Image Anal Stereol 2024;43: 131-137 doi: 105566/ias.3171
Original Research Paper

COMPARISON OF THE VISUAL SCORING METHOD AND SEMI-AUTOMATIC IMAGE ANALYSIS FOR EVALUATING STAINING INTENSITY OF HUMAN CARTILAGE SECTIONS

ARMIN ALIBEGOVIĆ[1], NEJC UMEK[✉,2,#], LUKA PUŠNIK[2], INIGO ZUBIAVRRE MARTINEZ[3,#]

[1]Institute of Forensic Medicine, Faculty of Medicine, University of Ljubljana, Korytkova 2, 1000 Ljubljana, Slovenia; [2]Institute of Anatomy, Faculty of Medicine, University of Ljubljana, Korytkova 2, 1000 Ljubljana, Slovenia; [3]Institute of Clinical Medicine, University of Tromsø, 9037 Tromsø, Norway.

[#]The authors contributed equally
e-mail: armin.alibegovic@mf.uni-lj.si, nejc.umek@mf.uni-lj.si, luka.pusnik@mf.uni-lj.si, inigo.martinez@uit.no

(Received February 26, 2024; revised April 4, 2024; accepted April 17, 2024)

ABSTRACT

Accurate estimation of postmortem interval (PMI) is crucial in forensic medicine. The hyaline cartilage, being predominantly composed of a dense extracellular matrix and partly resistant to factors influencing protein degradation, can be utilized for analyzing PMI intervals. Various staining methods are available for cartilage staining for PMI evaluation; however, the conventional visual scoring method for assessing staining intensity is susceptible to evaluator bias. This study compared the visual scoring method with a modified Bern score with semi-automatic image analysis. The cartilage samples were obtained from human cadavers with known time of death. Forty-five histological slices were prepared and stained using Alcian blue, Safranin-O with Fast green, Safranin-O without Fast green, Masson trichrome, and Sirius red. Ten evaluators visually scored each sample on a scale of 0 to 3. A semi-automatic analysis was conducted on the same images using the deconvolution plugin of the ImageJ software by three independent evaluators. Linear regression was used to assess the correlation between the mean grey value and the mean Bern score from all evaluators. The results showed strong correlations across all evaluated staining techniques ($r \geq 0.77$, $p < 0.0001$), with Masson trichrome staining exhibiting the highest correlation. The intra-class correlation coefficients between the independent semi-automatic assessments were excellent for all five stainings ($ICC \geq 0.965$). Accordingly, semi-automatic image analysis can be a suitable replacement for the visual scoring method, particularly when no procedural artifacts are present.

Keywords: Bern score, cartilage, histology, postmortem interval, semi-automatic analysis, visual scoring method.

INTRODUCTION

Precise determination of the time of death is imperative in forensic medicine (Dell'Aquila *et al.*, 2021). Various traditional techniques, such as body temperature measurement, electrical or chemical supravital stimulation of skeletal muscles, or rigor and livor mortis assessment, are commonly used for estimating the postmortem interval (PMI). However, the reliability of these methods, which are based on early physical postmortem changes, decreases as the PMI extends (Hayman and Oxenham, 2016; Shrestha *et al.*, 2024). Biochemical or histological changes can be utilized to assess the PMI, especially within isolated compartments less affected by putrefaction. Analytic techniques based on such indices have emerged as promising tools, offering the potential for a more objective and reliable estimation of the time of death (Wei *et al.*, 2020; Tomsia *et al.*, 2022; Madea *et al.*, 2001; Pigaiani *et al.*, 2020).

Several biochemical techniques that analyze protein degradation have been suggested for assessing PMI. Proteins undergo specific changes postmortem, some of which follow a regular and predictable pattern (Sacco *et al.*, 2022; Pittner *et al.*, 2022; Zissler *et al.*, 2020). Many studies have investigated protein degradation in various tissues with different techniques, such as Western blotting, liquid chromatography-mass spectrometry, and immunohistochemistry (Zhang *et al.*, 2020; Choi *et al.*, 2019; Alibegović *et al.* 2020). While the significance of protein degradation in PMI assessment has been extensively emphasized, practical challenges arise due to varied methodologies, results obtained from both animal and human specimens, and the examination of different

tissues exposed to diverse factors influencing their integrity (Ehrenfellner *et al.*, 2017; Choi *et al.*, 2019; Zissler *et al.*, 2020).

Hyaline cartilage is considered ideal in PMI estimation for its stability and slower decomposition rate due to absence of vasculature (Chang *et al.*, 2024). As a diffusion-dependent isolated compartment, the tissue exhibits partial resistance to external factors impacting protein degradation (Tomsia *et al.*, 2022). With a comparatively low cell density in contrast to the dense extracellular matrix, it proves particularly valuable for late postmortem analysis (Alibegović, 2014). Nevertheless, while cartilaginous tissue is utilized for age estimation, its potential in forensic medicine remains underexplored. Various protein staining methods have been demonstrated for PMI detection in cartilage tissue (Alibegović 2014; Alibegović *et al.* 2020; Chang *et al.*,2024; Rogers *et al.*, 2011); however, the main drawbacks when assessing the staining intensity are the subjectivity of the evaluator and the categorical nature of the results. To categorize evaluator estimates, a semi-quantitative grading scale can be employed for the assessment of intensity, with one such being the modified Bern score—a four-grading scale employed for the assessment of the cartilage staining intensity (Power *et al.*, 2021). Accordingly, this study aimed to compare the visual scoring method to a semi-automatic approach in the assessment of PMI with stained cartilage sections, using a modified Bern score.

MATERIAL AND METHODS

Ethical Approval

The study was approved by the National Medical Ethics Committee of the Republic of Slovenia (Permit numbers: 25/09/08 and 30/06/15). The study adhered to national confidentiality standards to safeguard sensitive patient health information and was conducted in accordance with the principles outlined in the Helsinki Declaration. The part of this study in which different examiners assessed the cartilage sections with the visual Bern score and correlated the results to PMI intervals has already been published (Alibegović *et al.*, 2020).

Sample Harvesting

The hyaline cartilage samples were harvested from three fresh cadavers of young deceased men, aged between 20 and 45, who had experienced sudden death in a traffic accident, with precise documentation of the time of death. Although the postmortem medical data was limited, none of the subjects had any recognized joint pathology. Promptly after the confirmation of death, the body was transported to a morgue with an environmental temperature set at 4 ± 2 °C. Cartilage harvesting from the knee was performed during the autopsy for all cadavers within the initial 24 hours postmortem. Using an arthroscopic mosaicplasty instrument (Helipro, Lesce, Slovenia), three osteochondral cylinders were obtained in aseptic conditions from the femoral condyle of each cadaver and promptly placed in a DMEM/F12 medium (Capricorn Scientific GmbH, Ebsdorfergrund, Germany) containing vancomycin, gentamicin, and amphotericin B (all procured from Gibco, Paisley, UK). The samples were sealed without media substitution, mimicking a postmortem environment. The osteochondral cylinders were then stored in a cooling room at a temperature of 11 ± 2 °C or temperature of 35 ± 2 °C until further analysis at different time points. At 2, 12, and 36 days postmortem, one osteochondral cylinder was cut with a vibratome and put into the 10 % solution of neutral buffered formaldehyde (Sigma-Aldrich, Steinheim, Germany) for approximately 24 hours. A more detailed description of the sample harvesting protocols can be found in the previously published article (Alibegović *et al.*, 2020).

Preparation and Staining of Histological Samples

Immediately following formaldehyde fixation, the samples were washed in increasing concentrations of ethanol solutions and subsequently embedded in paraffin blocks. The samples were oriented in the split plane, ensuring the presence of all layers (superficial, middle, and deep) in each histological section. Histological slices, with a thickness of 4-5 μm, were prepared using a microtome, with 2-4 slices obtained from each paraffin block for each staining procedure.

Five different histological staining techniques were employed. Alcian blue and Safranin-O stains were used for acidic polysaccharides (glycosaminoglycans), with Safranin-O partly combined with Fast green (with or without Fast green). Masson's trichrome stain and Sirius red were utilized for collagen fibers. The Alcian blue and Masson's trichrome stains were automated, while Safranin-O (with or without Fast green) and Sirius red staining were performed manually (Figure 1). The detailed staining procedures are described in a previously published article (Alibegović *et al.,* 2020).

Manual Assessment of the Samples

The samples were evaluated with the light microscope (Nikon Eclipse 80i, Nikon, Tokyo, Japan) under ×100 magnification. Ten evaluators, professor of biosciences, pathology and forensic medicine specialists and residents, and laboratory technicians, assessed samples

using a modified Bern score, a well-established method for grading cartilage staining. The staining intensity receives a score between 0 and 3, where 0 describes an absence of staining or extremely pale staining, while number 1 represents weak staining, 2 is moderate staining, and 3 is intensive sample staining. Prior to each evaluation, the assessors were provided with reference samples demonstrating the different staining intensities.

Semi-Automatic Image Analysis

The slides for digital image analysis were visualized using a light microscope (Axio Zoom.V16, Zeiss, Oberkochen, Germany) under ×100 magnification. The scanning was performed with a digital camera (AxioCam MRm, Zeiss, Oberkochen, Germany) and ZEN imaging software platform (AxioVision, Zeiss, Jena, Germany), with 16-bit images of whole sections acquired at a resolution of 1388 × 1040. All similarly stained sections were captured with identical image settings. The digital intensity of staining analysis was performed using the ImageJ software (National Institutes of Health, Bethesda, Maryland, United States) and image processing packages Fiji (Schindelin *et al.*, 2012). Using the deconvolution plugin, the blue, red, and green channels were split (Ruifrok and Johnston, 2001). The blue channel was used to quantify the intensity of Alcian blue and Masson's trichrome staining, while the red channel was employed to quantify the intensity of Safranin-O with Fast green, Safranin-O without Fast green, and Sirius red staining. All color channels were converted into greyscale 8-bit images. Staining intensity was assessed by measuring opacity, which was determined as the average mean grey intensity along five transverse lines across the sample, avoiding folded or destroyed parts in the captured slides of stained cartilage (Fig. 1). Background intensity was calculated as the average mean grey values of four different areas in the image not occupied by the cartilage section. The manual assessments were performed by three independent, single-blinded evaluators.

Statistical Analysis

The data are presented as means with 95 % confidence intervals (CI) unless otherwise stated. Linear regression was utilized to determine the correlation between the mean grey value measured using digital image analysis and the mean Bern score from all evaluators for each assessed slide. Statistical analysis and graphing were performed with GraphPad Prism 10 (GraphPad Software; LLC, San Diego, USA). Intraclass correlation coefficients (ICC) were calculated with absolute values using SPSS software (IBM SPSS Statistics, Chicago, USA). Statistical significance was considered at a P-value below 0.05.

RESULTS

In total, 45 histological slices stained with the different techniques were evaluated and compared (Fig. 1). Strong positive correlations ($r \geq 0.77$) were noted between the intensity staining evaluated by digital image analysis and visual scoring assessment (Fig. 2). The coefficient of determination was highest for Masson trichrome staining, and slightly lower, but comparable for Alcian blue, Safranin-O with or without Fast green, and Sirius red staining. The subtraction of background intensity did not significantly improve the magnitude of correlation (Table 1). The ICC for semi-automatic assessment was excellent for all five histological stainings (Table 2).

Table 1. *Linear regression of visual scoring method and semi-automatic image analysis: the effect of background.*

Staining	Slope equation	Slope (95% CI)	r	R squared
Alcian blue BG	32.95x + 164.40	32.95 (25.89-40.00)	0.82	0.67
Alcian blue sBG	31.53x + 102.60	31.53 (23.94-39.12)	0.78	0.62
Masson trichrome BG	40.56x + 70.25	40.56 (34.86-46.26)	0.89	0.79
Masson trichrome sBG	38.86x + 14.01	38.86 (33.75-43.96)	0.90	0.81
Safranin-O + Fast Green BG	30.91x + 157.50	30.91 (24.60-37.22)	0.81	0.66
Safranin-O + Fast Green sBG	29.13x + 143.50	29.13 (22.17-36.09)	0.77	0.59
Safranin-O BG	33.80x + 146.60	33.80 (27.39-40.20)	0.84	0.71
Safranin-O sBG	33.73x + 130.20	33.73 (27.08-40.38)	0.83	0.69
Sirius red BG	55.08x + 51.17	55.08 (44.08-66.08)	0.82	0.68
Sirius red sBG	49.50x + 36.60	49.50 (38.68-60.32)	0.80	0.64

BG – background; sBG – subtracted background; CI – confidence interval. Comparison of the mean visual scores with mean semi-automatic scores. The differences in slopes with and without the background are statistically non-significant.

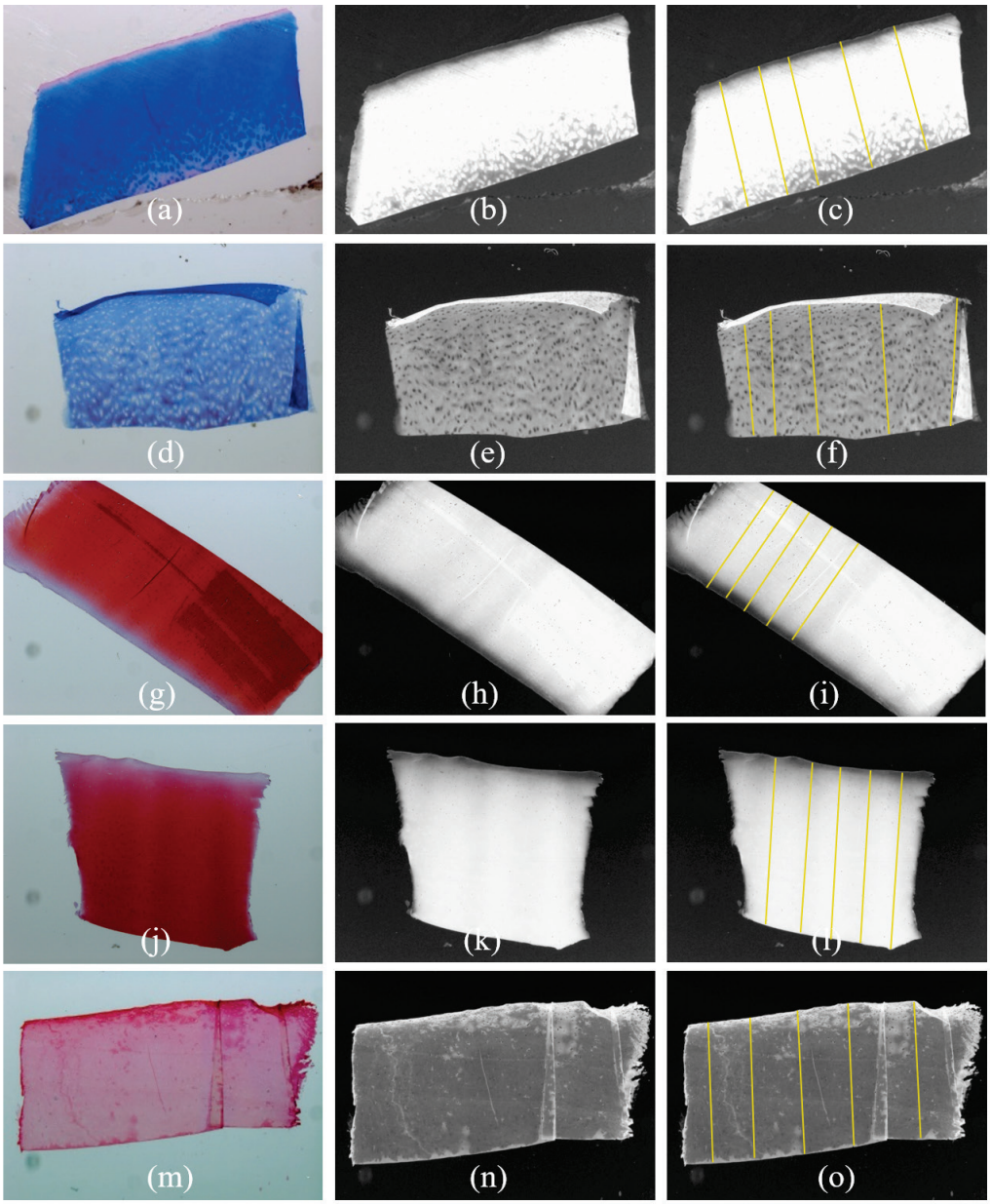

Fig.1. *Comparison of different stains with measuring the opacity. The histological stained slices and their corresponding images with mean grey values are shown for* ***(a-c)*** *Alcian blue,* ***(d-f)*** *Masson trichrome,* ***(g-i)*** *Safranin-O with Fast green,* ***(j-l)*** *Safranin-O without Fast green, and* ***(m-o)*** *Sirius red. The last column depicts images with grey values, each containing five transverse lines across the sample that are avoiding folded parts or artefacts.*

Table 2. *Intra-class correlation between independent raters for semi-automatic image analysis.*

Staining	ICC (95% CI)
Alcian blue	0.965 (0.943-0.980)
Masson trichrome	0.979 (0.966-0.988)
Safranin-O + Fast Green	0.984 (0.974-0.990)
Safranin-O	0.967 (0.947-0.980)
Sirius red	0.978 (0.970-0.983)

CI – confidence interval; ICC – intra-class correlation coefficient. The comparison of absolute values with background not being subtracted.

DISCUSSION

In this study, we compared the visual scoring of staining intensities of cartilage histological sections with semi-automatic image analysis. The results demonstrate strong correlations between both approaches for all the assessed staining techniques, with the highest correlation noted for Masson's trichrome staining. The semi-automatic method, even without background subtraction, appears to be a suitable substitute for the visual scoring method.

The cartilage primarily comprises collagen and proteoglycans, and postmortem degradation impacts the

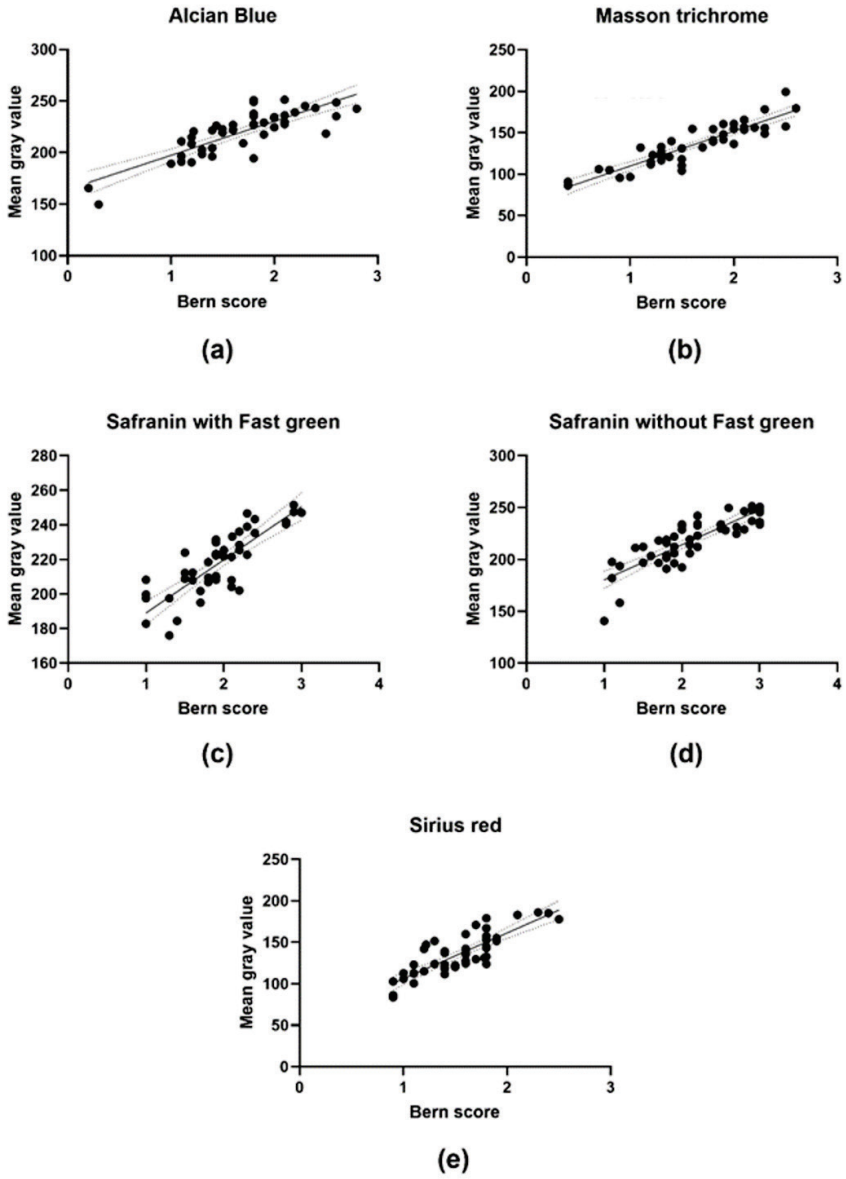

Figure 2. *Correlation between visual score and semi-automatic image analysis with background for* ***(a)*** *Alcian blue,* ***(b)*** *Masson trichrome,* ***(c)*** *Safranin-O with Fast green,* ***(d)*** *Safranin-O without Fast green, and* ***(e)*** *Sirius red staining. The figure depicts the correlation with 95 % confidence intervals between the mean grey value measured by digital image analysis (y-axis) and the mean Bern score from different evaluators using the visual scoring method (x-axis). The highest correlation coefficient was detected for Masson's trichrome staining. All the assessed correlations were statistically significant ($p < 0.0001$).*

histochemical staining intensity of these macromolecules. Because of its robust and resistant properties, it proves to be a reliable tissue for PMI detection (Alibegović, 2014; Chang *et al*., 2024). Considering the various staining protocols for cartilage, it is crucial to use identical methods when assessing PMI, particularly for meaningful result comparisons (Alibegović *et al*., 2020). It is imperative that time frames are considered because some stains, such as Alcian blue, cause overstaining with prolonged exposure and consequently impact the validity of the results (Rigueur and Lyons, 2014). This is especially important in automatic or digital analysis. In manual analysis, the experienced evaluator can partially compensate for such differences based on previous experience and less experienced evaluators can lack this ability. The same goes for the semi-automatic method, where the evaluator's input is crucial to processing the data to some extent. It is imperative to recognize the learning curve that probably exists regardless of the semi-automatic or manual approach.

The high-reliability between different evaluators indicate that the semi-automatic method could potentially require only one evaluator. Therefore, it is less time-consuming compared to the visual scoring method, which

involves multiple evaluators. Conversely, the subjectivity of raters in manual assessment can lead to poorer agreement between the raters. In the manual scoring method, only poor, fair, or moderate agreement has been noticed between the raters (Alibegović *et al.*, 2020). In addition, the semi-automatic method yields a numerical result on a larger scale, enhancing objectivity (Rizzardi *et al.*, 2012). Conversely, a visual scoring method gives a categorical value on the scale and thus does not allow intermediate choices.

The comparison of image analysis with or without the background subtraction do not suggest the necessity of subtracting the background. Furthermore, if the acquisition settings are the same for all the analyzed images and there are no impurities in the optical axis that would create artifacts, this step could be less important (Zupančič *et al.*, 2022; Zupančič *et al*., 2023). Moreover, the newer software can automatically address such issues with bright field corrections. Notably, tissue folding on a slice could impact the semi-automatic assessment; hence, visual inspection and evaluation should be preferred in such cases. In the future, applying digital image analysis aided by artificial intelligence or neural networks could help overcome the current challenge of recognizing the artifacts. Similar analysis with semi-automatic evaluation of staining intensities could be employed for other organ assessment as well (Alabbasi *et al.*, 2022).

We acknowledge that this study had some limitations. First, for the comparison with the visual scoring method, the mean value of all evaluators was employed, thus reflecting the group average, not the individual scores. Second, the results of the semi-automatic quantitative method are relative compared to the values obtained with the visual scoring method; however, this study aimed to assess the usefulness of the semi-automatic method and not to evaluate the referential values.

CONCLUSION

In conclusion, the semi-automatic method can be a suitable replacement for the visual scoring method, particularly when no procedural artifacts are present. The excellent agreement between the raters suggests that it may be feasible to rely on just one evaluator. Future studies with a larger sample size should establish reference intervals for detecting specific PMI.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the Slovenian Research and Innovation Agency (ARIS), Slovenia, (Grant No. P3-0043 and P2-0109).

The authors declare that have no competing interests in the submission of this manuscript.

We are grateful to Chiedozie K. Ugwoke for proofreading the manuscript.

REFERENCES

Alabbasi SF, Viramontes AC, Diaz FJ, Weedn VW (2022). Loss of Nuclear Basophilic Staining as a Postmortem Interval Marker. Am J Forensic Med Pathol 43:142–146.

Alibegović A (2014). Cartilage: A new parameter for the determination of the postmortem interval? J Forensic Leg Med 27:39–45.

Alibegović A, Blagus R, Martinez IZ (2020). Safranin O without fast green is the best staining method for testing the degradation of macromolecules in a cartilage extracellular matrix for the determination of the postmortem interval. Forensic Sci Med Pathol 16:252–258.

Chang LR, Marston G, Martin A (2024). Anatomy, Cartilage. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing.

Choi KM, Zissler A, Kim E, Ehrenfellner B, Cho E, Lee S, Steinbacher P, Yun KN, Shin JH, Kim JY, Stoiber W, Chung H, Monticelli FC, Kim JY, Pittner S (2019). Postmortem proteomics to discover biomarkers for forensic PMI estimation. Int J Legal Med 133:899–908.

Dell'Aquila M, De Matteis A, Scatena A, Costantino A, Camporeale MC, De Filippis A (2021). Estimation of the time of death: where we are now? Clin Ter 172:109–112.

Ehrenfellner B, Zissler A, Steinbacher P, Monticelli FC, Pittner S (2017). Are animal models predictive for human postmortem muscle protein degradation? Int J Legal Med 131:1615–1621.

Hayman J, Oxenham M (2016). Supravital Reactions in the Estimation of the Time Since Death (TSD). In: Human Body Decomposition. London: Elsevier, 1–12.

Madea B, Kreuser C, Banaschak S (2001). Postmortem biochemical examination of synovial fluid — a preliminary study. Forensic Sci Int 118:29–35.

Pigaiani N, Bertaso A, De Palo EF, Bortolotti F, Tagliaro F (2020). Vitreous humor endogenous compounds analysis for post-mortem forensic investigation. Forensic Sci Int 310:110235.

Pittner S, Merold V, Anders S, Lohner L, Amendt J, Klinger M, Hausmann R, Kissling S, Monticelli F, Geissenberger J, Zissler A, Steinbacher PA (2022). A standard protocol for the analysis of postmortem muscle protein degradation: process optimization and considerations for the application in forensic PMI estimation. Int J Legal Med 136:1913–1923.

Power L, Acevedo L, Yamashita R, Rubin D, Martin I, Barbero A (2021). Deep learning enables the

automation of grading histological tissue engineered cartilage images for quality control standardization. Osteoarthritis Cartilage 29:433–443.

Rigueur D, Lyons KM (2014). Whole-mount skeletal staining. Methods Mol Biol 1130:113-121.

Rizzardi AE, Johnson AT, Isaksson Vogel R, Pambuccian SE, Henriksen J, Pn Skubitz A, Metzger GJ, Schmechel SC (2012). Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. Diagn Pathol 7:42.

Rogers CJ, Clark K, Hodson BJ, Whitehead MP, Sutton R, Schmerer WM (2011). Postmortem degradation of porcine articular cartilage. J Forensic Leg Med 18:52–56.

Ruifrok AC, Johnston DA (2001). Quantification of histochemical staining by color deconvolution. Anal Quant Cytol Histol 23:291-9.

Sacco MA, Cordasco F, Scalise C, Ricci P, Aquila I (2022). Systematic Review on Post-Mortem Protein Alterations: Analysis of Experimental Models and Evaluation of Potential Biomarkers of Time of Death. Diagnostics 12(6):1490.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. (2012). Fiji: An open-source platform for biological-image analysis. Nat Methods 9(7):676–82.

Shrestha R, Kanchan T, Krishan K (2024). Methods of Estimation of Time Since Death. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing.

Tomsia M, Cieśla J, Pilch-Kowalczyk J, Banaszek P, Chełmecka E (2022). Cartilage Tissue in Forensic Science—State of the Art and Future Research Directions. Processes 10:2456.

Wei W, Michu Q, Wenjuan D, Jianrong W, Zhibing H, Ming Y, Bo J, Xia L (2020). Histological changes in human skin 32 days after death and the potential forensic significance. Sci Rep 10:18753.

Zhang Y, Liu L, Ren L (2020). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) determination of cantharidin in biological specimens and application to postmortem interval estimation in cantharidin poisoning. Sci Rep 10:10438.

Zissler A, Stoiber W, Steinbacher P, Geissenberger J, Monticelli FC, Pittner S (2020). Postmortem Protein Degradation as a Tool to Estimate the PMI: A Systematic Review. Diagnostics 10:1014.

Zupančič B, Ugwoke CK, Abdelmonaem MEA, Alibegović A, Cvetko E, Grdadolnik J, Šerbec A, Umek N (2023). Exploration of macromolecular phenotype of human skeletal muscle in diabetes using infrared spectroscopy. Front Endocrinol (Lausanne) 14:1308373.

Zupančič B, Umek N, Ugwoke CK, Cvetko E, Horvat S, Grdadolnik J (2022). Application of FTIR Spectroscopy to Detect Changes in Skeletal Muscle Composition Due to Obesity with Insulin Resistance and STZ-Induced Diabetes. Int J Mol Sci 23:12498.