopted as "learning analytics" (Brown, 2009), which involves collecting, analyzing, modeling, synthesizing, and communicating data on learners and their environments, for better understanding and improved learning (Pecaric *et al.*, 2017; Viberg *et al.*, 2018). Learning analytics can produce and transmit information about students' progress as learners, enabling the teacher to regulate and improve existing pedagogical practices (Lim *et al.,* 2021). Therefore, these data have a potentially profound impact on developing efficient teaching strategies, and identifying at-risk student populations and difficult subject areas. The use of learning analytics can also shed light on the factors that hinder, or aid, students' success.

## Research Questions

This study processed data on students' behavior with digitized slides, perceptions, and performances, using different predictive algorithms (Cytomine®), to acquire learning analytics for the purpose of answering the following questions:

- Is it possible to identify the types of students' behavior, in the context of the virtual microscope, favorable for academic success?
- Is it possible to predict students' performance in Histology examinations?
- Are learning analytics relevant tools for regulating an instructional design suitable for excellent histology teaching?

# Material and Methods

## Study Population

Cytomine® users' data on students' behavior with digitized slides, perceptions, and performances, were saved. Data were collected from 552 students enrolled in a histology course at the Faculty of Medicine, University of Liège. These students were divided into four groups: cohort 1 (320 first-year students of medicine, and 48 first-year students of dental sciences), enrolled in the second semester of the academic year 2016-2017; cohort 2 (34 second-year students of biomedical sciences), enrolled in the first semester of the academic year 2016-2017; cohort 3 (94 first-year students of medicine, and six first-year students of dental sciences), enrolled in the second semester of the academic year 2017-2018; and cohort 4 (50 second-year students of biomedical sciences), enrolled in the first semester of the academic year 2017-2018. Each cohort was studied independently to consider potential differences that might exist between groups, to validate the prediction model and to analyze the predictor variables separately in order to correlate them specifically with their own examination.

## Instructional Design

To characterize instructional design, the Biggs constructive model (Biggs, 1996; Biggs & Tang, 2014), which aligns learning outcomes, teaching methods, and assessment tasks, was employed. Histology theory was taught through *ex cathedra* courses, consisting of presentation slides with didactic images from existing literature. Practical work was divided into five sections, reflective of the five major histological tissue families. Specific learning outcomes were listed for each biological tissue type. These outcomes can be further separated into four distinct levels, from the least to the most complex, according to Bloom's taxonomy (Zaidi *et al.*, 2017):

- Identification of cells, tissues, and histological specificities
- Application of a diagnostic approach
- Building a three-dimensional representation of histological structures
- Working out relationships between morphological structures, and their specific functions

These learning outcomes were taught through a hybrid teaching method, with initial online sessions followed by face-to-face ones. Online sessions were broadcasted within the Massive Open Online Course (MOOC), "Introduction to Histology: Exploration of the human body tissues," available in French and English on the *France Université Numérique* platform (https://www.fun-mooc.fr/en/) (Defaweux *et al.*, 2019). It includes the use of the Cytomine® virtual microscope.

Each digitized histological slide, including route, identification, and introductory slides, fulfilled educational purposes. Route slides (Fig.1A) are teacher-annotated slides that enable students to follow marked-out routes, indicating the structures or cells of interest. A histological question was asked at each point on the route, and detailed feedback (text and pictures) was accessible. The identification slides (Fig.1B) are exercise slides that require students to associate a term with each tag, to identify the structures to which the annotations point. Finally, the introductory slide provides a tutorial on the use of a virtual microscope. A pedagogical alignment connecting teaching methods, pedagogical objectives, and evaluation methods was conceived (Fig.1C), to highlight reflection and coherence in instructional design.

## Collection and Analysis of Perception Data

The students in the different cohorts participated in the same online survey at the end of the course. The survey included four sections: identification, students' perceptions of practical work, students' perceptions of the use of Cytomine®, and students' perceptions of achievement of educational outcomes. Perception data and the questionnaire used for data collection are part of an article published by our team (Pesesse *et al.*, 2023) dedicated to the MOOC and its implementation.



## User Data Collection

The Cytomine platform includes a large amount of data on user lists, image lists, and projects, with annotations and descriptions added by teachers. It also includes extensive data on user behavior, such as the number of connections to the platform on a specific image, students' exploration paths on an image, and the consultation of annotations. These data are saved on a secure main server, and can be retrieved locally. The data collected on the Cytomine® platform, relating to students' behavior useful for this study, mostly included visualizations of reference annotations (tags added by the teachers), users' positions in an image, and magnification. Users' positions on an image were recorded every five seconds, when the user altered the magnification level, or stopped moving on the image. A list of the images of interest included in the instructional design was produced in CSV format for import. Each histological image was matched to its educational purpose, that is, as a route, identification, or introductory slide.

## Performance Data Collection

Data on students' performances in both theoretical and practical examinations were collected. The practical written examination was composed of an open question, 10 identification questions, and 10 multiple-choice questions (MCQs). The theoretical written examination was composed of an open question and 30 MCQs. The results were then encoded in a Microsoft® Excel® 2016 table and converted into CSV format.



## Machine Learning

Data imported from Cytomine® servers to a local server can be processed by Machine Learning algorithms, that predict on the basis of previous observations. In our setting, an observation was a student, and each observation was described by input variables corresponding to the aforementioned collected data. The Machine Learning model used in this study was based on an ensemble of regression tree models, wherein observations are recursively split into sub-sets, according to tests performed on input variables. This regression tree structure was preferred because it facilitates a better understanding of how the variables influence results. The current model was precisely based on the extra-tree regressor method (Geurts *et al.,* 2006), wherein the decisional node at each tree randomly tests all the available variations for selecting combinations, thereby avoiding over-adjustments and under-adjustments. This Machine Learning approach is also based on an ensemble method that combines several trees for more precise predictions on small datasets. Thus, the extra-tree regressor method makes the most accurate predictions, by averaging all the tree predictions. An ensemble of 10000 trees was used for the experiments.

Three publicly available scripts (Vanhee & Hoyoux, 2022) were required to sequentially import, process, and analyze the data for exploitable learning analytics. The "download_data" script allows the Cytomine® web application to import all, or selected data, to a local server. The "data_manager" script allows the elementary interpretation of data, by organizing it according to predefined parameters; it generates important information such as gaze maps, and scan paths. Following the application of these two scripts, perception, and performance input variables, were manually added to the generated data. Finally, the "data_learning" script generated the predictive statistics for drawing conclusions through the Machine Learning algorithms, in the form of histograms showing the most relevant input variables of behavioral parameters and histological images (exploiting the importance of tree-based variables), and correlation diagrams of user behavior and performance data.

The algorithm used on the user data on students' behavior on Cytomine® included 53 parameters, grouped into seven types, according to their relevance:

- The number of digitized histological slides visited in total, or per module
- Total number of positions on digitized histological slides per module, and per slide, with or without magnification

- Average, and median, numbers of positions on one histological slide, with or without magnification.
- The average and median level of magnification
- The total time spent viewing the slide, per module, and per slide
- The mean, and median, time spent in viewing slides
- The time spent in, and order of, clicking annotations by students

# Results

## Digitized Histological Slides' Predictive Value for Student Performance.

Based on performance data of Cytomine® users, the Machine Learning algorithms identified those histological slides that were predictive of the scores obtained in the theoretical and practical histology examinations, as well as the global grade for each cohort (Fig.2). Consequently, of the total 45 slides part of the histology course, the 25 most predictive slides were identified. Among these, one appeared in "Unit 1" (Introduction), four in "Unit 2" (Epithelia), three in "Unit 3" (Epithelial Glands), seven in "Unit 4" (Connective Tissues), six in "Unit 5" (Muscle Tissues), and four in "Unit 6" (Nervous Tissues); 14 of them were route slides, ten were identification slides, and one was as an introductory slide in the virtual microscope tutorial. Five histological slides predicted the theoretical, practical, and global scores (Table 1).

## Predictive Variables of Students' Performances, Related to the Examination of Digitized Histological Slides

Based on data on Cytomine® users' perceptions and performances, the Machine Learning algorithms highlighted those variables linked to the use of histological slides and annotations that were most predictive of each cohort's grades (practical, theory and global) in the histology examination (Fig.3).

Among the 10 most predictive variables for each cohort, 35% of the entries corresponded to annotation scores per histological slide. This score was calculated on the basis of the tag area's magnification, and time devoted to this. Further, 19% of the entries corresponded to the annotations' order of visit (for example, whether the student consulted annotation 1 before annotation 20. The order of these annotations was defined by the teachers and part of the slides' educational route. Moreover, 19% corresponded to the number of positions while magnifying on a slide, 7.5% corresponded to the number of positions without magnification on a slide, 5.8% corresponded to user scores (calculated on the basis of annotation scores, and were overall reflective of the use of all annotations on a slide), 3.3%, corresponded to the overall number of positions with magnification, 1.7% corresponded to annotation scores per slide over a learning module, and finally, 1.7% of the total time corresponded to time spent on each slide. Survey questions related to forum usage, and self-assessment of time spent on learning modules (Pesesse *et al.*, 2023), significantly influenced the predictive algorithm of cohort 4 (Table 2).

## Predictive Variables Corresponding to Examination Questions for Each Cohort

The most predictive variables related to slide annotations were associated with several examination questions for each cohort, which were studied independently. In cohort 1, for the score of annotation number 1219665 on slide number 1219597, an identification tag of a striated cardiac muscle tissue, cut longitudinally on a heart slide, was the variable related to the most predictive annotation of practical, theoretical, and global grades. Their Pearson correlation coefficients, obtained from the “data learning” script, were 0.44, 0.42, and 0.44, respectively. There were four MCQ-type questions on striated cardiac muscles in the practical examination (Fig.4).

In cohort 2, the variable related to the most predictive annotation of the practical and global scores was the order in which the annotations were consulted. This was reading annotation 1 before annotation 2 for image 1216615 (hypodermis slide: route slide), the Pearson correlation coefficients of which were 0.65 and 0.69, respectively. On the same slide, the score of annotation 3 was the most predictive variable for the theory score, with a Pearson correlation coefficient of 0.65. These annotations refer to the adipose tissue, that was the subject of a MCQ in the practical examination (Fig.4).

In cohort 3, the variable related to the most predictive annotation of the practical score was the order of visiting annotation 19 before annotation 20 of image 1218337, with a Pearson correlation coefficient of 0.20. These annotations are identification tags on the gastrointestinal junction slides. There were six MCQs on the gastrointestinal junction in the practical examination (Fig.4).

In cohort 4, the variable related to the most predictive annotation of the practical, theoretical, and global scores was the order of visiting annotation 12 before annotation 13 of image 1219137, with Pearson correlation coefficients of 0.5, 0.48, and 0.55, respectively. These annotations include astrocyte and oligodendrocyte identification tags on a spinal cord identification slide. There was an open question on these nerve cells in the practical and theoretical examinations (Fig.4).

## A Model Validated by the Predictability of Student Performance

The quality of the extra-tree regressor predictive model was assessed by leave-on-out cross-validation. The model learns on n-1 observations, and is validated on the $n^{th}$ observation; this operation is repeated n times (Marée *et al.*, 2019).

After evaluating more than 2500 variables related to Cytomine® use by students, this model predicted the grade that a student would obtain on examinations (practical, theoretical, and global), with an error margin of less than 0.5 out of 20 points, that corresponds to the difference between the estimated, and actual median scores. The quality of this prediction model depends on the cohort size, and amount of data provided. Thus, for the current model, the difference between the actual and predicted medians, and Pearson correlation index, varied from one cohort to another. Thus, for cohort 1 (Fig.5), the practical examination score was predicted with an error margin of 0.34 out of 20 points, and a Pearson correlation coefficient of 0.56. The theoretical score was predicted with an error margin of 0.06 out of 20 points, with a Pearson correlation coefficient of 0.52. The overall score was predicted with an error margin of 0.27 out of 20 points, with a Pearson correlation coefficient of 0.56. As illustrated in Table 3, the model was validated by four cohorts of different sizes.

# Discussion

With the COVID-19 pandemic, and consequent forced transition to distance learning, digital technology has integrated into education at an unprecedented pace. However, is it relevant to teaching in all situations? For nearly 25 years, there has been a debate between promoters (Kulik and Kulik, 1987), and detractors (Russell *et al.*, 1999), of digital technology in educational contexts. This illustrates the need of clear empirical evidence for understanding issues related to this, and determining the possibilities and limitations of such use.

Machine Learning and Deep Learning have facilitated the exploitation of big data generated by digital teaching tools, by increasing the power of predictive algorithms for the ultimate objective of improving education (Kusunose *et al.* 2019; Sin & Muthu, 2015). This study objectified students' behavior on the virtual microscope Cytomine®, and utilized their perceptions and performances as learning analytics. To this end, thousands of variables related to these parameters were collected and processed by predictive algorithms, that is, tree regression models. The relevance of the collected dataset was linked with the disciplinary field, learning tasks, learners, and pedagogical model (Cirigliano *et al.*, 2020; Kolachalama *et al.*, 2018; Rienties *et al.*, 2020).

## A Digital Environment Centered on Learning Outcomes and Students

Specific histological slides and behavioral variables were identified as predictors of students' performances. For each module, the students were required to view distinct identification and route slides relevant to the algorithm. Identification slides allowed students to practice the histological diagnostic steps, by determining the slide's origin and nature. The route slides were marked with annotations that included questions related to the histological structures viewed on the slide, followed by the correct answers. For these types of slides, the three most predictive variables were annotation score per image, order of visiting the annotations, and number of positions with or without magnifying an image. These variables directly depended on the students' investigation of slides, and the actions they requested teachers for achieving learning outcomes. Measurable and quantifiable learning outcomes were evaluated through practical histological examinations. Thus, we demonstrated that the implementation and use of a virtual microscope are relevant for achieving specific educational objectives, and validating instructional designs based on the Biggs model (Biggs, 1996; Biggs and Tang, 2014). In a context wherein most education programs are decreasing the number of hours dedicated to teaching (Gribbin *et al.*, 2022), teaching strategies must be useful, effective, and focused on learning outcomes (Kaliannan & Chandran, 2012; Eng-tat *et al.*, 2022). The enlightened use of digital tools can help build an efficient teaching environment.

Upon reflecting on the student-centered use of virtual microscopy, one can conclude that using Cytomine® lends meaning to students' learning experiences, and encourages commitment. Students who actively engaged in the suggested activities performed better in their exams. They spent more time on the slides, viewed them first at low and then at high magnification, and eventually left the marked trails to explore further on their own. They did not skip any step, and visited the route slide annotations in the order specified by the teachers, from the simplest to the most complex, to gradually master the subject. They consulted questions and answers, conducted self-assessment, and identified the elements to be revised. This trained them in the techniques, and made them knowledgeable about the slides. Activities requiring greater cognitive engagement allowed students to prepare better for their examinations, be more successful, and ultimately acquire the necessary skills for passing the examinations.

Learning analytics have also made it possible to utilize perception data for predicting students' grades. The most relevant perception data identified by the algorithms concerned the use of the histology course forum, and self-assessment of the time spent on learning modules. It would, therefore, be interesting to develop on the latter, for investigating whether a student's metacognitive reflection upon one's own learning method is directly impactful.

## Digital Environment as a Tool for Predicting Performance

Our approach allowed predicting students' performances by an empirical method that does not suffer from the weaknesses reported by Namoun & Alshanqiti (2021). Indeed, in their literature review on predicting students' performances from the learning outcome perspective, they cited various studies that suffered from limitations such as excessive generalization, lack of focus on results for evaluating performance, and poor quality of data and methodology. In our instructional design, based on Biggs's model (Biggs, 1996; Biggs & Tang, 2014), the students' results corresponded to the grades obtained in the practical evaluation of histology, conducted with specific educational objectives. An educational paradigm for achieving learning objectives is possible if the objectives are specifically identified, and their achievement is properly aligned with students' results (Premalatha, 2019). Moreover, our model predicted students' theoretical and overall grades, reflecting global pedagogical coherence wherein practical and theoretical teaching coincide.

Thus, our model allows awarding students a predictive score, based on their behavior on Cytomine®, with an error margin of less than 0.5 out of 20 points at different evaluations. Further studies are needed to determine whether, and how, communicating this information to students would be beneficial or harmful in improving performance (Hausman *et al.*, 2022). Thus, this study, and the implementation of an interface that allows visualizing students' engagement, could be considered for identifying, and helping, low-performing students (Verbet *et al.*, 2014). It would aid teachers in making necessary interventions early on in the learning process, such as advising students, monitoring their progress, developing intelligent tutoring systems, and devising institutional policies (Viberg *et al.*, 2018). Formative feedback is another important factor for successful student learning, development of professionalism in the field of morphological sciences (Camp *et al*., 2010; Youdas *et al*., 2013), and identification of students with difficulties in histology (Hortsch & Mangrulkar, 2015).

## Limitations of the Study

The data acquired on user behavior were limited to students in the first year of medical and dentistry studies and the second year of biomedical sciences studies. Also, in this specific context, learning outcomes have been precisely defined and corresponded specifically to assessment modalities described in a previous study (Pesesse *et al.,* 2023). Although this study ought to be transposed to other learning environments for students in other fields with different learning outcomes or for the use of virtual microscopes in other contexts, we should be mindful of how these differences affect the models used in this study.

The leave-one-out approach adopted in this study allows us to train models when the number of data points, in this case, students, is limited. In this approach, we train and refine our model on all but one student and then assess the prediction based on that student in an iterative manner. One limitation of this approach is that the same cohort is used for both training and prediction, though we do not have enough data to constitute two cohorts. While this study demonstrates the viability of our approach, further validation of the approach and model prediction errors on new cohorts will need to be carried out as future work.

Van der Niet and Bleakley (2019) warned the scientific community against technological solutionism, given the pedagogical and ethical questions raised by the increasing use of artificial intelligence in medicine and medical education. Indeed, for artificial intelligence instrumentalism in pedagogy, the risk lies in limiting research to evaluation of performance only, that shifts focus from skill development and complex learning processes.

It is necessary to take advantage of this opportunity to assist students in their learning processes as we plan to collect further user data to assess the development of higher-level thinking skills rather than performance only.

## Conclusion

The analysis of big data generated by digital education environments, as learning analytics, allows choosing and implementing technology wisely, without assuming it to be some sort of silver bullet that performs everything. This manner of using artificial intelligence offers unique opportunities for either predicting students' performances more accurately, or identifying and explaining the factors for such predictions. This study demonstrated that virtual microscopy is a powerful digital tool for teaching practical histology courses, if used relevantly by teachers and students. It empirically demonstrated the value of using a digital tool for pedagogical coherence.

## Acknowledgements

The authors wish to thank Renaud Hoyoux from Cytomine Corporation.

## Ethical Agreements

These datasets were collected and processed in accordance with the Humanities and Social Sciences Ethics Committee (opinion 180901), and in compliance with the General Data Protection Regulation.

## References

Biggs J. 1996. Enhancing teaching through constructive alignment. High Educ 32(3):347-364.

Biggs J, Tang C. 2014. Constructive alignment: An outcomes-based approach to teaching anatomy. In: Chan LK, Pawlina W (Editors). Teaching anatomy: A practical guide. 1st Ed. New York, NY: Springer International Publishing. p 31-38.

Brown JS. 2000. Growing Up: Digital: How the Web Changes Work, Education, and the Ways People Learn. Chang Mag High Learn 32(2):11-20.

Camp CL, Gregory JK, Lachman N, Chen LP, Juskewitch JE, Pawlina W. 2010. Comparative efficacy of group and individual feedback in gross anatomy for promoting medical student professionalism. Anat Sci Ed 3(2), 64-72.

Chimmalgi M, Hortsch M. 2022. Teaching Histology Using Self-Directed Learning Modules (SDLMs) in a Blended Approach. Med Sci Educ 32(6):1455-1464.

Cirigliano MM, Guthrie CD, Pusic M V. 2020. Click-level Learning Analytics in an Online Medical Education Learning Platform. Teach Learn Med 32(4):410–21.

Cytomine community documentation site. Cytomine Community Documentation. (accessed in May 2023). http://www.cytomine.org/

Defaweux V, Delbrassine D, Dozo BO, Etienne AM, Centi V, D'Anna V, Multon S, Pesesse L, Stassart C, Van de Poël JF, Wagener A, Weatherspoon A, Verpoorten D. 2019.To combine a MOOC to a regular face-to-face course-A study of three blended pedagogical patterns. CEUR Workshop Proc 2356(2017):210–7.

Eng-Tat A, Liu P, Jitesh R, Chandrika M, Satish RL, Lian WQD, Huppertz B, Khamuani M, Minarcik JR. 2022. The challenges of histology education. Clin Anat 1–7.

García M, Victory N, Navarro-Sempere A, Segovia Y. 2019. Students' Views on Difficulties in Learning Histology. Anat Sci Educ 12:541–549.

Gatumu MK, Macmillan FM, Langton PD, Headley PM, Harris JR. 2013. Evaluation of Usage of Virtual Microscopy for the Study of Histology in the Medical, Dental, and Veterinary Undergraduate Programs of a UK University. Anat Sci Educ. 7(5):389-398.

Geurts P, Ernst D, Wehenkel L. 2006. Extremely randomized trees. Mach Learn 63(1):3-

42.

Gribbin W, Wilson EA, McTaggart S, Hortsch M. 2022. Histology education in an integrated, time-restricted medical curriculum: Academic outcomes and students' study adaptations. Anat Sci Educ 15(4):671-684.

Hausman M, Dancot J, Pétré B, Guillaume M, Detroz P. 2022. 'I don't know if people realize the impact of their words': how does feedback during internship impact nursing student learning? Assess Eval High Educ 1-13.

Herodotou C, Aristeidou M, Scanlon E, Kelley S. 2022. Virtual Microscopes and online learning: Exploring the perceptions of 12 teachers about pedagogy. Open Learn: J Open Distance e-Learn 1-25.

Herodotou C, Muirhead DK, Aristeidou M, Hole MJ, Kelley S, Scanlon E, Duffy M. 2020. Blended and online learning: a comparative study of virtual microscopy in Higher Education. Interact Learn Environ 28(6), 713-728.

Hortsch M, Mangrulkar RS. 2015. When students struggle with gross anatomy and histology: A strategy for monitoring, reviewing, and promoting student academic success in an integrated preclinical medical curriculum. Anat Sci Educ 8(5):478-483.

Hortsch M, Koney NK, Oommen AM, Yohannan DG, Li Y, de Melo Leite ACR, Girão-Carmona VCC. 2023. Virtual Microscopy Goes Global: The Images Are Virtual and the Problems Are Real. Adv Exp Med Biol 1421:79-124.

Hortsch M. 2023. The Michigan Histology website as an example of a free anatomical resource serving learners and educators worldwide. Anat Sci Educ 16(3):363-371.

Husmann PR, O'Loughlin VD, Braun MW. 2009. Quantitative and qualitative changes in teaching histology by means of virtual microscopy in an introductory course in human anatomy. Anat Sci Ed 2(5):218-226.

Hussein IH, Raad M. 2015 Once Upon a Microscopic Slide: The Story of Histology. Journal of Cytology & Histology 06(06).

Kaliannan M, Chandran SD. 2012. Empowering students through outcome-based education (OBE). Res. Educ 87, 50–63.

Kolachalama VB, Garg PS. 2018. Machine learning and medical education. npj Digit Med 1(1):2–4.

Kulik JA, Kulik CLC. 1987. Review of recent research literature on computer-based instruction. Contemporary Educational Psychology 12(3):222-230.

Kusunose K, Haga A, Abe T, Sata M. 2019. Utilization of Artificial Intelligence in Echocardiography. Circ J 83(8):1623-1629.

Lee LMJ, Goldman HM, Hortsch M. 2018. The virtual microscopy database—sharing digital microscope images for research and education. Anat Sci Educ 11:510–515.

Lim LA, Gentili S, Pardo A, Kovanović V, Whitelock-Wainwright A, Gašević D, Dawson S. 2021. What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. Learn Instr 72.

Maity S, Nauhria S, Nayak N, Nauhria S, Coffin T, Wray J, Haerianardakani S, Sah R, Spruce A, Jeong Y, Maj MC, Sharma A, Okpara N, Ike CJ, Nath R, Nelson J, Parwani AV. 2023.

Virtual Versus Light Microscopy Usage among Students: A Systematic Review and Meta-Analytic Evidence in Medical Education. Diagnostics 13(3):558.

Maree R, Vanhee L, Rubens , Weatherspoon A, Pesesse L, Quatresooz P, Multon S, Defaweux V. 2019. New Cytomine Modules for User Behavior Analytics in Digital Pathology. J Pathol Inform 10:68-70.

Marée R, Rollus L, Stévens B, Hoyoux R, Louppe G, Vandaele R, Begon JM, Kainz P, Geurts P, Wehenkel L. 2016. Collaborative analysis of multi-gigapixel imaging data using Cytomine. Bioinformatics 32:1395–1401.

Mcbride JM, Drake RL. 2018. National Survey on Anatomical Sciences in Medical Education. Anat Sci Educ 14:7-14.

Mione S, Valcke M, Cornelissen M. 2013. Evaluation of virtual microscopy in medical histology teaching: Virtual Microscopy in Histology Teaching. Anat Sci Ed 6(5):307-315.

Multon S, Pesesse L, Weatherspoon A, Florquin S, Poel J-F Van de, Martin P, Vincke G, Hoyoux R, Marée R, Verpoorten D, Bonnet P, Quatresooz P, Defaweux V. 2018. A Massive Open Oneline Course (MOOC) on pratical histology: A goal, a tool, a large public ! Return on a first experience. Ann Pathol 38(2):76–84.

Multon S, Weatherspoon A, Schaffer P, Quatresooz P, Defaweux V. 2015. Practical histology in tune with the times. Med Educ 49(11):1166-1167.

Namoun A, Alshanqiti A. 2021. Predicting student performance using data mining and learning analytics techniques: A systematic literature review. Appl Sci 11(1):1–28.

Pecaric M, Boutis K, Beckstead J, Pusic M. 2017 A Big Data and Learning Analytics

Approach to Process-Level Feedback in Cognitive Simulations: Acad Med 92(2):175-184.

Pesesse L, Dozo BO, Quatresooz P, Weatherspoon A, Multon S, Pirotte M, Tonus C, Vandenbosch R, Bonnet P, Verpoorten D, Defaweux V. 2023. Designing a Massive Open Online Course (MOOC) in face-to-face sessions. A blended design to teach practical histology. Morphologie. S1286-0115(23)00064-4.

Premalatha, K. 2019 Course and program outcomes assessment methods in outcome-based education: A review. J. Educ 199(3):111-127.

Rienties B, Køhler Simonsen H, Herodotou C. 2020 Defining the Boundaries Between Artificial Intelligence in Education, Computer-Supported Collaborative Learning, Educational Data Mining, and Learning Analytics: A Need for Coherence. Front Educ 5:128.

Russell K, Weinberger E, Stone A. 1999. Preserving digital scholarship: the future is now. Learned Publishing 12(4):271-280.

Shmueli G. 2010. To Explain or to Predict? Statist Sci 25(3).

Sin K, Muthu L. 2015. Application of big data in education data mining and learning analytics – a literature review. IJSC. 05(04):1035-1049.

van der Niet AG, Bleakley A. 2021. Where medical education meets artificial intelligence: 'Does technology care?' Med Educ 55(1):30–6.

Vanhee, L. & Hoyoux R. (2022). Learning Analytics scripts. GitHub. https://github.com/geektortoise/Master-Thesis

Verbert K, Govaerts S, Duval E, Santos JL, Van Assche F, Parra G, Klerkx. 2014. Learning dashboards: an overview and future research opportunities. Pers Ubiquit Comput 18,1499-1514.

Viberg O, Hatakka M, Bälter O, Mavroudi A. 2018. The current landscape of learning analytics in higher education. Comput Hum Behav 89:98-110.

Wu YH, Chiang CP. 2022. Comparison of virtual microscopy and real microscopy for learning oral histology laboratory course among dental students. J Dent Sci 17(3):1201-1205.

Youdas JW, Krause DA, Hellyer NJ, Rindflesch AB, Hollman JH. 2013. Use of individual feedback during human gross anatomy course for enhancing professional behaviors in doctor of physical therapy students: Use of Individual Feedback During Human Gross Anatomy. Anat Sci Ed 6(5):324-331.

Zaidi NB, Hwang C, Scott S, Stallard S, Purkiss J, Hortsch M. 2017. Climbing Bloom's taxonomy pyramid: Lessons from a graduate histology course: Bloom's Taxonomy Histology Tool. Anat Sci Ed 10(5):456-464.



## Notes on contributors

FRIES ALLYSON, DVM, is a Ph.D. student and a teaching assistant in the Department of Biomedical and Preclinical Sciences, Faculty of Medicine, University of Liège (Belgium). She oversees practical histology teaching to Bachelor students in Medicine, Biomedical Sciences and Dentistry. Her research questions the implementation of digital learning environments in Anatomy and Histology courses.

PIROTTE MARIE, M.Sc. was a teaching assistant in the Department of Biomedical and Preclinical Sciences, Faculty of Medicine, University of Liège (Belgium). She was in charge of practical histology teaching to the first years of medical studies, Biomedical Sciences, and Dentistry.

VANHEE LAURENT, M.Sc. was a student engineer at the Montefiore Institute of the University of Liège who made his master thesis in the field of computer sciences specialized in intelligent systems

BONNET PIERRE, M.D., Ph.D. is a professor of Anatomy in the Department of Biomedical and Preclinical Sciences, Faculty of Medicine, University of Liège (Belgium). He is a urologist surgeon and the director of the research center which is interested in digital didactics in the morphological sciences field.

QUATRESOOZ PASCALE, M.D., Ph.D. is a professor of Histology and Histopathology in the Department of Biomedical and Preclinical Sciences, Faculty of Medicine, University of Liège (Belgium). She is a dermatopathologist and has a strong interest in technology-enhanced learning and how it can be used to support learner gain.

DEBRUYNE CHRISTOPHE, Ph.D., is a Professor at the Montefiore Institute of the University of Liège in Belgium. He teaches database related courses to Bachelor and Master students in Mathematics, Informatics and Engineering. He conducts research in ontology and knowledge engineering, data modelling, Linked Data, and declarative approaches to data integration.

MAREE RAPHAEL, Ph.D, is a senior researcher in Machine Learning, Collaborative Software Development, Human-Centered Computing at the Montefiore Institute of the University of Liège

DEFAWEUX VALERIE, Ph.D. is an associate professor of Histology and Anatomy in the Department of Biomedical and Preclinical Sciences, Faculty of Medicine, University of Liège (Belgium). Her research questions the implementation of digital learning environments in Anatomy and Histology.



## Figure Legends

Figure 1A: Educational purpose of digitized histological slides on Cytomine®, the route slide

Routing the slide interface on a Cytomine® virtual microscope: Sign-posted learning paths are shown directly on the histological slides. These paths include numbered tags associated with the questions. The recommended magnification levels are described in this section. The standardized feedback includes text, images, and drawings associated with each question.

Figure 1B: Educational purpose of digitized histological slides on Cytomine®, the identification slide

Identification of the slide interface using a Cytomine ® virtual microscope: Exercise slides train students to apply the diagnostic criteria. The student must associate a term with each tag, for identifying the structures indicated by the tags.

Figure 1C : Instructional design diagram presenting the alignment between the purposes of digitized histological slides, learning outcomes, and evaluation methods.

The route slide trains in identifying histological structures, diagnostic approaches, 3D representations, and relationships between morphological structures and their specific functions. The identification slide identifies the histological structures and diagnostic objectives. These objectives were specifically evaluated through a practical histological examination.

Figure 2: Digitized histological slides with predictive value for students' performance

On the y-axis, the histograms illustrate the importance of the predictive images of practical, theoretical, and global grades for cohort 1. On the x-axis, the names of the predicted slides isolated by the algorithms are situated. Histological slides identified on the y-axis were classified according to their educational purpose (Introductory, Route, or Identification) in table 1 for each cohort.

Figure 3: Predictive variables of students' examination performance, related to their behavior with digitized histological slides

The histogram corresponds to the set of predictive variables identified by the Machine Learning algorithms, presented on the x-axis, for the practice grade of cohort 1. The predictive strengths of the variables are represented on the y-axis. The correlation graph provided by the algorithms presents the most predictive variables isolated from the histogram and practical exam scores. The predictive variables (x-axis) were pooled et classified in table 2 for each cohort.

Figure 4: Relation between predictive variables provided by Machine Learning algorithms, and selected assessment questions for each cohort

For each cohort, an image, annotation, or some other predictive variable may have been associated with the subject specifically addressed by an exam question.

Figure 5: Scatter plot between the predicted and actual grades (practical, theorical and global) for cohort 1.

For each grade, the Pearson correlation coefficient is presented in a table, along with the delta score between the actual and predicted median scores.

Table 1 : Digitized histological slides with predictive value for students' performance

Histological slides highlighted in the histogram in figure 2 (for cohort 1) were classified according to their educational purpose (Introductory, Route, or Identification), that have predictive value for the grades obtained by each cohort in the histology certification exams: one, if the slide was predictive of one of the three categories of histology exam scores (practical, theory, and overall scores); two, if the slide was predictive of two categories; and three,  if the slide was predictive of all three categories.

Table 2: Predictive variables of students' examination performance, related to their behavior with digitized histological slides

The most predictive variables provided by the Machine Learning algorithms were listed for each cohort. The algorithm used on the user data on students' behavior on Cytomine® included 53 parameters, grouped into seven types, according to their relevance. Among the 10 most predictive variables were identified.

Table 3 : Scores (in theory, in practice and overall) obtained and scores predicted by the extra-tree regressor model for each cohort

For each cohort, the Pearson correlation coefficient is presented, along with the delta score between the actual and predicted median scores.

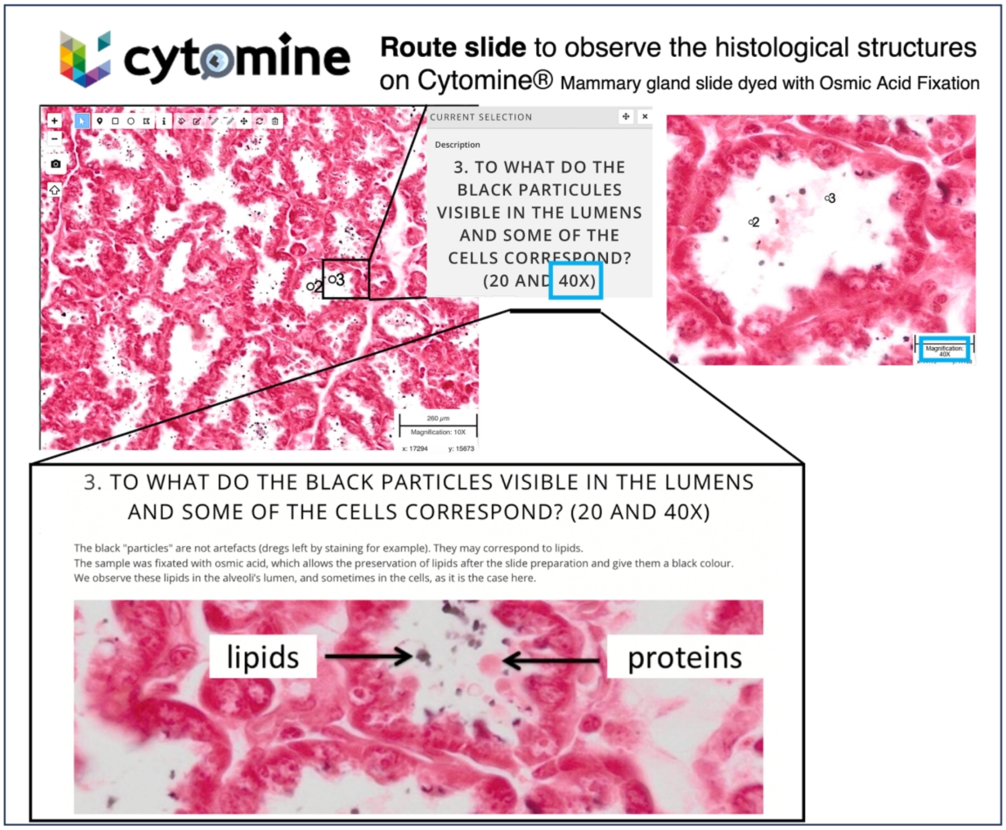


Figure 1A: Educational purpose of digitized histological slides on Cytomine®, the route slide
Routing the slide interface on a Cytomine® virtual microscope: Sign-posted learning paths are shown directly on the histological slides. These paths include numbered tags associated with the questions. The recommended magnification levels are described in this section. The standardized feedback includes text, images, and drawings associated with each question.

175x143mm (300 x 300 DPI)

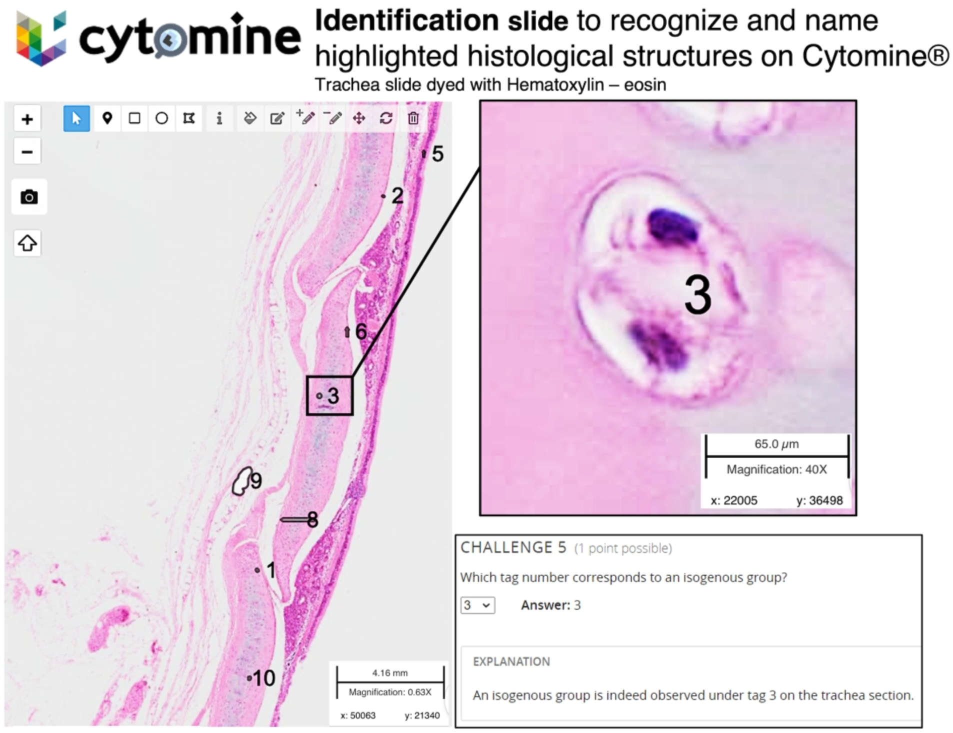


Figure 1B: Educational purpose of digitized histological slides on Cytomine®, the identification slide Identification of the slide interface using a Cytomine ® virtual microscope: Exercise slides train students to apply the diagnostic criteria. The student must associate a term with each tag, for identifying the structures indicated by the tags.

175x132mm (330 x 330 DPI)

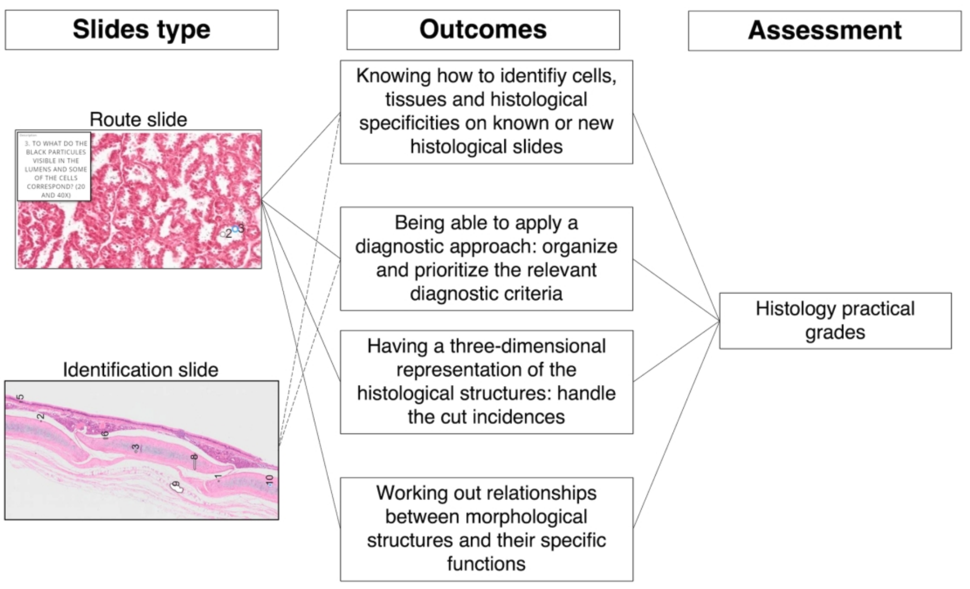


Figure 1C : Instructional design diagram presenting the alignment between the purposes of digitized histological slides, learning outcomes, and evaluation methods.
The route slide trains in identifying histological structures, diagnostic approaches, 3D representations, and relationships between morphological structures and their specific functions. The identification slide identifies the histological structures and diagnostic objectives. These objectives were specifically evaluated through a practical histological examination.

175x103mm (300 x 300 DPI)

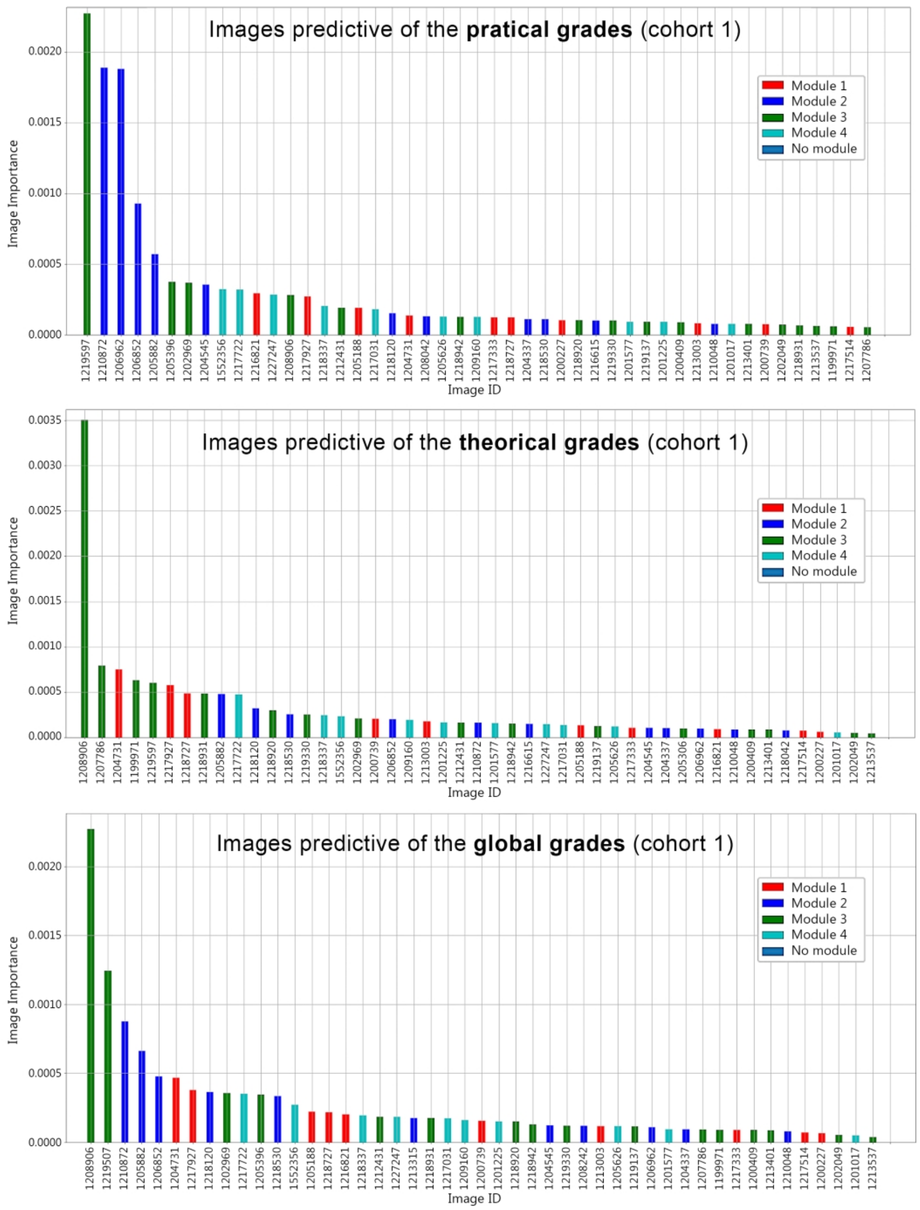


Figure 2: Digitized histological slides with predictive value for students' performance
On the y-axis, the histograms illustrate the importance of the predictive images of practical, theoretical, and global grades for cohort 1. On the x-axis, the names of the predicted slides isolated by the algorithms are situated. Histological slides identified on the y-axis were classified according to their educational purpose (Introductory, Route, or Identification) in table 1 for each cohort.

175x226mm (330 x 330 DPI)

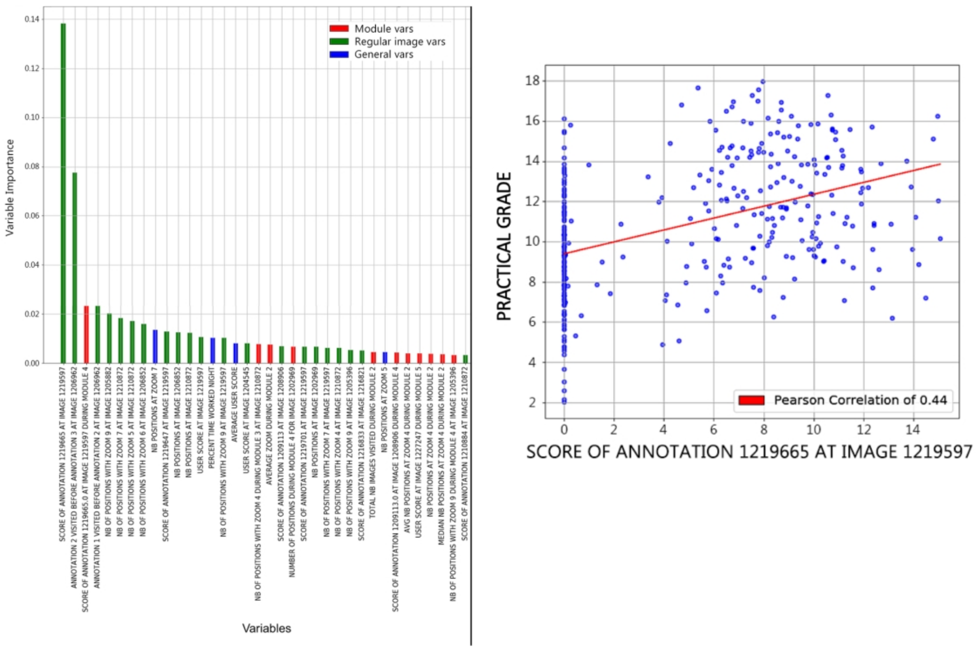


Figure 3: Predictive variables of students' examination performance, related to their behavior with digitized histological slides

The histogram corresponds to the set of predictive variables identified by the Machine Learning algorithms, presented on the x-axis, for the practice grade of cohort 1. The predictive strengths of the variables are represented on the y-axis. The correlation graph provided by the algorithms presents the most predictive variables isolated from the histogram and practical exam scores. The predictive variables (x-axis) were pooled et classified in table 2 for each cohort.

174x112mm (300 x 300 DPI)

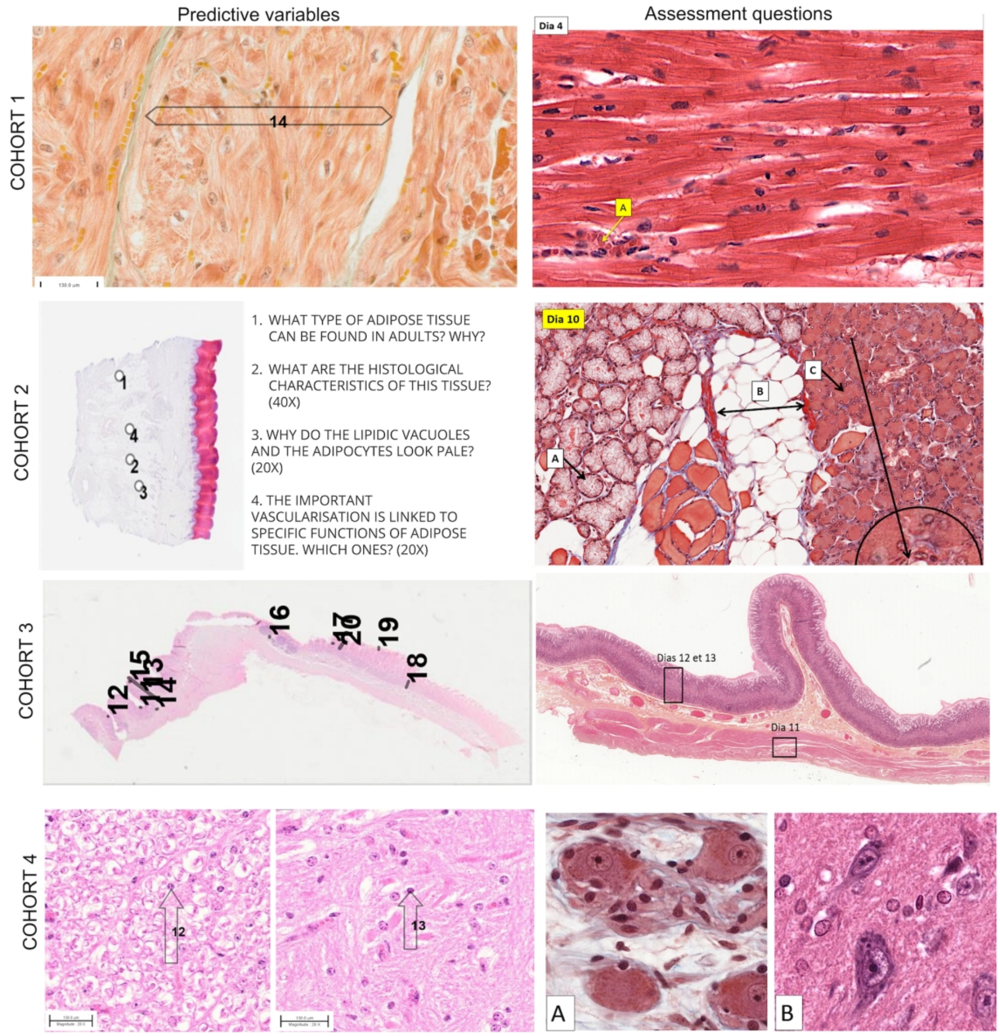


Figure 4: Relation between predictive variables provided by Machine Learning algorithms, and selected assessment questions for each cohort
For each cohort, an image, annotation, or some other predictive variable may have been associated with the subject specifically addressed by an exam question.

175x181mm (300 x 300 DPI)

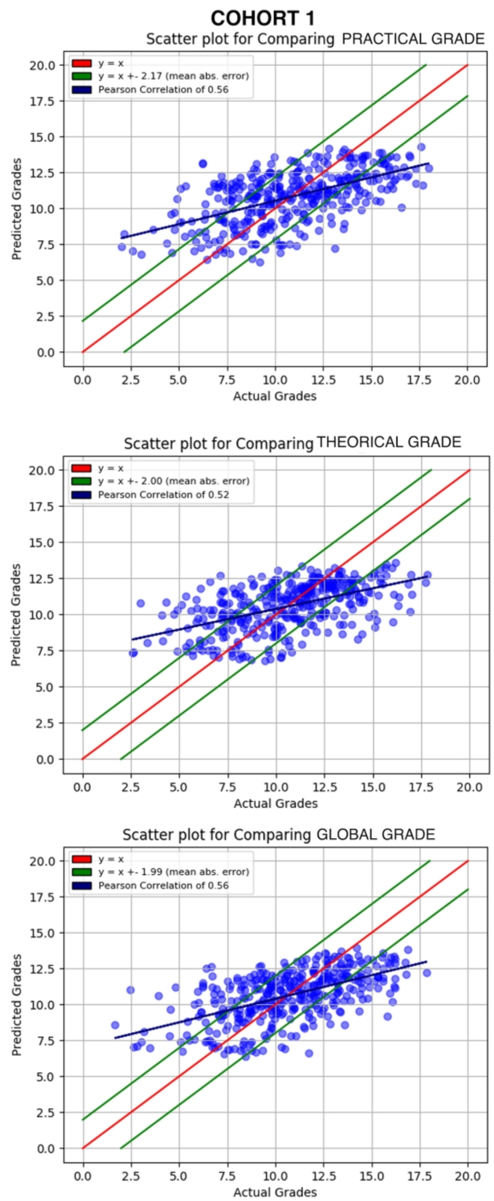


Figure 5: Scatter plot between the predicted and actual grades (practical, theorical and global) for cohort 1. For each grade, the Pearson correlation coefficient is presented in a table, along with the delta score between the actual and predicted median scores.

96x226mm (330 x 330 DPI)

| Unit | Name of the slide (reference) | Slide type (Introduction, Route (R), Identification (I)) | Cohort 1 | Cohort 2 | Cohort 3 | Cohort 4 |
|---|---|---|---|---|---|---|
| 1 | Tissue composition (1227247) | Introduction | | 1 | | |
| 2 | Small intestine (1209160) | R | | | 2 | 1 |
| | Œsophagus (1205626) | R | | 2 | | |
| | Lip (1217722) | I | | 3 | 1 | 2 |
| | Gastro-intestinal junction (1218337) | I | | 3 | 1 | 2 |
| 3 | Fundus (1217514) | R | | | 2 | |
| | Pancreas (1204731) | R | 1 | | | |
| | Lip (1217927) | I | | | | 1 |
| 4 | Small intestine (1206962) | R | 1 | | | |
| | Hypodermis (1216615) | R | | 3 | | |
| | Trachea (1206852) | R | 2 | | | |
| | Tibia (1205882) | R | 2 | 2 | | 1 |
| | Rat tail (1210872) | R | 2 | | | |
| | Trachea (1218530) | I | | | 2 | |
| | Lip (1218120) | I | | | 1 | |
| 5 | Rabbit tongue (1200409) | R | | | 1 | 1 |
| | Myocardium (1202049) | R | | | | 1 |
| | Heart (1213537) | R | | | | 1 |
| | Eyelid (1213401) | R | | 1 | | |
| | Anorectal junction (1218942) | I | | | 2 | |
| | Heart (1219597) | I | 3 | | 3 | |
| 6 | Vascular bundle (1207786) | I | 1 | | | |
| | Nerve (1208906) | I | 2 | | | |
| | Small intestine (1199971) | I | 1 | | | |
| | Bone marrow (1219137) | R | | | | 3 |

175x224mm (300 x 300 DPI)

| Performance predictor | Cohort 1 | Cohort 2 | Cohort 3 | Cohort 4 | Percentage from the total (%) | Total |
|---|---|---|---|---|---|---|
| Annotations score per image | 11 | 5 | 20 | 6 | 35,0 | 42 |
| Annotations viewing order | 4 | 5 | 2 | 12 | 19,2 | 23 |
| Number of positions with magnification on one image | 6 | 11 | 1 | 5 | 19,2 | 23 |
| Number of positions on one image | 1 | 5 | 3 | | 7,5 | 9 |
| User score | 2 | 1 | 4 | | 5,8 | 7 |
| Number of positions with magnification | 1 | 2 | | 1 | 3,3 | 4 |
| Annotation score per image over the duration of a module | 2 | | | | 1,7 | 2 |
| Time spent on an image | | 1 | | 1 | 1,7 | 2 |
| Average number of positions while magnifying | 1 | | | | 0,8 | 1 |
| Number of positions while magnifying during a module | 1 | | | | 0,8 | 1 |
| Median number of positions while magnifying during a module | 1 | | | | 0,8 | 1 |
| Medium magnification used during the module | | | | 1 | 0,8 | 1 |
| User perception data | | | | 4 | 3,3 | 4 |

175x182mm (300 x 300 DPI)