International Journal of Computer Assisted Radiology and Surgery

Design implications of repurposing a radiomics research platform for education: The case of QuantImage v2

--Manuscript Draft--

Manuscript Number:	CARS-D-24-00084
Full Title:	Design implications of repurposing a radiomics research platform for education: The case of QuantImage v2
Article Type:	Short communication
Keywords:	Artificial intelligence; Machine learning; Medical education; Oncology; Radiology; Radiomics; User study
Corresponding Author:	Jakub Mlynář HES-SO Valais-Wallis School of Management: HES-SO Valais-Wallis Haute Ecole de Gestion Sierre, SWITZERLAND
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	HES-SO Valais-Wallis School of Management: HES-SO Valais-Wallis Haute Ecole de Gestion
Corresponding Author's Secondary Institution:	
First Author:	Jakub Mlynář
First Author Secondary Information:	
Order of Authors:	Jakub Mlynář
	Roger Schaer
	Adrien Depeursinge
	Daniel Abler
	Mario Jreige
	John O. Prior
	Florian Evéquoz
Order of Authors Secondary Information:	
Funding Information:	HES-SO Valais Wallis (RCSO-ISNET- 119387) — Prof. Florian Evéquoz

Abstract:

Purpose

This paper addresses the complex relationships between AI, medical education and research, embodied in QuantImage v2, a software platform for radiomics. Existing research has indicated lack of user studies and physicians' low confidence in the diagnosis and treatment planning proposed by AI models. In this context, we explore how QuantImage, designed as a tool for radiomics research, can be repurposed for education.

Methods

Specifying the possible educational uses and required technical adjustments of QuantImage, we have organized user studies in the form of collective trial sessions at the Centre hospitalier universitaire vaudois (CHUV) in Lausanne, Switzerland. These 13 sessions, in which pairs of novice users worked with the platform together with an expert tutor, have been video recorded, transcribed, and analyzed in detail, focusing on troubles that participants encountered.

Results

Based on the analyses, we formulate actual and potential design implications. Actual changes already implemented in the platform include a paradigm shift in feature selection and highlighting central elements of the user interface, which are both

	motivated by the aims of making the work with QuantImage more accessible to users with varying levels of expertise. Potential improvements include implementation of three different modes for the use of the platform: basic, advanced, and tutorial mode, with the last one reflecting specific needs of higher education and broadening the relevance of the obtained skills beyond the particular platform. Conclusions Understanding the practical work with AI-based models is key for making radiomics respectable in oncology and radiology and overcoming the resistance to their adoption in clinical practice. Taking into account that different user groups and use scenarios place different requirements on the design of QuantImage opens the possibility of using the platform to leverage timely reflections of the practical role of AI in medicine.

Design implications of repurposing a radiomics research platform for education: The case of QuantImage v2

Author information

Jakub Mlynář*[1], Roger Schaer[1], Adrien Depeursinge[1,2], Daniel Abler[2], Mario Jreige[2,3], John O. Prior[2,3], Florian Evéquoz[1]

[1] Institute of Informatics, School of Management, HES-SO Valais-Wallis University of Applied Sciences and Arts Western Switzerland, Sierre, Switzerland; [2] Nuclear Medicine and Molecular Imaging Department, Lausanne University Hospital (CHUV), Lausanne, Switzerland; [3] Faculty of Biology and Medicine, University of Lausanne (UNIL), Lausanne, Switzerland

* Corresponding author: jakub.mlynar@hes-so.ch

Abstract

Purpose

This paper addresses the complex relationships between AI, medical education and research, embodied in QuantImage v2, a software platform for radiomics. Existing research has indicated lack of user studies and physicians' low confidence in the diagnosis and treatment planning proposed by AI models. In this context, we explore how QuantImage, designed as a tool for radiomics research, can be repurposed for education.

Methods

Specifying the possible educational uses and required technical adjustments of QuantImage, we have organized user studies in the form of collective trial sessions at the Centre hospitalier universitaire vaudois (CHUV) in Lausanne, Switzerland. These 13 sessions, in which pairs of novice users worked with the platform together with an expert tutor, have been video recorded, transcribed, and analyzed in detail, focusing on troubles that participants encountered.

Results

Based on the analyses, we formulate actual and potential design implications. Actual changes already implemented in the platform include a paradigm shift in feature selection and highlighting central elements of the user interface, which are both motivated by the aims of making the work with QuantImage more accessible to users with varying levels of expertise. Potential improvements include implementation of three different modes for the use of the platform: basic, advanced, and tutorial mode, with the last one reflecting specific needs of higher education and broadening the relevance of the obtained skills beyond the particular platform.

Conclusions

Understanding the practical work with AI-based models is key for making radiomics respectable in oncology and radiology and overcoming the resistance to their adoption in clinical practice. Taking into account that different user groups and use scenarios place different requirements on the design of QuantImage opens the possibility of using the platform to leverage timely reflections of the practical role of AI in medicine.

Keywords

Artificial intelligence; Machine learning; Medical education; Oncology; Radiology; Radiomics; User study

Statements and Declarations

The authors declare that they have no competing interests directly or indirectly related to the work submitted for publication. All funding sources have been disclosed in section "Funding Information".

Design implications of repurposing a radiomics research platform for education: The case of QuantImage v2

1. Introduction

“Artificial Intelligence” (AI) has entered numerous everyday and expert environments, including medical settings [3, 11]. It can be highly relevant for diagnostics and data analysis and has also been implemented in the field of medical imaging analysis known as *radiomics* – extracting and exploiting large-scale quantitative features from medical imaging using AI algorithms [4]. At the same time, “AI-facilitated health care requires education of clinicians” [7]. The complex relationships between medical education and research, embodied in a software platform for radiomics called QuantImage v2, provide the background and motivation for this study.

Despite recent improvements in radiomics, little is known about the physicians’ and researchers’ actual conduct while interpreting and assessing radiomic models [17], pointing to a lack of user studies [2]. One of the main challenges of adopting radiomics in clinical practice is the limited interpretability of resulting radiomic models, leading to physicians’ low confidence in the diagnosis and treatment planning proposed by the model [9]. Development of any discipline is anchored through its teaching, and it is necessary to “consider adapting the education curricula to include more radiomics related topics” [13]. Systematic implementation of radiomics in university teaching is crucial because higher education is the principal environment where future physicians and researchers can establish reliance in radiomic results, trust in AI-based systems, and become aware of their limitations.

In this paper, we explore how QuantImage, designed as a tool for radiomics research, can be repurposed for education in this emerging field. QuantImage allows extracting several types of features from PET/CT images, providing a simple and user-friendly environment that can be further adjusted for more refined analyses [1]. It allows clinical researchers with no programming background to develop and validate radiomics models using their own data, which can be easily exported from the hospital information system. While many other radiomics libraries or software are focusing solely on quantitative image feature extraction, QuantImage also includes building and validating machine learning (ML) algorithms fed with the extracted image features and provided outcomes (i.e. labels). An intermediate feature visualization functionality allows validating scientific hypotheses concerning the relationship between image features and the considered clinically relevant outcomes. The latter functionality enables not only building robust and interpretable models, but also opens avenues for educating radiologists, nuclear physicians and other clinical researchers on the significance and relevance of quantitative image (radiomics) features and models.

Our study focused on changes needed to adjust QuantImage for education. The platform was introduced in trial sessions to explore its possible use in education by identifying aspects of QuantImage that can be utilized and modified for teaching purposes. Analysing video materials from the trial sessions described below, some of our questions were: Which aspects of the platform are suitable for research but not necessarily for education? Which functionalities are unclear for novice users? What kinds of misunderstandings occur? On this background, the applied objective of our study was to transform our findings from the trial sessions into specific design implications.

2. Methods

Specifying the possible educational uses and technical adjustments required for QuantImage [1], we have organized user studies in the form of collective trial sessions at the Centre hospitalier universitaire vaudois (CHUV) in Switzerland. In these sessions, novice users (students of medicine, physicians and technicians) reviewed and interpreted radiomic results together with an expert tutor. Their tasks were based on developing a diagnostic model of Pulmonary Lymphangitic Carcinomatosis (PLC), a condition linked to very poor prognosis in the context of non-small-cell lung cancer. Despite being associated with increased peritumoral uptake in 18F-FDG PET as well as peribronchovascular thickening in CT images, diagnosing PLC is very difficult for human readers/radiologists and usually requires an invasive and risky lung biopsy [6]. A collection of >100 cases with annotated Volumes Of Interest (VOIs) was assembled at CHUV. The Quantimage's Feature Explorer can be used to identify the three features that are most predictive of PLC. While an initial model based on all 630 radiomic features led to a cross-validated AUC of 0.74, the model based on the identified feature subset resulted in an improved AUC of 0.81 while using fewer radiomics features. The three identified features (SUVmax and two CT texture parameters) are consistent with the hypothesis that increased peritumoral uptake in 18F-FDG PET results in higher SUVmax and peribronchovascular thickening in CT images is altering the parenchymal CT texture, thus indicating an interpretable and trustworthy radiomics prediction model.

In 2022/2023, we organized 13 trial sessions that were video recorded from two complementary angles. Video-based studies of healthcare and hospital contexts, including oncology and radiology, constitute an established area [8, 14, 15]. The major advantage of observational methods is that they allow for discovery of practices that are normally taken for granted and therefore not accessible by questionnaire/interview methods. The video recordings were transcribed and analyzed in accordance with ethnomethodological conversation analysis [5]. All participants signed an informed consent and the transcripts have been anonymized. In the analysis, we have singled out consequential interactional phenomena while focusing on troubles encountered by the novice users, examining the interplay of talk, bodily conduct and material artifacts as the novices and experts interact with each other and the platform [10]. Based on our participation in the sessions and analysis of the video recordings, we formulate design implications for an educational version of QuantImage, leveraging on widely used heuristic usability principles [12] (see Table 1).

#	Principle	Description
1	Visibility of system status	Users should be informed about what is going on in an appropriate and timely fashion.
2	Match between system and the real world	The language used should be familiar to the user.
3	User control and freedom	There should be a way to return to a previous state of work after making a mistake.
4	Consistency and standards	Words and actions should have the same meaning across the platform.
5	Error prevention	Problems should be prevented from occurring.
6	Recognition rather than recall	Elements of design should be visibly recognizable, without needing to remember them.
7	Flexibility and efficiency of use	Interaction with the platform should be fluent and smooth.
8	Aesthetic and minimalist design	Only relevant information should be provided.
9	Recognize, diagnose, and recover from errors	Errors should be described in plain language and offer a solution.
10	Help and documentation	No need for additional explanation is ideal, but documentation should be provided.

Table 1. Summary of Nielsen's usability heuristics, based on [12].

3. Results

We present two sets of improvements to QuantImage as an outcome of the semi-experimental trial sessions. The first set (3.1) has been implemented in QuantImage in response to troubles and problems encountered by novice users. The second set (3.2) is presented as a list of potential changes that would need to be implemented in case the platform is systematically incorporated in education. We refer to numbered items from Table 1 as "principles".

3.1 Actual improvements: Changes implemented on the fly, based on feedback and observations from the trial sessions

***Feature selection paradigm shift**:* originally, feature selection was limited, as the user could only select modalities, VOIs and radiomics features independently, e.g. enabling or disabling shape features completely across all modalities and VOIs (see Figure 1). It became apparent that this was not sufficient, as users wanted to select features in a more fine-grained manner (e.g. selecting shape features only for a given modality and VOI). This feedback led to a significant transformation of the feature selection process, moving away from simple, independent checkboxes towards a hierarchical tree structure. This enabled users to select features at several levels of granularity (modality, VOI, feature family, feature subgroups, individual features), as well as across hierarchical levels (e.g. deselect shape features for all modalities and VOIs at once). This was a major change that is also highly domain-specific to radiomics, therefore not easily captured by generic usability heuristics; however, the most fitting one would be principle 7.

Figure 1 – (A) Old feature selection paradigm, which did not allow for a combination of parameters (e.g. use VOI "GTV N" only for modality "CT"). (B) New feature selection paradigm, which allows much more fine-grained feature selection.

Show description of features when hovering over their name*:* A comment that was given several times during the trials is that users generally do not have an exhaustive understanding of all the available radiomics features. Responding to principles 2 and 10, an improvement was therefore developed to show a description when the user hovers over a feature to give some insight as to what it represents (see Figure 2). The descriptions were derived from an established ontology [16].

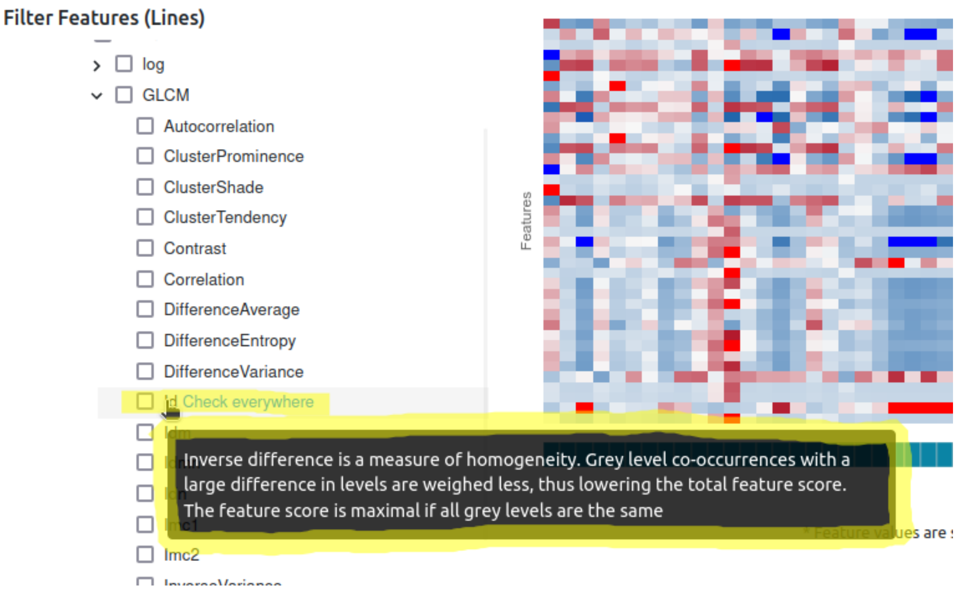

Figure 2 – Within the Feature Explorer Visualization tab, when a user hovers over the name of a radiomics feature, a more detailed definition is shown.

Highlight important user interface (UI) elements*:* Certain UI elements that are an integral part of the feature selection and model training workflow, such as the button that saves a selected feature set as a new collection have been adapted to be more prominent. Indeed, it was observed several times that despite being a mandatory step for training a model with a set of selected features, users were not aware that they needed to click on a link to create a new collection. This is connected to principles 1, 5 and 8. The link was therefore replaced by a larger and more visible button. Other mechanisms were implemented to make it clearer that feature collection creation is required before model training (e.g. warn the user before he navigates away from an in-progress selection, so it is not lost). Another UI improvement, based on principles 3 and 5, was making the button to undo a feature selection step more visible and displaying it in the section allowing for automatic feature selection (dropping correlated features and keeping only the N best-ranked features).

Adding new features to a created collection*:* Another functionality that was requested several times is the ability to add back features that were previously removed when creating a new feature collection. Indeed, it was initially only possible to filter out features, but not add back removed features. Aligned with principles 6 and 7, this made it quite tedious to create a new collection that included all the features of an existing collection, but with one additional feature (the user had to go back to the original feature set and manually reproduce the slightly modified version, requiring him/her to remember all the features selected previously). Users can now always see the complete feature tree and add removed features back to a collection (see Figure 3).

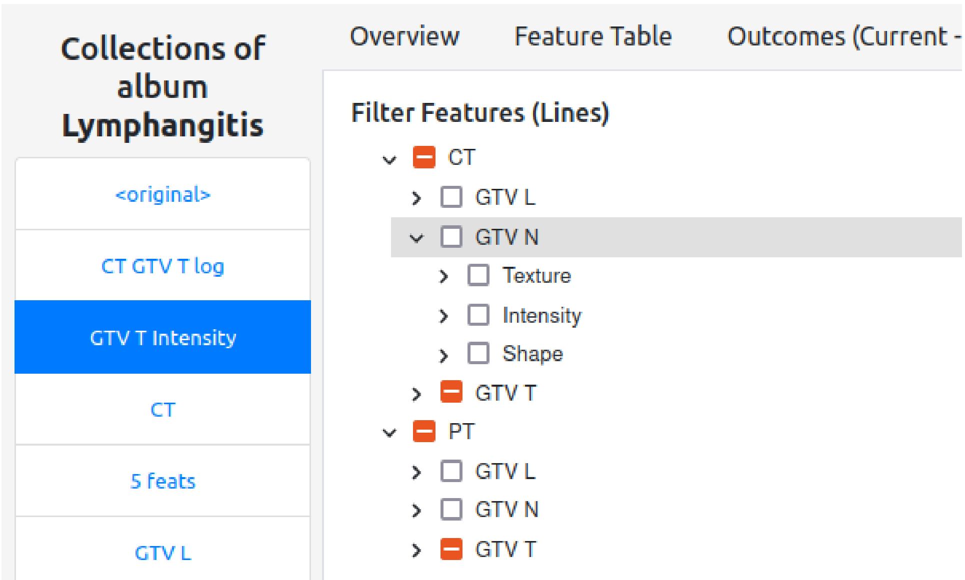

Figure 3 – Within the Feature Explorer, with a feature collection selected, the user can easily add back features that were previously removed when creating the collection.

***Remove a feature by clicking on it in the feature heatmap**:* Another minor improvement to the feature selection process was suggested during a trial. The user, having identified a feature to remove, wanted to simply click on it in the feature heatmap to deselect it, rather than look for it in the adjacent feature tree. This change, related to principles 7 and 8, was implemented to make it easier to remove unwanted features that are visually identified in the heatmap.

3.2 Potential improvements: Future substantial changes that could be implemented in the platform to improve its use in education

***Implementing different modes in the platform**:* It could be interesting to have different modes for the platform, based on the users' proficiency with radiomics, ML, statistics, etc. Using the platform as a clinical research tool would benefit from distinguishing a "basic" mode that does not display advanced options or complicated concepts, and an "expert" mode that gives users more control on all the intricacies of the feature extraction and model training processes (e.g. configuring the type of cross-validation performed). There are already some advanced items available in the feature extraction screen, where an experienced user can define a custom configuration to set parameters related to image padding, normalization, etc. This could be extended to other aspects of the platform and streamlined into global modes that affect various parts of the UI, responding to several of Nielsen's heuristic principles (mainly 5, 7, 8 and 10).

***Developing a "tutorial" mode**:* A current bottleneck for expanding the use of the platform as an educational tool is the introduction of the various features, which were explained and demonstrated at the start of every trial session so far. A video highlighting the main elements of the platform is also available, but ideally there should be an explanatory "tutorial" mode that would guide a new user step-by-step through all the main elements of the platform interactively. Using a demo dataset, the user would have to first extract radiomics features using a specific configuration and set of VOIs, then upload patient outcomes for a sample classification use case, visualize the feature heatmap and create a new collection using a subset of features, and finally training a machine learning model using those features. The tutorial, in line with principles 5 and 10, should ensure that the user performs all the required actions in the correct order, as well as provide instructions and explanations for each step. Most importantly, this mode would also allow the user/student to reflect on the more generally applicable aspects of using AI-based technology in medical work, and to make practical sense of working with AI in medicine while understanding its limits and potential.

4. Conclusion

Understanding the practical work with AI-based models is key for making radiomics respectable in oncology and radiology and overcoming the resistance to their adoption in clinical practice. Reporting on an observational, semi-experimental study with pairs of novice users that explores the possible educational use of the radiomics platform QuantImage v2, we have reflected on actual improvements that were made on the basis of the trial sessions, as well as potential improvements of the platform that would be needed to make it a full-fledged educational tool. Our findings were based on observing recurrent troubles in the users' collaboration that were formulated by themselves as part of their work or noticed by onlooking designers of the platform.

Changes already implemented in the platform include a paradigm shift in feature selection and highlighting central elements of the user interface, which are both motivated by the aims of making the work with QuantImage more accessible to users with varying levels of expertise. At the same time, we have also realized that different user groups and use scenarios place different requirements on the design of the platform. In future versions of QuantImage, the requirements of educational settings would be best taken into account by implementing three different modes in the platform: basic, advanced, and tutorial mode. *Basic mode* would include minimal information required for skilled operation, automatizing most of the radiomics workflow. *Advanced mode* would allow the users to adjust settings of some of the processes automatized in the basic mode. Finally, *tutorial mode* – specifically designed for the use of the platform in higher education – would include some of the more general information about radiomics that is taken for granted by experts and not included in basic or advanced mode. It would also prompt users to reflect on the limitations and potential of AI-based technologies, broadening the relevance of the obtained skills beyond the single software platform to the implementation of ML and AI in radiology more widely. This opens the possibility of using QuantImage for leveraging broader knowledge of this kind, more than timely in the current era when AI-powered devices are becoming ubiquitous.

References

[1] Abler D, Schaer R, Oreiller V, Verma H, Reichenbach J, Aidonopoulos O, Evéquoz F, Jreige M, Prior JO, Depeursinge A (2023) QuantImage v2: A comprehensive and integrated physician-centered cloud platform... . European Radiology Experimental 7. https://doi.org/10.1186/s41747-023-00326-z

[2] Antoniadi AM, Du Y, Guendouz Y, Wei L, Mazo C, Becker BA, Mooney C (2021) Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: A systematic review. Applied Sciences 11: 5088. https://doi.org/10.3390/app11115088

[3] Briganti G, Le Moine O (2020) Artificial Intelligence in Medicine: Today and Tomorrow. Frontiers in Medicine 7: 27. https://doi.org/10.3389/fmed.2020.00027

[4] Guiot J, Vaidyanathan A, Deprez L, Zerka F, Danthine D, Frix A-N, Lambin P, Bottari F, Tsoutzidis N, Miraglio B, Walsh S, Vos W, Hustinx R, Ferreira M, Lovinfosse P, Leijenaar RTH (2022). A review in radiomics: Making personalized medicine a reality... . Medicinal Research Reviews 42: 426–440. https://doi.org/10.1002/med.21846

[5] Heath C, Hindmarsh J, Luff P (2010) Video in Qualitative Research. Sage, London

[6] Jreige M, Dunet V, Letovanec I, Prior JO, Meuli RA, Beigelman-Aubry C, Schaefer N (2020) Pulmonary Lymphangitic Carcinomatosis: Diagnostic Performance of High-Resolution CT... . J Nucl Med 61: 26-32. https://doi.org/10.2967/jnumed.119.229575

[7] Keane PA, Topol EJ (2021) AI-facilitated health care requires education of clinicians. The Lancet 397. https://doi.org/10.1016/S0140-6736(21)00722-4

[8] Keel S (ed) (2023) Medical and Healthcare Interactions. Routledge, London

[9] Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, Sun K, Longfei L, Li B, Wang M, Tian J (2019). The applications of radiomics in precision diagnosis and treatment of oncology: Opportunities and challenges. Theranostics 9: 1303–1322. https://doi.org/10.7150/thno.30309

[10] Mlynář J, Depeursinge A, Prior JO, Schaer R, Martroye de Joly A, Evéquoz F (2023) Making sense of radiomics: Insights on human–AI collaboration... . Frontiers in Communication. https://doi.org/10.3389/fcomm.2023.1234987

[11] Nardini C (2020) Machine learning in oncology: A review. Ecancermedicalscience 30: 1065. https://doi.org/10.3332/ecancer.2020.1065

[12] Nielsen J (1993) Usability Engineering. Academic Press, San Diego

[13] Peeken JC, Bernhofer M, Wiestler B, Goldberg T, Cremers D, Rost B, Wilkens JJ, Combs SE, Nüsslin F (2020) Radiomics in radiooncology: Challenging the medical physicist. Physica Medica 48: 27–36. https://doi.org/10.1016/j.ejmp.2018.03.012

[14] Rystedt H, Ivarsson J, Asplund S, Johnsson AÅ, Båth M (2011) Rediscovering radiology: New technologies and remedial action... . Social Studies of Science 41: 867–891. https://doi.org/10.1177/0306312711423433

[15] Singh S, Cortez D, Maynard D, Cleary JF, DuBenske L, Campbell TC (2017) Characterizing the nature of scan results discussions: Insights into why patients misunderstand their prognosis. Journal of Oncology Practice 13: 231–239. https://doi.org/10.1200%2FJOP.2016.014621

[16] Shi Z, Traverso A, van Soest J, Dekker A, Wee L (2019) Technical Note: Ontology-guided radiomics analysis workflow (O-RAW). Med Phys 46: 5677–5684. http://doi.org/10.1002/mp.13844

[17] Verma H, Mlynář J, Schaer R, Reichenbach J, Jreige M, Prior JO, Evéquoz F, Depeursinge A (2023) Rethinking the role of AI with physicians in oncology: Revealing perspectives from clinical and research workflows. In Proceedings of CHI '23, Article 17. New York: ACM. https://doi.org/10.1145/3544548.3581506