QUANTITATIVE ASSESSMENT OF MYOCARDIAL FIBROSIS BY DIGITAL IMAGE ANALYSIS: *an adjunctive tool for pathologist "ground truth"*

João Abecasis[1,2], MD, joaoabecasis@hotmail.com
Nuno Cortez-Dias[3], MD, PhD, cortezdias@yahoo.com
Daniel Gomes Pinto[2,4], MD, danielgomespinto@gmail.com
Pedro Lopes[1], MD, pedro_fagalopes@hotmail.com
Márcio Madeira[5], MD, madeira.marcio@gmail.com
Sancia Ramos[6], MD, sframos@chlo.min-saude.pt
Victor Gil[7,8], MD, PhD, victorgilmd@gmail.com
Nuno Cardim[2], MD, PhD, cardimnuno@gmail.com
Ana Félix[2,9], MD, PhD, ana.felix@nms.unl.pt

[1]Cardiology Department, Hospital de Santa Cruz, Lisboa, Portugal
[2] Nova Medical School, Lisboa, Portugal
[3] Cardiology Department, Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte, Faculdade de Medicina de Lisboa, Portugal
[4] Pathology Department, Hospital Garcia de Orta, Almada, Portugal
[5] Cardiac Surgery Department, Hospital de Santa Cruz, Lisboa, Portugal
[6] Pathology Department, Hospital de Santa Cruz, Lisboa, Portugal
[7] Hospital da Luz, Lisboa, Portugal
[8] Faculdade de Medicina, Universidade Católica, Lisboa
[9] Pathology Department, IPOFG, Lisboa, Portugal

Short title: Myocardial fibrosis quantification by an *AI* algorithm

Corresponding Author
João Abecasis, MD
joaoabecasis@hotmail.com
+351 914054977
Calçada da Palma de Baixo, 8, 4º B, 1600-175 Lisboa

Type of article: original article
Word Count: 6007 words.

ABSTRACT

AIMS: Myocardial fibrosis (MF) is a common pathological process in a wide range of cardiovascular diseases. Its quantity has diagnostic and prognostic relevance. We aimed to assess if the complementary use of an automated artificial intelligence software might improve the precision of the pathologist´s quantification of MF on endomyocardial biopsies (EMB).

METHODS AND RESULTS: Intraoperative EMB samples from 30 patients with severe aortic stenosis submitted to surgical aortic valve replacement were analysed. Tissue sections were stained with Masson´s trichrome for collagen/fibrosis and whole slide images (WSI) from the experimental glass slides were obtained at a resolution of 0.5µm using a digital microscopic scanner. Three experienced pathologists made a first quantification of MF excluding the subendocardium. After two weeks, an algorithm for Masson´s trichrome brightfield WSI (at *QuPath* software) was applied and the automatic quantification was revealed to the pathologists, who were asked to reassess MF, blinded to their first evaluation. The impact of the automatic algorithm on the inter-observer agreement was evaluated using Bland-Altman type methodology.

Median values of MF on EMB were 8.33% [IQR 5.00-12.08%] and 13.60% [IQR 7.32-21.2%], respectively for the first pathologist´s and automatic algorithm quantification, being highly correlated (R^2: 0.79; $p<0.001$). Inter-observer discordance was relevant, particularly for higher percentages of MF. The knowledge of the automatic quantification significantly improved the overall pathologist´s agreement, which became unbiased within the spectrum of MF severity.

CONCLUSIONS: The use of an automated artificial intelligence software for MF quantification on EMB samples improves the reproducibility of measurements by experienced pathologists. By improving the reliability of the quantification of myocardial tissue components, this adjunctive tool may facilitate the implementation of imaging-pathology correlation studies.

LIST OF ABBREVIATURES

MF: myocardial fibrosis

ECM: extracelular matrix

AS: aortic stenosis

EMB: endomyocardial biopsy

AVR: aortic valve replacement

CMR: cardiac magnetic resonance

ECV: Extracellular volume

WSI: whole slide images

INTRODUCTION

Myocardial fibrosis (MF), meaning the excessive deposition of extracellular matrix (ECM) components in the myocardium, is a common pathological process in a wide spectrum of chronic heart diseases [1,2]. Being associated with the disruption of normal myocardial structure, it is behind the mechanistic base for adverse cardiac remodeling [3]. Indeed, it is a key contributor to heart failure and its progression, having recognized prognostic implications in both ischemic and non-ischemic cardiac conditions [4]. In this regard, several attempts have been recently made in trying to integrate cardiac fibrosis assessment into distinct clinical scenarios, namely for heart failure management, patients risk stratification and even therapeutic intervention before symptoms development [5].

Endomyocardial biopsy (EMB) histopathology remains the gold-standard method to diagnose MF. Additionally, in theory, the application of histomorphometry parameters at specific fibrosis-stained sections is the most accurate technique for MF quantification [6]. However, EMB samples may not be representative of a non-uniform, sometimes patchy, pathological process. Besides, it has low feasibility, as determined by limited availability, invasiveness and need for expertise in cardiac pathology interpretation [1]. Thereby, non-invasive imaging methods have been developed.

Multiparametric cardiac magnetic resonance (CMR) is currently the best imaging modality that offers a direct, whole heart assessment of myocardial fibrosis. Myocardial T1 mapping and associated techniques yield improved myocardial characterization through the ability to quantify signal intensity for each voxel in the myocardium[7]. Extracellular volume (ECV) fraction, as derived from combined pre- and postcontrast T1 mapping, seems to be particularly sensitive to extracellular space expansion. Previous single centre studies in patients with severe aortic stenosis (AS) have shown good correlations between ECV and histological diffuse fibrotic burden, as assessed by quantitative morphometry [7] [8]. As for other clinical contexts,

ECV obtained using CMR is poised as an ideal imaging surrogate marker for predicting diffuse MF [9]. Nevertheless, correlation results from distinct clinical cohorts are far from uniform, and some of these studies did not find associations between the ECV and collagen volume fraction in EMB samples [10]. Myocardial infiltration, oedema, and inflammation remain important potentially confounding sources of increased ECV, as imaging markers do not measure fibrous tissue directly, but the total interstitial space instead. Additionally, there are qualitative aspects related to the composition of fibrous tissue, such as the type of collagen fibres and molecular organization, with potential impact on non-invasive assessment of extracellular myocardial component [1]. Finally, and no less than important, these correlation studies with histopathology rely on the accuracy of MF quantification provided by the pathologist, which is totally operator-dependent and highly sensitive to the experience level [11]. Recently, digital algorithms for automatic morphometry started to be developed, aiming to improve the reproducibility of MF quantification. Such complementary tools would fulfil an unmet need, increasing the precision of MF invasive quantification.

Our aim is to assess if the use of an automated artificial intelligence software for the quantification of MF on EMB, might improve individual pathologist´s quantification and agreement, in terms of precision.

METHODS

1. Study design

Present analysis was conducted on EMB samples from 30 patients undergoing elective surgical aortic valve replacement at our tertiary center between April 2019 and January 2022 because of isolated severe symptomatic AS, defined according to European guidelines on valvular heart disease [12]. This is part of a correlation research protocol involving both pre- and post-operative LV structural and functional assessment by multimodality imaging and myocardial histopathology study from EMB, specially addressing cardiomyocyte adaptation and extracellular matrix remodeling and fibrosis. Study approval was granted by the ethical committee of Nova Medical School University (number *61/2018/CEFCM*) conforming to the principles of the Helsinki declaration. All participants gave written informed consent before inclusion.

We excluded patients with congenital AS or previous diagnosis of sub/supra valvular aortic stenosis, concomitant severe non-aortic valve dysfunction, moderate and severe aortic regurgitation, previous cardiac surgery, active endocarditis, previous history of myocardial infarction, myocarditis, ischemic and non-ischemic cardiomyopathy including cardiac amyloidosis and other infiltrative diseases, chronic kidney disease with glomerular filtration rate below 30mL/min/1.73m^2, non-cardiac inflammatory disease, active infection, under immunosuppressive and chronic anti-inflammatory therapy, under chemotherapy and with previous chest radiotherapy.

EMB samples were obtained either from intraoperative septal biopsy as per protocol design (harvested with a scalpel from the basal interventricular septum, preferably with endocardium included) or from complementary septal myectomy, performed by the surgical team at the time of surgical AVR because of asymmetric septal hypertrophy or at surgeon´s discretion.

From a total of 150 patients included in the study protocol, 119 (79.3%) were submitted to AVR and EMB tissue samples were obtained in 112. Biopsy was not performed in 6 patients owing to the reported risk from the surgical team (thin interventricular septum); 1 patient had a very small biopsy sample with too scarce myocardium for analysis. For this particular analysis we used the EMB samples from 30 patients randomly selected from the overall study population (Figure 1).

2. Histomorphological studies

In a pre-analytical phase, biopsies were fixed in 10% neutral buffered formalin (*JTBaker*®) for 24 to 48 hours at room temperature (20ºC). Formalin fixed, paraffin (*VWR International*, USA) embedded tissues were processed in *Sakura´s "Tissue-Tek VIP"* and cut into 3µm thick sections. Tissue section adhesion time and temperature were constant for one hour at 70ºC. Sections were stained with Masson´s trichrome for collagen/fibrosis assessment according to a standard protocol. Whole slide images (WSI) from the experimental glass slides were obtained at a resolution of 0.5µm using a digital microscopic scanner (ScanScope® AT2 brightfield scanner TM, Aperio Technologies, Vista, CA, USA) at a 40x objective magnification, and stored in a tiled TIFF or PNG format (Figure 2A). The actual scan resolution (effective pixel size in the sample plane) at 40x is 0.25µm/pixel. The accompanying software allows the user to navigate through the captured WSI at any given digital zoom up to 540x.

3. Pathologists´ visual analysis and quantification

The extent of cardiac fibrosis was independently evaluated, at the same 30 sections from WSI, by three highly experienced pathologists (more than five years of clinical experience with routine evaluation of cardiac pathology). A five percent interval quantification was attributed. After automatic algorithm quantification (see below), each one of the cases with automatic

quantification, expressed as a %, was shown to the pathologists. A new blinded quantification of the same cases was requested and registered for all the pathologists after two weeks (Figure 1). Dense endocardial fibrosis was excluded from both visual analysis and automatic quantification (Figure 2 B).

4. Automatic quantification algorithm, artificial intelligence

The algorithm for automatic image processing, analysis and quantification algorithm was developed using the software platform *QuPath 0.3.0*. This widely used software, short for Quantitative Pathology, allows the development of tools for digital pathology image analysis, being specifically designed for WSIs without the need for cropping or down-sampling images [13] [14]. It allows the analysis of immunohistochemistry (brightfield or fluorescent) and haematoxylin and eosin (H&E) images, through a pattern recognition deep learning algorithm that distinguishes spatial and morphological features based on structures (classes) provided by the user [15]. Ultimately, the definition of specific thresholds for specific tissue compartment (for instance collagen fibers, vessel lumen, cardiomyocytes) and empty spaces, provides the base for quantification and proportion estimation (applying pixel counting and boundary detection, using specific colour values).

Our specific algorithm/classifier for the quantification of MF on *Masson´s* Trichrome brightfield images was developed as follows (transition from WSI through algorithm classification in Figure 3 and 4): 1) new project and training set created; 2) digital slides added to a training set; 3) colour stain vectors estimation – modal *RGB* (Figure 3 A); selection of objects for analysis, excluding dense endocardial fibrosis (Figures 3 C and 4 B); 4) pixel classification loading according to specific classifier, previously created and tested ("*Monsta"*); specific task application algorithm for automatic analysis (Figure 3 D, E and 4 C, D). For quantification, the total tissue component was split into fibrosis and a remaining part, containing

cardiomyocytes, empty spaces, vessels, blood cells (detailed identification of fibrosis in white and empty spaces in orange/red in Figure 3B). All components were assessed in pixels and a final table containing both absolute areas and correspondent proportions is displayed [Figure 4 A; percentage of fibrosis = (area of fibrosis / area of cardiomyocytes + area of fibrosis) x 100]. Figure 4 depicts specific cases where the algorithm proved to appropriately identify and classify other tissue components beyond fibrosis, correctly integrate artifacts on image analysis and provide better identification of an extensive diffuse pattern of fibrosis, inherently interpreted as a challenging quantification.

The whole analysis was performed on i*Mac 27*® (3.8 GHz, Intel Core i7, 8-GB RAM).

5. Statistical analysis

Categorical variables are presented as count (percentage) and difference between groups were analysed by chi-square or Fisher´s exact tests. Continuous variables with normal and non-normal distribution were described by the mean ± SD or median ± interquartile range (IQR), and compared using Student´s t test or Mann-Whitney U test, respectively. Analysis of variance and Kruskal-Wallis testing were performed for comparison among multiple groups. For paired data – first and second quantification from each pathologist, paired Student's t-test and Wilcoxon signed rank test were used.

The agreement between the quantification of fibrosis by the three pathologists was evaluated using Bland-Altman plots, including the estimation of interobserver differences and search for systematic proportional bias using linear regression analysis [16]. Additionally, reliability analysis was used to further characterize the interobserver agreement both at individual and average measurement level.

A gold-standard MF quantification was calculated as the average of the three individual pathologists' quantifications. The correlation between the quantification of fibrosis by the

automatic algorithm and the gold-standard value was assessed using linear regression analysis. Similarly, linear regression analysis was used to evaluate the correlation between each pathologist and the gold-standard MF quantification.

To compare the agreement of the three pathologists' quantification before and after knowing the result of the automatic algorithm (first and second observation), we adopted a modification to the Bland-Altman plot methodology for the construction of just one plot: the difference between each observer and the average quantification for the subject (gold-standard value) was estimated, rather than the mean difference between each two observers. The 95% limits of agreement with the mean are estimated as $\pm 1.96 \times s$, s being an estimate of the standard deviation for all possible observers and that can be evaluated as the square root of the variance of the differences plot [17].

A two-sided p-value <0.05 was considered statistically significant. The statistical analysis was performed with IBM SPSS Statistics 26.0 (IBM Corp, Armonk, NY, USA).

RESULTS

1. Population characteristics

This study was conducted on EMB samples from 30 patients undergoing elective surgical AVR, with a median age of 73 (68-77) years. Demographic and clinical characteristics of the population are presented in Table 1.

2. Myocardial fibrosis quantification by the Pathologists versus the Automatic Algoritm (QuPath)

MF was quantified in the tissue samples by the three experienced pathologists (Table 2). '*Gold-standard*' MF value, defined as the average of the experienced pathologists' initial quantification, had a median value of 8.33% [IQR 5.00-12.08%]. Overall, the evaluations were highly correlated (Figure 6), but at an individual sample-to-sample level, significant discordance between operators was noticed, ranging the interobserver mean difference from -2.42 to 14.00 ($p<0.001$ to $p=0.044$).

The automatic *QuPath* algorithm was able to quantify MF in all the tissue samples and the median MF quantification was 13.60% [IQR 7.32-21.2%]. The *QuPath* MF measurements were highly correlated with the "gold-standard" pathologist quantifications (R^2: 0.79; $p<0.001$ – Supplemental Figure 1). At an individual sample-to-sample level, the *QuPath* MF estimation was significantly higher than the 'gold-standard' pathologists quantification ($p<0.001$), but the mean difference was just 3.91% (therefore smaller than the maximum interobserver variability) and presented no evidence of proportional bias ($\beta=0.015$; $p=0.87$).

3. Impact of the automatic quantification of fibrosis in the experienced interobserver agreement

The interobserver agreement regarding MF quantification, evaluated by the interobserver difference and intraclass correlation coefficients, was at most moderate at the initial evaluation – Table 3 and 4. In detail, the intraclass correlation coefficients ranged at the level of the average of operator measurements between 0.43 and 0.78; and were notably worse at the sample-by-sample individual level, ranging between 0.27 and 0.64. Furthermore, the discordance between experienced operators, as evaluated by Bland-Altman analysis, significantly increased with the increased severity of tissue fibrosis ($p<0.001$).

After 2 weeks, the automatic MF quantification results were revealed to the experienced pathologists, who were asked to re-quantify MF, blinded to their initial evaluation. The interobserver agreement improved to good or excellent in all cases – Figure 7. Indeed, the intraclass correlation coefficients increased both at the average level to 0.84-0.93 and at the individual sample-by-sample level to 0.73-0.87. In addition, the impact of the MF severity in the interobserver disagreement was noticeably attenuated.

DISCUSSION

The main findings of our study were that: (1) the use of an automatic algorithm as an adjunctive tool for MF quantification at individual EMB samples improves the overall agreement between experienced pathologist´s quantifications, indistinctive of the MF severity; and (2) in the absence of such tool, the inter-observer variability is relevant and, importantly, the inter-observer discordance increases for higher percentages of fibrosis. Finally, to the best of our knowledge this is the first study to address the use of *QuPath* to automatically quantify MF.

Last decade investigations have proved to revolutionized what used to be illustrative microscopic imaging, merely escorting cardiac research articles focused on physiological, biochemical, or molecular biological techniques [18]. Recent developments in cardiac imaging and in what specifically regards the study of cardiac fibrosis, led to the surge of comparative studies with histopathology of myocardial samples [19]. These started to be comparisons of late gadolinium enhancement and T1 mapping techniques, as markers of replacement and interstitial fibrosis, respectively, with visual score based semi-quantifications by the pathologist. However, comparisons of results between institutions were often considered inaccurate, as supposed to be highly dependent on the analyst, thus subjective and hardly reproducible. Because of this, distinct studies underpinned the comparison with histology on several morphometric methods. Tissue quantification estimated as ratios, expressed as per unit of the reference areas and volumes such as collagen volume fraction, soon gained ground, gradually becoming increasingly computed and automatic based assessments [20]. In effect, stereology is the gold standard method for quantification on microscopic sections, but its use on cardiac research is scarce owing to the complexity of sampling methodology and extensive time of analysis [21]. Because of this, both semi-automatic and automatic analysis of digital myocardial microscopy were compared with stereology, which proved to be strongly correlated [6,22]. Notwithstanding, inconsistencies between studies that use distinct semi-automatic and automatic methods for

fibrosis quantifications are still being reported[1], and these may be mirrored from divergent correlation results with non-invasive imaging markers.

The *QuPath* algorithm used in this study was able to produce a MF quantification highly correlated with the average value obtained from three experienced pathologists, here considered as our intra-study "gold-standard" evaluation. Although the *QuPath* algorithm produced slight overestimation of MF, the mean difference was just 3.91%, noticeably smaller than the maximum interobserver variability. Furthermore, its accuracy was consistent all over the spectrum of MF severity. Nevertheless, our aim was not to develop a tool to replace the pathologist´s intervention in MF quantification, but to test a new bioimage analysis software, created to improve the inter-observer quantification discrepancy. A single cohort evaluation was narrowly designed to cancel additional sources of recognized bias for quantification, such as EMB location, laboratorial processing and potential influence of tissue shrinkage, sample thickness, type of stain, and modes of digitalization [21]. Our findings are new as they showed that the application of an open software framework, *QuPath*, with extensive visualization tools and capacity for the analysis of a specific biomarker expression on WSI, significantly improves the overall agreement of MF quantification by traditional pathological assessment.

Our most important methodological commitment was to use *QuPath*, an extensible platform for both pathologists and researchers, to specifically address a commonly inquired biomarker in cardiac imaging correlation studies. This was made instead of using other open-source tools or web platforms that respectively rely on comprised libraries or collaborative analyses [13]. These last resources have proved to be useful and yielded feasible analyses across cardiac imaging correlation studies [9,10]. However, they lack immediate wide availability for researchers and sometimes rely on laborious manual evaluations for image down sampling and cropping, which may introduce variability and restrict the reproducibility of the analysis. The main advantage of *QuPath´*s use line in its ability to create and apply similar analysis in a reproducible batch

processing manner, as a machine learning method. This rests on the creation of individual hierarchical “object” based data model, with the assignment of classifications, measurements and links that allow fast query and manipulation through built-in command and scripting [13,14]. Our created project workflow was proficient for the automatic quantification of MF, as it has good correlation to the first quantification of the individual pathologist. More important, it reduced inter-observer variability of a second quantification after knowledge of the automatic results i.e., increased the precision of the measurements, becoming unbiased according to the degree of MF. Due to inherent methodological restrains we did not specifically assess the application of the algorithm to EMB of normal myocardium. This could have been useful in trying to evaluate the potential added value and consistency of the automatic algorithm for the quantification of MF in patients with presumed low grades of tissue fibrosis as well. Still, we specifically excluded subendocardial areas from the analysis, typically with high amount of fibrosis, which could be expected to reduce interobserver agreement, and frequently considered as nonrepresentative myocardial fields.

How the application of a deep learning network provided by *QuPath* impacts the precision of quantification in digital pathology is still a matter of debate. In our cohort of EMB we believe that predefined pixel counting and boundary detection using specific colour values is particularly valuable for intercellular space assessment and classification. This could translate in increased sensitivity for the quantification of interstitial fibrosis, meaning overall better quantification of EMB with higher levels of tissue heterogeneity. As this is frequently dependent on this type of fibrosis in the clinical context of severe AS [1,10] , one might deduce two additional explanations for our findings: 1) MF is underrated by the pathologist, when in comparison to *QuPath*, due to reduced sensitivity of human observation to interstitial fibrosis; 2) interobserver reproducibility is worse for cases with higher percentages of MF, possible owing to substantial increased degrees of interstitial fibrosis in these samples. Moreover, and

as depicted by our examples in figure 5, we verified that the automatic algorithm appropriately discriminated other tissue components, such as blank spaces, adipose tissue, vessels, red blood cells and even artifacts. These last findings, as well as perivascular fibrosis, might be particularly difficult to integrate in the pathologist quantification process. As stated, and contrary to automatic algorithms analysis, in which a predefined threshold for tissue components classification works for any and every tissue section, individual pathologist´s "human eye" assessment and "thresholds" must be individually determined again for each new section [14].

In clinical terms, the use of an open-source software like *QuPath*, with artificial intelligence development through deep learning techniques, may be of extended value in clinical correlation studies, further beyond cardiovascular pathology. Our analysis may be the first to specifically demonstrate how to improve MF quantification in EMB samples, reducing inter-observer variability, resorting to the use of *QuPath*, as previously recognized for other medical specialties such as cancer pathology. As *QuPath*`s developments and algorithms do support the *Digital Imaging and Communications in Medicine* (DICOM) standard format for WSI, one might extrapolate its use for the quantification of cardiac imaging derived (big) data, always keeping the demonstrated capacity to mine the image components, uncovering subtleties in a precise manner. The increased precision of measurements in both pathology and imaging studies may better assess distinct results across institutions and cohorts of patients, eventually improving the agreement of cardiac imaging/histology correlation studies. Ultimately this could mean more consensus in the definition of non-invasive imaging surrogate markers of MF.

LIMITATIONS

Contrary to some correlation studies we did not specifically assess the impact of the automatic algorithm quantification across distinct stains for collagen, such as Picrosirius red or Azan. This may be important when different institutional protocols are compared. Variations in staining intensity, which is related to specificity stain affinity for collagen fibers and further ECM components, staining hue, contrast, and tendency to fade over time may all impact the level of colour threshold definition for the algorithm and overall quantification.

Additionally, we did not evaluate the effect of pathologist´s fatigue on the quantification and interobserver variability and improvement after the knowledge of the automatic quantification. Timeframe of evaluation could be an issue for large volumes of samples and imaging data, as this was not our case. Even so, we might suppose that the consistency of the automatic quantification, irrespective of the number of evaluations, may improve the impact on precision for large volume of data, meaning increased automation. In this way, more comprehensive studies involving distinct patterns of MF, beyond valvular pathology, with large number of EMB samples, should be proposed for the application of *QuPath* methodology.

Finally, our study was not intended to measure the accuracy of the automatic method for the quantification of MF. This would need to be an agreement study, namely with stereology, aimed for the calibration of the automatic algorithm.

CONCLUSIONS (*graphical abstract*)

The use of an automatic digital pathology algorithm for the quantification of MF on EMB samples significantly improves the reproducibility of measurements by experienced pathologists. This was achieved through a machine learning imaging hierarchical protocol, developed on *QuPath,* an open-source software. As precision increases and automation is inferred, this quantification tool should be pursued for wide availability in the assessment of MF, a decisive biomarker on a huge range of cardiovascular conditions.

ACKNOWLEDGEMENTS: the authors would like to thank to the technician Fernanda Silva, from Pathology Department, IPOFG, Lisboa, for microscopy slides digitization, and to Fundação Champalimaud, Lisboa, for the opportunity to use computer vision for the first applications of *QuPath.*

SOURCES OF FUNDING

Nothing to declare.

DISCLOSURES

None.

REFERENCES

1. López B, Ravassa S, Moreno MU, José GS, Beaumont J, González A, Díez J. Diffuse myocardial fibrosis: mechanisms, diagnosis and therapeutic approaches. *Nat Rev Cardiol* [Internet]. 2021 [cited 2022 Sep 7];18:479–498. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33568808
2. Frangogiannis NG. Cardiac fibrosis. *Cardiovasc Res* [Internet]. 2021 [cited 2022 Sep 7];117:1450–1488. Available from: https://academic.oup.com/cardiovascres/article/117/6/1450/5949044
3. González A, Schelbert EB, Díez J, Butler J. Myocardial Interstitial Fibrosis in Heart Failure: Biological and Translational Perspectives. *J Am Coll Cardiol*. 2018;71:1696–1706.
4. Díez J, de Boer RA. Management of cardiac fibrosis is the largest unmet medical need in heart failure. *Cardiovasc Res* [Internet]. 2022 [cited 2022 Sep 7];118:e20–e22. Available from: https://academic.oup.com/cardiovascres/article/118/2/e20/6318397
5. Montesi SB, Désogère P, Fuchs BC, Caravan P. Molecular imaging of fibrosis: Recent advances and future directions. *Journal of Clinical Investigation*. 2019;129.
6. Daunoravicius D, Besusparis J, Zurauskas E, Laurinaviciene A, Bironaite D, Pankuweit S, Plancoulaine B, Herlin P, Bogomolovas J, Grabauskiene V, et al. Quantification of myocardial fibrosis by digital image analysis and interactive stereology. *Diagn Pathol* [Internet]. 2014 [cited 2022 Aug 31];9. Available from: https://pubmed.ncbi.nlm.nih.gov/24912374/
7. Chin CWL, Pawade TA, Newby DE, Dweck MR. Risk Stratification in Patients With Aortic Stenosis Using Novel Imaging Approaches. *Circ Cardiovasc Imaging* [Internet]. 2015 [cited 2022 Sep 25];8. Available from: https://pubmed.ncbi.nlm.nih.gov/26198161/
8. Park SJ, Cho SW, Kim SM, Ahn J, Carriere K, Jeong DS, Lee SC, Park SW, Choe YH, Park PW, et al. Assessment of Myocardial Fibrosis Using Multimodality Imaging in Severe Aortic Stenosis: Comparison With Histologic Fibrosis. *JACC Cardiovasc Imaging*. 2019;12:109–119.
9. Karamitsos TD, Arvanitaki A, Karvounis H, Neubauer S, Ferreira VM. Myocardial Tissue Characterization and Fibrosis by Imaging. *JACC Cardiovasc Imaging* [Internet]. 2020 [cited 2022 Sep 25];13:1221–1234. Available from: https://pubmed.ncbi.nlm.nih.gov/31542534/
10. Treibel TA, López B, González A, Menacho K, Schofield RS, Ravassa S, Fontana M, White SK, Disalvo C, Roberts N, et al. Reappraising myocardial fibrosis in severe aortic stenosis: An invasive and non-invasive study in 133 patients. *Eur Heart J*. 2018;39:699–709.
11. Laurinavicius A, Laurinaviciene A, Dasevicius D, Elie N, Plancoulaine B, Bor C, Herlin P. Digital image analysis in pathology: benefits and obligation. *Anal Cell Pathol (Amst)* [Internet]. 2012 [cited 2022 Sep 4];35:75–78. Available from: https://pubmed.ncbi.nlm.nih.gov/21971321/
12. Vahanian A, Beyersdorf F, Praz F, Milojevic M, Baldus S, Bauersachs J, Capodanno D, Conradi L, de Bonis M, de Paulis R, et al. 2021 ESC/EACTS Guidelines for the management of valvular heart disease. *Eur Heart J* [Internet]. 2022;43:561–632. Available from: https://academic.oup.com/eurheartj/article/43/7/561/6358470

13. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, McQuaid S, Gray RT, Murray LJ, Coleman HG, et al. QuPath: Open source software for digital pathology image analysis. *Sci Rep* [Internet]. 2017 [cited 2022 Nov 16];7. Available from: https://pubmed.ncbi.nlm.nih.gov/29203879/
14. Humphries MP, Maxwell P, Salto-Tellez M. QuPath: The global impact of an open source digital pathology system. *Comput Struct Biotechnol J* [Internet]. 2021 [cited 2022 Oct 11];19:852–859. Available from: https://pubmed.ncbi.nlm.nih.gov/33598100/
15. Eckstein J, Renner A, Zittermann A, Fink T, Sohns C, Niehaus K, Bednarz H, Neumann JM, Piran M, Kellner U, et al. Impact of left atrial appendage fibrosis on atrial fibrillation in patients following coronary bypass surgery. *Clin Cardiol* [Internet]. 2022 [cited 2022 Oct 11];Available from: https://pubmed.ncbi.nlm.nih.gov/35864729/
16. Krouwer JS. Why Bland-Altman plots should use X, not (Y+X)/2 when X is a reference method. *Stat Med* [Internet]. 2008 [cited 2022 Oct 12];27:778–780. Available from: https://pubmed.ncbi.nlm.nih.gov/17907247/
17. Christensen HS, Borgbjerg J, Børty L, Bøgsted M. On Jones et al.'s method for extending Bland-Altman plots to limits of agreement with the mean for multiple observers. *BMC Med Res Methodol* [Internet]. 2020 [cited 2022 Oct 12];20. Available from: https://pubmed.ncbi.nlm.nih.gov/33308154/
18. Mühlfeld C, Nyengaard JR, Mayhew TM. A review of state-of-the-art stereology for better quantitative 3D morphology in cardiac research. *Cardiovasc Pathol* [Internet]. 2010 [cited 2022 Nov 17];19:65–82. Available from: https://pubmed.ncbi.nlm.nih.gov/19144544/
19. Frangogiannis NG. Cardiac fibrosis. *Cardiovasc Res* [Internet]. 2021 [cited 2022 Jun 28];117:1450–1488. Available from: https://pubmed.ncbi.nlm.nih.gov/33135058/
20. Vasiljević JD, Popović ZB, Otašević P, Popović Z v., Vidaković R, Mirić M, Nešković AN. Myocardial fibrosis assessment by semiquantitative, point-counting and computer-based methods in patients with heart muscle disease: a comparative study. *Histopathology* [Internet]. 2001 [cited 2022 Nov 16];38:338–343. Available from: https://pubmed.ncbi.nlm.nih.gov/11318899/
21. Schipke J, Brandenberger C, Rajces A, Manninger M, Alogna A, Post H, Mühlfeld C. Assessment of cardiac fibrosis: a morphometric method comparison for collagen quantification. *J Appl Physiol (1985)* [Internet]. 2017 [cited 2022 Nov 17];122:1019–1030. Available from: https://pubmed.ncbi.nlm.nih.gov/28126909/
22. Hadi AM, Mouchaers KTB, Schalij I, Grunberg K, Meijer GA, Vonk-Noordegraaf A, van der Laarse WJ, Beliën JAM. Rapid quantification of myocardial fibrosis: a new macro-based automated analysis. *Cell Oncol (Dordr)* [Internet]. 2011 [cited 2022 Nov 22];34:343–354. Available from: https://pubmed.ncbi.nlm.nih.gov/21538025/

NOVELTY AND SIGNIFICANCE

"What is known?"

- Myocardial fibrosis is a common pathological process, with prognostic implication, in a wide range of chronic heart diseases.
- The methods for its quantification are diverse, both in imaging and invasive, histology-based, assessments.
- The clinical impact of this biomarker is difficult to compare across correlation studies with distinct methodologies.

"What new information does this article contribute?"

- The use of an Artificial Intelligence algorithm from an open-source platform for the quantification of myocardial fibrosis on digital images improves the human/pathologist´s reproducibility of measurements.
- Wide application of Artificial Intelligence tools may positively impact the reliability of correlation studies for cardiac biomarkers involving large number of patients.

TABLES

Table 1 – *Demographic and clinical characteristics of the population.*

Number of patients N = 30	
CLINICAL CHARACTERISTICS	
Age, years	73 (69-78)
Male	13 (43.3%)
BSA m^2	1.80 ± 0.18
Hypertension	26 (86.6%)
Diabetes mellitus	7 (23.0%)
Dyslipidemia	23 (76.6%)
Smoking history	4 (13.3%)
Previous stroke	1 (0.03%)
NYHA functional class	
I	2 (6.7%)
II	18 (60%)
III	10 (33.3%)
Anginal symptoms	8 (26.7%)
Syncope	5 (16.7%)
LABORATORY RESULTS	
Creatinine, mg/dL	0.89 (0.70-1.12)
Estimated glomerular filtration rate, mL/min	69.5 (52.0-88.1)
Troponin, ng/L	13 (10-22)
IMAGING DATA	
Aortic valve mean gradient, mmHg (TTE)	66.5 ± 21.7
Aortic valve area, cm^2 (TTE)	0.68 ± 0.18
Left Ventricular Ejection Fraction, % (CMR)	58.4 ± 8.0
Left Ventricular Indexed Mass, g/m^2 (CMR)	85.5 (63-119)
Delayed enhancement (number of patients)	20 (66.7%)
Delayed enhancement mass, %	3.0 (2.7-5.9)
Native T1, ms	1061.5 ± 44.5
Extracellular volume, %	24.5 (23.0-27.5)

607 Values are median (interquartile range), mean±standard deviation, BSA - body surface area, NYHA – New York Heart Association,
608 TTE – transthoracic echocardiography; CMR – cardiac magnetic resonance.
609

610 ***Table 2 –*** *Summary statistics for MF quantification (quantification in %).*

	FIRST QUANTIFICATION			***AUTOMATIC QUANTIFICATION QuPath***	***SECOND QUANTIFICATION***		
	OBS A	***OBS B***	***OBS C***		***OBS A***	***OBS B***	***OBS C***
Number of observations	30	30	30	30	30	30	30
Mean	19.5	5.5	7.9	14.9	11.2	8.3	8.6
Median	15.0	2.5	7.5	13.6	10	7.5	7.5
Standard deviation	15.1	4.8	9.5	9.5	6.8	7.6	9.1
Range	10.0	5.0	5.0	13.7	10.0	5.0	5.0
Minimum	5.0	2.5	7.5	4.0	5.0	2.5	2.5
Maximum	80.0	27.5	52.5	51.2	40.0	42.5	50.0

611 *OBS* – observer (pathologist)
612

613 ***Table 3 –*** *Bland-Altman for the agreement of MF quantification before and after the knowledge*
614 *of automatic algorithm quantification.*

BLAND-ALTMAN ANALYSIS							
		Inter-observer Mean difference	**95% CI of the mean difference**	**Interobserver 95% CI range**	**One-sample T-test P-value**	**LINEAR REGRESSION**	
						Beta coefficient	**P-value**
AGREEMENT BEFORE AUTOMATIC QUANTIFICATION	*Obs A vs. B*	14.00	9.64; 18.36	-8.89; 36.89	<0.001	0.93	<0.001
	Obs A vs. C	11.58	8.57; 14.60	-4.25; 27.41	<0.001	0.73	<0.001
	Obs B vs. C	-2.42	-4.76; -0.07	-14.72; 9.88	0.044	-0.79	<0.001
	Obs A, B & C	0	-1.67; 1.67	-15.68; 15.68	NS (1.0)	0	NS (1.0)
AGREEMENT AFTER AUTOMATIC QUANTIFICATION	*Obs A vs. B*	2.83	1.00; 4.66	-6.77; 12.43	0.004	-1.19	NS (0.32)
	Obs A vs. C	2.58	0.70; 4.46	-7.26; 12.42	0.009	-0.49	0.006
	Obs B vs. C	-0.25	-1.92; 1.42	-9.01; 8.51	NS (0.76)	-0.35	NS (0.10)
	Obs A, B & C	0	-0.21; 0.21	-5.94; 5.94	NS (1.0)	0	NS (1.0)

615 *Obs* – observer (pathologist); CI – confidence interval.
616

617 ***Table 4 –*** *Reliability analysis for the agreement of MF quantification before and after the*
618 *knowledge of the automatic algorithm quantification. As depicted, there is an increased*
619 *agreement for the three observers A, B & C at both individual and average measurement*
620 *analysis (highlighted) before and after the knowledge of the automatic algorithm*

quantification. Individual measurement analysis is describing the agreement between each of the measurements, independent from the observer, and the mean classification from the three pathologists. This provides the evaluation per each measurement instead as per observer.

RELIABILITY ANALYSIS					
		Individual measurement analysis		**Average measurement analysis**	
		Intraclass correlation coefficient	**P-value**	**Intraclass correlation coefficient**	**P-value**
AGREEMENT BEFORE AUTOMATIC QUANTIFICATION	*Obs A vs. B*	0.274	0.003	0.430	0.003
	Obs A vs. C	0.573	<0.001	0.728	<0.001
	Obs B vs. C	0.636	<0.001	0.777	<0.001
	Obs A, B & C	0.737	<0.001	0.849	<0.001
AGREEMENT AFTER AUTOMATIC QUANTIFICATION	*Obs A vs. B*	0.729	<0.001	0.843	<0.001
	Obs A vs. C	0.778	<0.001	0.875	<0.001
	Obs B vs. C	0.867	<0.001	0.929	<0.001
	Obs A, B & C	0.925	<0.001	0.961	<0.001

FIGURES WITH FIGURE LEGENDS

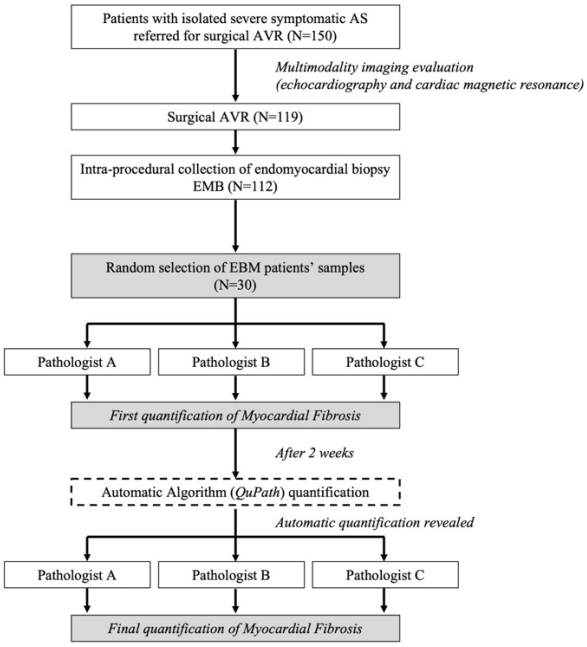


Figure 1. *Study flow-chart.*

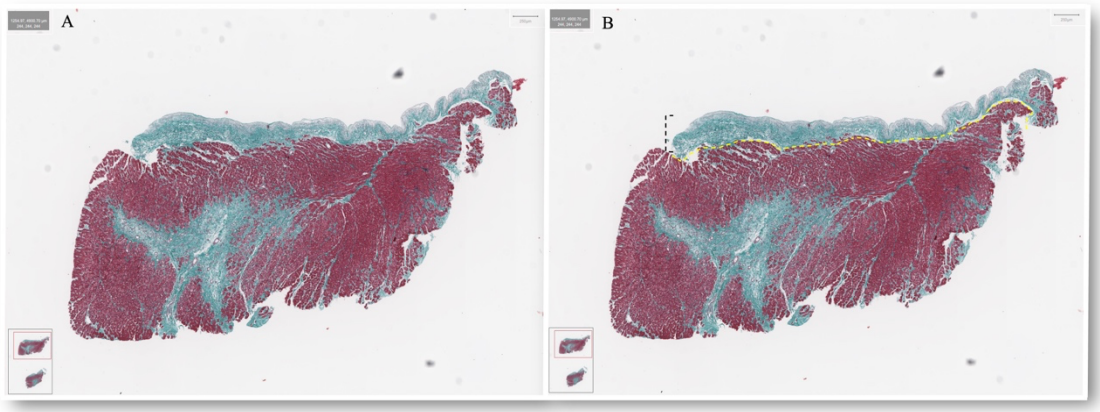


Figure 2. *A) Whole slide image (WSI), as obtained from digital microscopic scanner, also called digital slide system or virtual microscope, used to convert ordinary glass slides into digital slides; B) Specific image selection with the exclusion of dense endocardial fibrosis.*

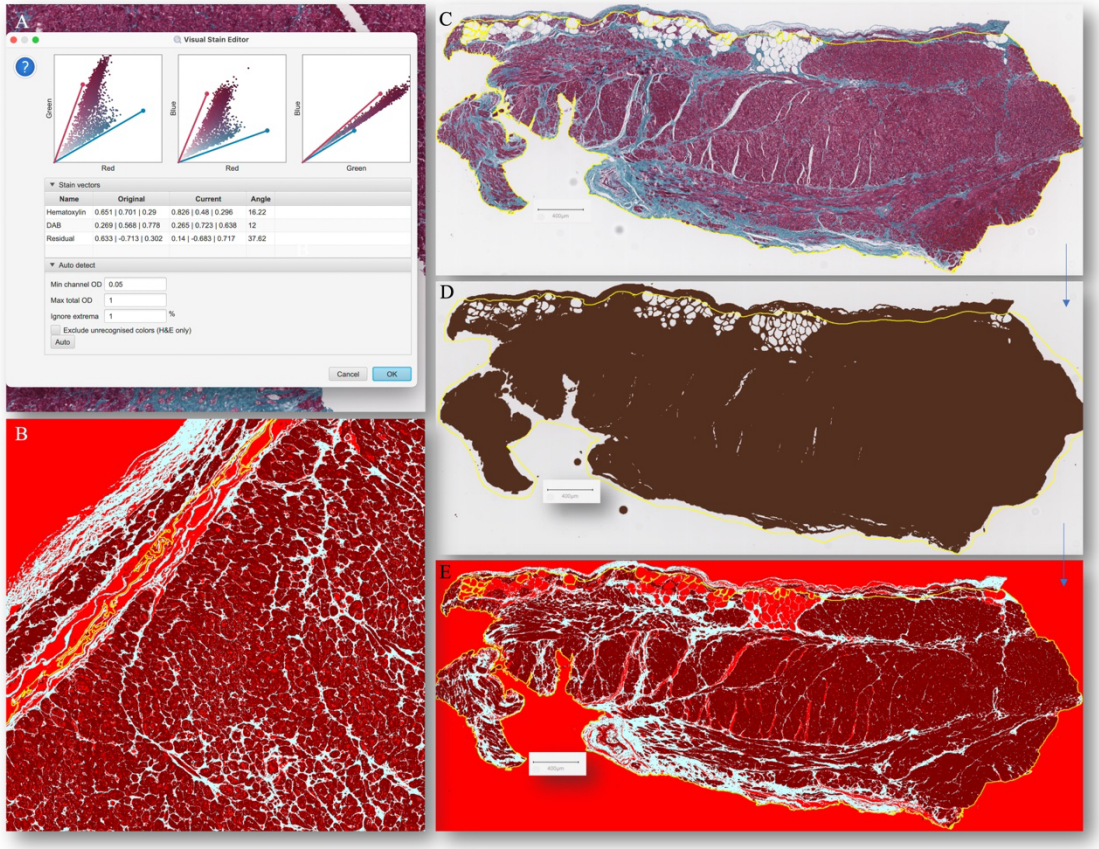


Figure 3. *Example of an automatic quantification, including previous definition of appropriate colour stain vectors – A; complete classification of tissue compartments is displayed from C through E; B – detailed view with specific colour coded tissue components – white is fibrosis.*

A

Key	Value
Image	1020759.svs
Name	Tissue
Class	Tissue
Parent	Image
ROI	Geometry
Centroid X µm	3492.1918
Centroid Y µm	7901.4082
MONSTA: Cardiomyocytes %	76.9298
MONSTA: Cardiomyocytes are...	5191197
MONSTA: Fibrosis %	18.6879
MONSTA: Fibrosis area µm^2	1261056.625
MONSTA: Void %	4.3823
MONSTA: Void area µm^2	295716.4688
Area µm^2	6748213.0236
Perimeter µm	24658.3233

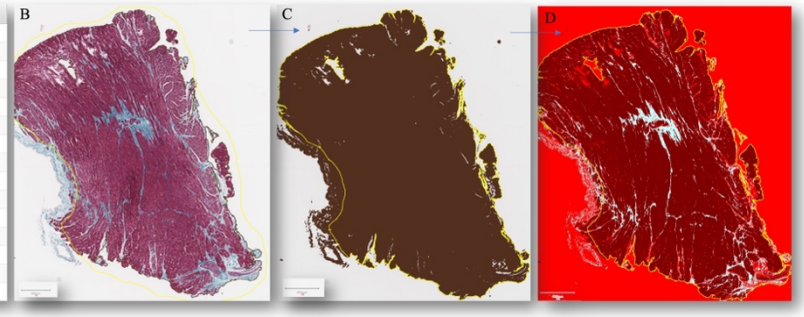


Figure 4. *Example of an automatic quantification and table (A), displaying absolute areas and proportions of each tissue components.*

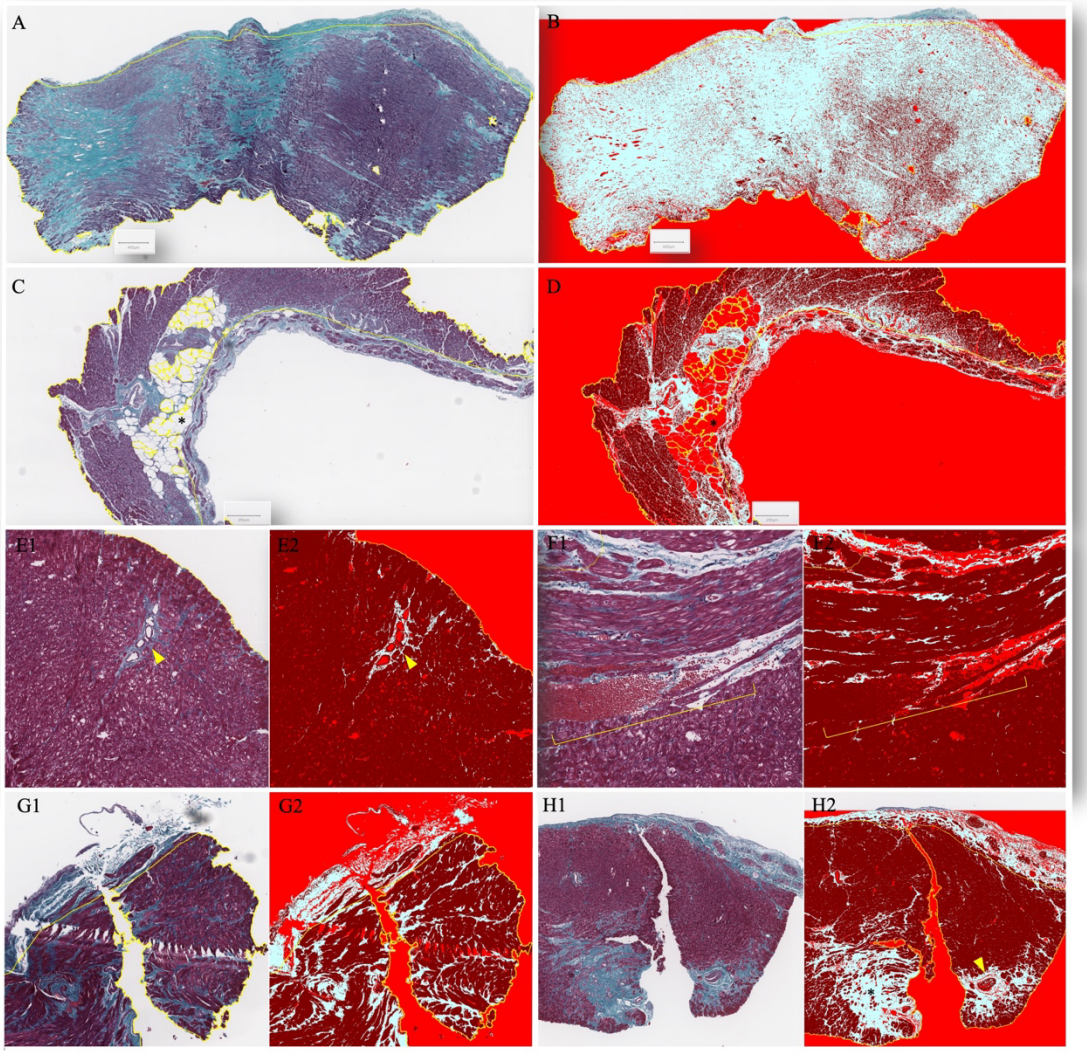


Figure 5. *A, B – case with an extensive diffuse fibrosis and corresponding algorithm classification; C, D – example of exclusion of adipose tissue (*) from the analysis, as appropriately classified as an empty space. E1, E2 – example with identification of*

intramyocardial vessels, corresponding lumina, and perivascular fibrosis (arrowhead). F1, F2 – single case example with identification of a big vessel with red blood cells inside (marker); in this case fibrotic component is correctly identified although red blood cells are matched with cardiomyocytes for the same coded colour. G1, G2 – example of an appropriately identified artifact by the algorithm. H1, H2 – example of a challenging case with a big empty space, extensive fibrosis with scar (*) and perivascular fibrosis (arrowhead).

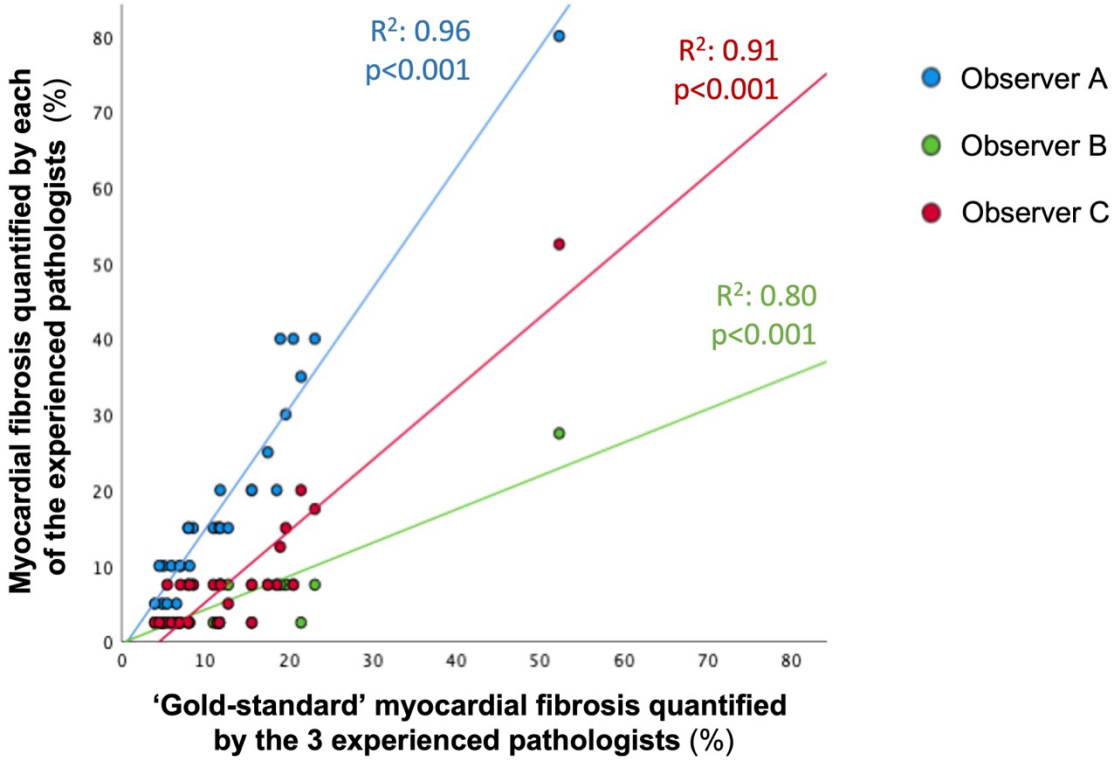
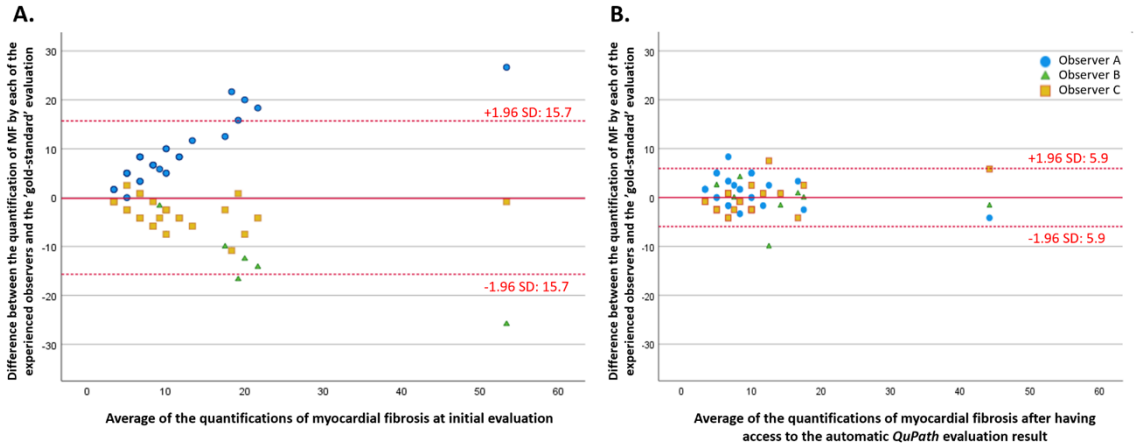


Figure 6. *Correlation plot assessing the individual quantification of each of the pathologist versus the mean "gold-standard" quantification by the three pathologists.*



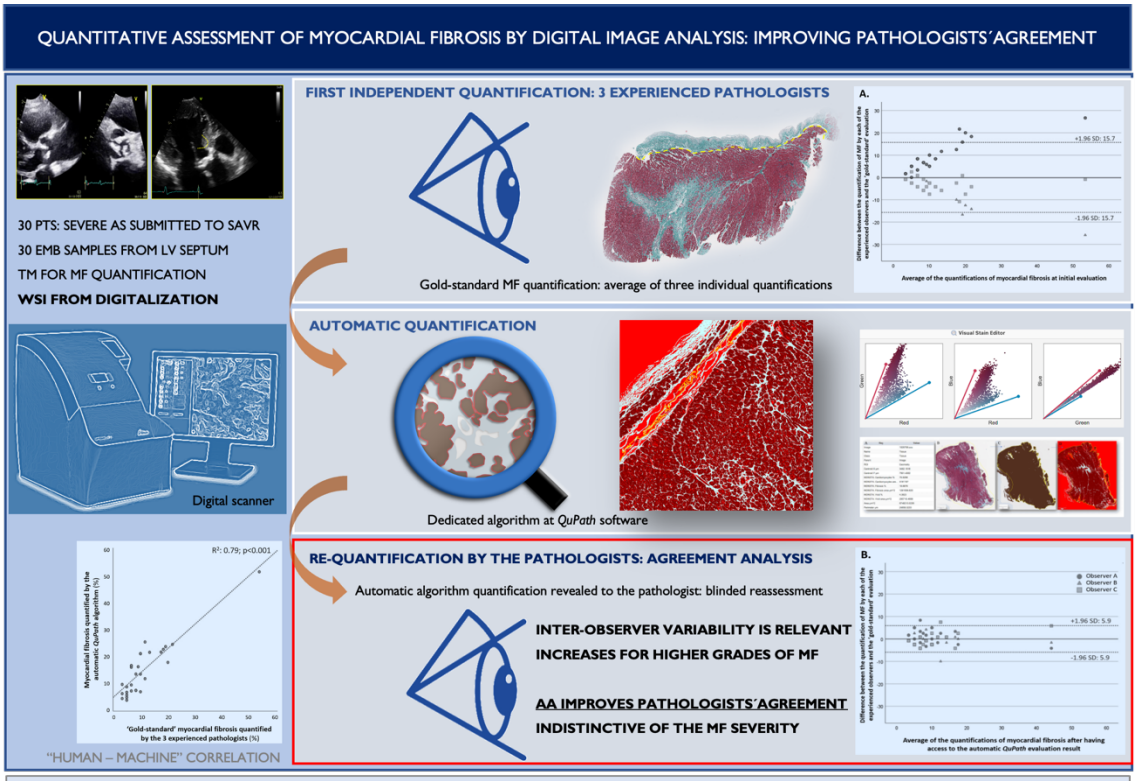
654

655 ***Figure 7.*** *Bland-Altman plot of agreement between the quantification of myocardial fibrosis*
656 *between the 3 experienced observers before (A.) and after (B.) knowing the result of the*
657 *automatic QuPath evaluation. The difference between each observer and the overall mean for*
658 *the specific observation was estimated and displayed before and after the knowledge of the*
659 *algorithm automatic quantification.*

660

661

662



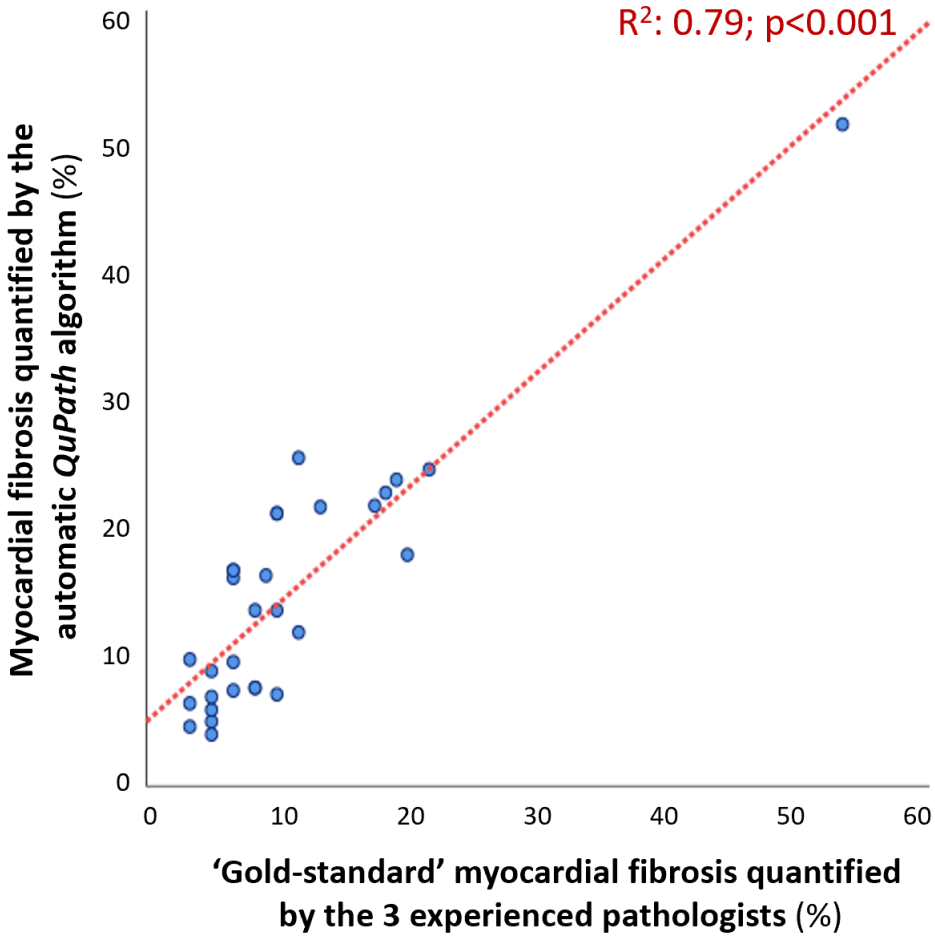
PTS – patients; AS – aortic stenosis; SAVR – surgical aortic valve replacement; EMB – endomyocardial biopsy; LV – left ventricular; MF - myocardial fibrosis; TM – *Masson´s Trichrome* stain; AA – automatic algorithm

Graphical abstract.

677 **SUPPLEMENTARY DATA**

678 ***Supplemental Figure S1.*** *Correlation plot for the overall quantification of the pathologists*

679 *versus automatic quantification from the QuPath.*



680