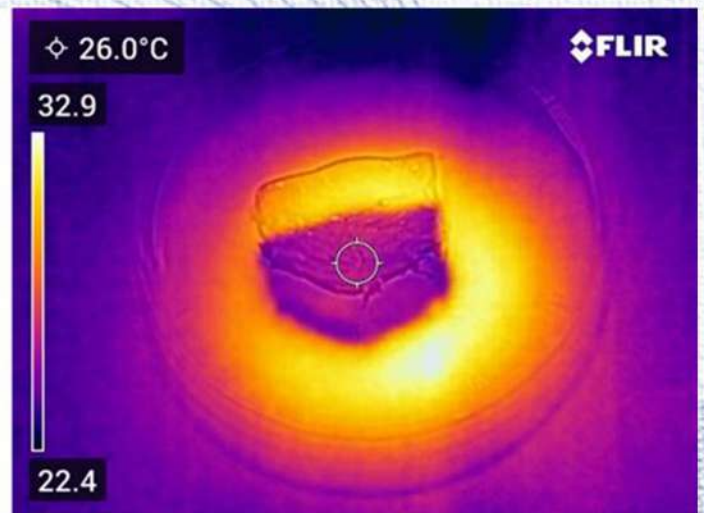
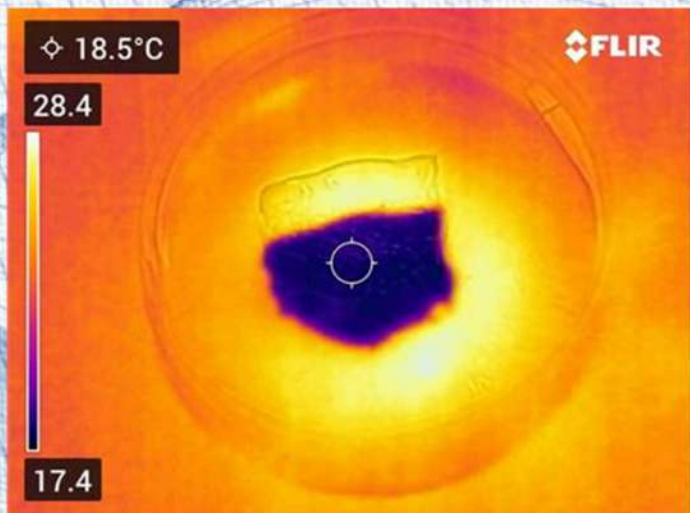


Journal

of Biomedical Photonics & Engineering



Editor-in-Chief:
V V Tuchin

Journal of Biomedical Photonics & Engineering

Volume 12, Issue 3 (September 2026)

Editor-In-Chief

Valery V. Tuchin – Saratov State University, Russian Federation;
Institute of Precision Mechanics and Control RAS, Russian Federation; a corr.-member of RAS

Deputy Editor-In-Chief

Valery P. Zakharov – Samara National Research University, Russian Federation

Postal address:

39B Lukacheva st., room #301

Samara 443086, Russian Federation

Used Figures from this issue on the Cover

Traian V. Chirila et al., Photocrosslinking of Tarsal Collagen for Treating Eyelid Laxity: Applying the Exposure Reciprocity Law to Establish an Optimal Irradiant Profile, 030301, Fig. 1.

Contents

Reviews

In Vivo Toxicity of Upconversion Nanoparticles. Overview. Part I 030201
Irina Y. Yanina, Vyacheslav I. Kochubey

In Vivo Toxicity of Upconversion Nanoparticles. Overview. Part II 030201
Irina Yanina, Vyacheslav I. Kochubey

Articles

Photocrosslinking of Tarsal Collagen for Treating Eyelid Laxity: Applying the Exposure Reciprocity Law to Establish an Optimal Irradiant Profile 030301
Traian V. Chirila, Shuko Suzuki, Nadja E. Pop, Claire Cattelot, Alexandra I. Manta

Polarized Time-Resolved Fluorescence Reveals Enzyme-Specific Anisotropy Dynamics of Bound NADPH 030302
Ioanna A. Gorbunova, Denis A. Volkov, Maxim E. Sasin, Oleg S. Vasyutinskii

Doppler Ultrasound Indices of the Major Arteries as Markers of Vascular Tone during Thermal Stress 030303
Anatoly V. Skripal, Konstantin V. Mashkov, Andrey A. Sagaidachnyi, Elena S. Olenko

A Graph Path Interpretability Metric for Explainable Classification of Dermoscopic Skin Lesions 030304
Aleksandra N. Zhdanova

Leakage-Free Biomedical Acoustic Signal Classification of Infant Cry for Reliable Neonatal Monitoring 030305
Humam Abdullah Khudair, Ziyad K Farej, Omar Younis Alani

Development of a Method for the Intelligent Analysis of Raman Spectra of Blood Serum 030306
Irina A. Pimenova and Irina A. Matveeva

In Vivo Toxicity of Upconversion Nanoparticles. Overview. Part I

Irina Yu. Yanina* and Vyacheslav I. Kochubey

Institute of Physics, Saratov State University, 83 Astrakhanskaya str., Saratov 410012, Russian Federation

*e-mail: irina-yanina@list.ru

Abstract. This review synthesizes current research on the *in vivo* dark toxicity of upconversion nanoparticles (UCNPs) for biomedical applications, including luminescent imaging, diagnostics, and photoinduced therapy. The aim is to assess key toxicity determinants – nanoparticle type, surface coating, concentration, and dose relative to body weight – and to establish safe exposure thresholds. Analysis of comprehensive histological and biochemical studies indicates that UCNPs generally exhibit low toxicity under tested conditions. The most non-toxic concentration identified across studies is 2 mg/mL. For mice, the maximum permissible dose for short-term exposure (12 h) is 10 mg/kg, and for long-term exposure (120 days) is 50 mg/kg. For rats, a dose of 5 mg/kg is safe for both short- and long-term studies. While positively charged coatings (e.g., PEI) enhance cellular uptake, they may increase cytotoxicity compared to neutral or negatively charged coatings (e.g., PAA, PEG). Dissolution of UCNPs releasing rare-earth and fluoride ions contributes to dark toxicity, but biocompatible shells (silica, PEG) reduce this effect. Overall, UCNPs demonstrate favorable biocompatibility at established doses, supporting their continued development for *in vivo* theranostic applications. This review is the first of three parts; subsequent parts address phototoxicity and toxicity of multimodal particles.

Keywords: toxicity; upconversion nanoparticles; concentration; dose.

Paper #9595 received 17 Apr 2026; revised manuscript received 21 May 2026; accepted for publication 26 May 2026; published online 7 Sep 2026. doi: [10.18287/JBPE26.12.030201](https://doi.org/10.18287/JBPE26.12.030201).

1 Introduction

The field of *in vivo* cellular and molecular research has experienced significant methodological advancements in imaging over recent decades [1–5]. Among the most promising techniques is luminescent imaging, which offers a noninvasive means to elucidate tissue architecture, detect specific biomolecules, and monitor biological processes in real time [6–10].

The foundation of this approach lies in the use of highly luminescent agents. For these probes to be effective in biomedical applications, they must satisfy several key requirements. First, they should be amenable to chemical modification, allowing researchers to enhance their biocompatibility or confer targeting capabilities [11–13]. Second, they must exhibit low toxicity to ensure minimal adverse effects on living systems. Third, and critically, they should absorb and

emit light within the 650–1450 nm range – encompassing the red and near-infrared (NIR) regions of the spectrum.

This spectral window confers significant advantages for *in vivo* imaging. Notably, light in this range experiences minimal absorption by biological tissues (with the exception of water absorption bands), enabling excitation photons to penetrate deeply into living tissue [14]. This property stands in stark contrast to visible light, which is substantially attenuated by tissue components.

Another pivotal advantage of NIR visualization lies in its ability to improve the signal-to-noise ratio. Since endogenous autofluorescence predominantly occurs in the visible spectrum (400–650 nm), imaging in the NIR range drastically reduces background interference [15]. This spectral separation enhances detection sensitivity,

making it possible to reliably identify targets even at low concentrations.

Furthermore, NIR photons exhibit reduced scattering in biological tissues compared to visible light. This characteristic translates into superior spatial resolution and sharper image contrast, particularly beneficial when imaging dense or heterogeneous tissues. The combination of deep tissue penetration, low background noise, and high resolution positions NIR-emitting luminescent probes as invaluable tools for *in vivo* applications [16].

Ongoing research continues to focus on optimizing probe design – particularly through surface functionalization and fine tuning of emission properties – to further expand the utility of luminescent imaging in preclinical and potentially clinical settings.

Moreover, the luminescence properties of phosphors are sensitive to their environment, which enables their use as biosensors for detecting various biological parameters. These include: intracellular temperature (via luminescent thermometers); oxygen levels (for monitoring oxygen content in living systems); specific molecules (such as proteins, amino acids, antibiotics, and DNA) and various ions [17–24].

A special class of phosphors can transfer absorbed excitation energy to molecular triplet oxygen ($^3\text{O}_2$), promoting it to the excited singlet state ($^1\text{O}_2$) [25, 26]. Singlet oxygen is highly chemically reactive: when produced within cells, it can damage cellular structures and trigger cell death [27]. This mechanism underpins photodynamic therapy – one of the most promising approaches to cancer treatment [26].

Consequently, luminescent compounds have found wide-ranging applications in biology and medicine. This has spurred a significant increase in research publications focused on developing luminescent materials for biological imaging, disease diagnostics, and therapeutic applications in recent years.

Among luminescent materials, upconversion nanoparticles (UCNPs) stand out as particularly promising. Their emission arises from cooperative energy transfer processes. The key feature of UCNPs is that excitation requires the absorption of at least two photons, which enables excitation in the infrared (IR) region [15]. Meanwhile, the resulting luminescence exhibits shorter wavelengths — a phenomenon driven by the anti-Stokes process.

This unique mechanism provides UCNPs with an important additional advantage. Since biological tissues typically exhibit luminescence at longer wavelengths compared to the excitation wavelength, the background luminescence of biotissues at the UCNP emission wavelengths is substantially reduced [15]. Consequently, UCNPs not only retain the beneficial properties common to luminescent materials but also enable significantly improved signal-to-background ratios in biological imaging applications.

The use of nanoparticles in medicine and biology necessitates careful consideration of their potential toxicity. Due to their small size, nanoparticles

inherently pose toxicity risks – their high surface-area-to-volume ratio enhances reactivity and may lead to unintended biological interactions [28–30]. The influence of various factors on *in vivo* dark toxicity is shown in Fig. 1.

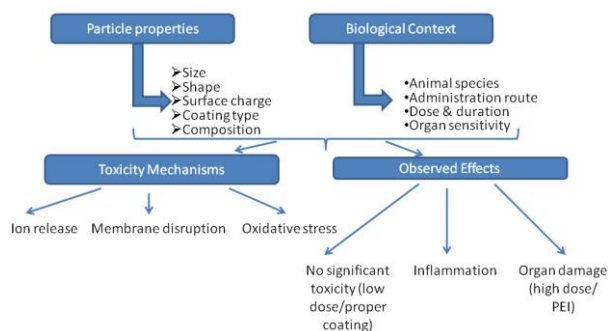


Fig. 1 Hierarchical diagram of factors affecting *in vivo* dark toxicity of UCNPs.

However, it is crucial to emphasize that the toxicity of UCNPs is not uniform; it depends on a range of physicochemical and physiological factors.

Physicochemical properties that influence toxicity include: particle size, shape, surface area, chemical composition and surface charge and functionalization (e.g., coating with biocompatible polymers) [28–30].

Physiological factors that modulate toxicity comprise: the specific disease context (e.g., inflamed vs. healthy tissue), individual genetic variations, metabolic state of the organism, route of administration and dose and duration of exposure [31].

Thus, assessing UCNP safety requires a comprehensive evaluation that integrates both material characteristics and biological context. This enables the design of safer nanomaterials and informed risk-benefit analyses for biomedical applications [32].

The recommended size for optimal penetration of NPs is below 100 nm. However, this size may also pose a risk of toxicity due to their potential to penetrate cellular structures and organs via the circulatory system.

One crucial aspect of using UCNPs in biomedical applications lies in ensuring their biocompatibility on cells and/or organisms [32]. To achieve this several chemical groups, including silica [33–35], polyethyleneimine (PEI) [36–39], polyvinyl pyrrolidone (PVP) [39], polyacrylic acid (PAA) [39–41], and polyethylene glycol (PEG) [42], have been used for this purpose. Additionally, UCNPs must be functionalized with different ligands that specifically target the desired cells and organs.

The increased surface area leads to higher solubility, causing compounds that would normally be insoluble to dissolve. As a result, toxicity can occur when these dissolved products – consisting of molecules and ions from the nanoparticle – enter living cells.

Increasing the relative surface area of nanoparticles not only affects their solubility, but also enhances their chemical activity. This can lead to increased toxicity, as the nanoparticles' chemical activity can cause harmful

effects. The toxicity of nanoparticles depends on their composition, the materials used to create their shells, and the surface charge of the nanoparticles [28–30].

When analyzing the causes of toxicity, it is also important to consider the surface charge of nanoparticles. This charge depends on the presence or absence of a shell around the nanoparticle. So-called “bare” nanoparticles, which are nanoparticles without a protective shell, can be obtained by either removing of the shell by reacting with acids (HCl) or by replacing it with other ligands (BF_4^-). However, it should be noted that the term “bare” is somewhat arbitrary, as these nanoparticles quickly acquire a negative charge when exposed to aqueous solutions due to the presence of OH^- ions.

It is important to note that the toxicity of nanoparticles is essential for their therapeutic use. This toxicity may lead to apoptosis or necrosis of the cells that are penetrated by the nanoparticles. The aim of targeted therapy is to specifically deliver nanoparticles to pathological cells, so that these cells experience the toxic effects of treatment.

However, it is also important to consider the potential undesirable toxic effects that can arise.

Toxicity can occur at different levels, including the cellular, subcellular, and molecular levels, as well as in altered organ function.

Nanoparticles injected into the bloodstream can reach the liver, spleen, lungs, heart, and brain [36]. The biologically inert shells of these nanoparticles allow them to circulate in the blood for a long time, leading to their accumulation in various organs. The presence of addressable molecules on the nanoparticles' shells can reduce their circulation time.

The toxicity of nanoparticles plays a crucial role in their application in photodynamic therapy and other biological uses and serves as a primary selection criterion. Depending on the intended application, particles with different types of coatings or composition are developed. There are several reviews on the toxicity of UCNPs [43–45]. However, the data presented in the literature are often contradictory, particularly regarding the maximum permissible concentration of nanoparticles. And despite the large number of studies on toxicity and antitumour effects in animals, the authors sometimes fail to provide important information such as the dose and concentration of particles used. On the other hand, as a rule, the authors do not investigate the mechanism of toxicity. The authors investigate the toxicity of a certain type of nanoparticles with a given coating for various cells, organisms and organs.

2 *In Vivo* Dark Toxicity of Nanoparticles

Nanoparticle toxicity can be divided into two main types: intrinsic toxicity and light-activated toxicity (phototoxicity).

Phototoxicity is usually caused by the generation of reactive oxygen species (ROS) or nitrogen. Both the nanoparticle surface itself and its shell can participate in the generation process. To increase activity, the shell

may include photodynamic dyes. In this case, photoexcitation of the dye occurs via photons of nanoparticle luminescence, which is excited by external radiation. The resulting products can cause DNA damage, which not only affects cell growth through protein oxidation but also impacts mitochondrial respiration [32].

As mentioned earlier, one of the causes of dark toxicity is the dissolution of nanoparticles. This process leads to the formation of ions of rare earth elements and fluorine, which can have a negative impact on cells and the body.

The dissolution of nanoparticles in the blood may generate yttrium ions, which could affect bone and bone marrow cell function [46]. Additionally, there is a report of the possibility of toxicity from the substitution of calcium ions (Ca^{2+}) with gallium ions (Ga^{3+}) [47].

The effect of lanthanides on blood composition is also well-known [45].

In the absence of lanthanide ions in UCNPs, their toxicity is reduced. Tian et al. intravenously administered 1200 mg/kg of upconversion nanocapsules to mice over 60 days [48]. Under these conditions, no significant toxicity was observed, as evidenced by body weight determination, histological and hematological analyses, and biochemical blood tests. Meanwhile, Yang et al. demonstrated that at a dose of 28 mg/kg, UCNPs caused temporary lung and liver damage in mice [49].

BALB/c mice were injected with 200 μL of high-dose solution (2.0 mg [Y]/kg mouse weight) and low-dose solution (0.2 mg [Y]/kg mouse weight) PEG-coated $\text{NaYF}_4:\text{Yb}^{3+}/\text{Er}^{3+}$ nanoparticles through the tail vein for the experimental group, while the control group received an equivalent amount of phosphate-buffered saline (PBS). Major organs were collected 24 and 72 h later, sectioned and subjected to haematoxylin & eosin (H&E) staining to observe histopathological changes for toxicity assessment [50]. There was no notable histopathological damage in the major organs of the control mice. The heart, liver, spleen, lungs, and kidneys of the experimental group exhibited general consistency with those of the control mice, showing normal tissue cell morphology and an absence of acute inflammatory responses. Moreover, no inflammatory reactions were detected in tissue sections taken at 24 h and 72 h post-injection, suggesting that neither short-term nor prolonged exposure to UCNPs induced organ damage in the mice. The metabolic and clearance pathways of UCNPs in mice were investigated by quantifying yttrium levels in feces and urine. The study revealed that UCNPs were primarily metabolized through fecal excretion within the initial four hours, shifting towards urinary excretion thereafter. Complete elimination was achieved within 72 h.

The dissolution rate of nanoparticles depends on the condition of their surface. If a nanoparticle's surface is not stabilized, it is likely to dissolve more easily than if it is protected by specialized coatings. It is important to consider the potential toxicity of these coatings,

however, as they may also affect the dissolution process.

Shells made of silica, zinc sulfide, or polymers are often used to enhance the biocompatibility of nanoparticles and decrease their solubility. However, research has shown that silica can interact with cells and cause inflammation, which may be linked to various diseases, such as rheumatoid arthritis, sclerosis, lupus, and chronic kidney disease [51, 52]. At the same time, the silica coating reduces the solubility of nanoparticles and the cytotoxicity caused by fluoride ion release during dissolution [53].

Studies on the biocompatibility of nanoparticles coated with a SiO₂ layer have shown that they exhibit low dark cytotoxicity, can penetrate cells, and remain inside for a considerable time. For instance, intravenous administration of coated with SiO₂ NaYF₄:25%Yb³⁺, 0.3%Tm³⁺ particles at a dose of 10 mg/kg to laboratory rats resulted in no adverse health effects or behavioral changes throughout the study (7 days) [33, 34]. Recently, it was shown that oral or intravenous administration of NaYF₄:Yb/Er@SiO₂ nanoparticles does not exhibit significant toxicity in mice, even at high doses of 100 mg/kg body weight [35]. The study revealed that the nanoparticles mainly accumulate in bones, the stomach, and intestines upon oral administration, whereas intravenous injection leads to accumulation primarily in the liver and spleen.

The Zhang group reported that NaYF₄ nanoparticles coated with SiO₂ are predominantly excreted from the bodies of mice by the 7th day after injection at a dose of 10 mg/kg [33]. A review of the properties and toxicity of functionalized UCNPs is provided in studies [54–56].

The experimental results show that the positively charged UCNPs@SiO₂-PEG-NH₂ has a stronger cytotoxic effect and better cell internalization ability than the negatively charged UCNPs@SiO₂-COOH and UCNPs@SiO₂-PEG. This charge-based difference in cytotoxicity may be caused by the redox stress caused by UCNPs@SiO₂, as well as the release of metal ions and ligands [57].

The toxicity of nanoparticles varies depending on the object they affect. In a study of the toxicity of UCNPs on keratinocytes, the toxicity was maximal when coated with PEI, weak cytotoxicity when coated with polymaleic acid (PMAO), and virtually non-toxic in the case of low molecular weight tetraethylammonium hydroxide (TMAH) [58]. When exposed to nerve cells, the toxicity differs. The TMAH coating allows nanoparticles to be hydrophilized by replacing oleic acid on their surface with a polymer, resulting in the appearance of hydroxide (OH⁻) radicals on the surface. This causes the nanoparticles to become negatively charged, similar to when coated with PMAO. However, the surface potential of PMAO is twice that of TMAH in this case.

There is no clear consensus on the toxicity of nanoparticles coated with PEI. PEI and nanoparticles based on it may cause toxic effects in cells, and the toxicity increases with molecular weight and branching

degree. These toxic effects are associated with the interaction between the cationic polymer and cell membranes, which can disrupt them. PEI can destroy cellular and mitochondrial membranes, leading to necrosis or apoptosis.

However, there are also studies showing that introducing PEI into nanoparticle's shells does not significantly affect their toxicity [59].

When PEI was used as a coating for UCNPs (NaYF₄:Yb³⁺,Er³⁺) and administered intravenously to laboratory rats at concentrations of 1.5, 10, 15, and 25 µg/mL and doses of 10 mg/kg, no toxic effects on cells or mice were observed [36]. Studies examined PEI-modified NaYF₄:Yb,Er nanoparticles with different administration methods in mice: intravenous, intraperitoneal, and intragastric (at doses of 10 mg/kg) [37, 38]. Following intraperitoneal administration, PEI@UCNPs accumulated significantly in the spleen for 30 days. In contrast, intragastric administration resulted in a decrease in accumulation over time in various body tissues, with particles primarily detected in the ileum and cecum, while their presence in other organs was relatively low. For intravenous administration, UCNPs demonstrated excretion from the body within 30 days and a gradual decrease in accumulation in the spleen.

Thus, high molecular weight PEI and high concentrations can be cytotoxic [60]. For medical purposes, modified forms of PEI are being developed, such as pegylation and acetylation, to reduce surface charge and toxicity.

For example, in a study, the effect of PEI, PVP, and PAA coatings on the cytotoxicity and cellular penetration of NaYF₄:Yb/Er nanoparticles was investigated [39]. These polymers carry different charges: PVP is neutral, PEI is positively charged, and PAA is negatively charged. Thus, this study aimed to explore the impact of surface charge on particle biocompatibility. NaYF₄:Yb/Er-PVP was synthesized via solvothermal methods, while NaYF₄:Yb/Er-PEI and NaYF₄:Yb/Er-PAA were obtained by replacing PVP through interactions with PEI and PAA in heated solution. Cytotoxicity tests on HeLa and U87MG (human glioblastoma) cell lines revealed that the least toxic material was the negatively charged NaYF₄:Yb/Er-PAA. At the same time, confocal microspectroscopy using 980 nm laser excitation on the same cell lines showed that NaYF₄:Yb/Er-PEI exhibited significantly higher penetration efficiency: NaYF₄:Yb/Er-PEI > NaYF₄:Yb/Er-PVP > NaYF₄:Yb/Er-PAA. This can be explained by the negative charge of the cell membrane, which has a greater affinity for positively charged NaYF₄:Yb/Er-PEI particles. For NaYF₄:Yb/Er-PEI, which demonstrated the highest penetration ability, the mechanism of cellular entry was studied, showing that the particles enter cells via clathrin-mediated endocytosis.

PAA is a non-toxic and biocompatible substance that can be biodegraded. However, there is still some uncertainty about the potential toxicity of PAA-coated

nanoparticles, as there is no clear consensus on this issue. Additionally, there is currently no direct evidence to support the toxicity of these nanoparticles themselves.

Intravenous administration of functionalized $\text{NaYF}_4:20\%\text{Yb}^{3+}$, $1\%\text{Tm}^{3+}$ -PAA particles at a dose of 15 mg/kg to mice did not result in toxic effects on cells or mice, with only minor hyperplasia observed in the per arteriolar lymphoid sheath of the white pulp [40]. Additionally, no significant toxicity was observed over a 115-day period [41].

In contrast, PEG-coated nanoparticles have been shown to be non-toxic due to the use of PEG to modify their surface. This modification process can improve the biophysical and chemical properties of the nanoparticles, such as reducing the adsorption of plasma proteins onto their surface and increasing their circulation time within the body.

In study, mice were intravenously injected with 200 μL of UCNP ($\text{NaYF}_4:\text{Yb}^{3+}/\text{Tm}^{3+}$ coated with PEG or functionalized with PAA) at a concentration of 2 mg/mL (dose of 20 mg/kg), and their condition was analyzed at various time points post-injection (3, 7, 20, 40, and 90 days, with five mice per time point) [42]. No noticeable toxic side effects were observed in laboratory animals.

UCNPs $\beta\text{-NaYF}_4:20\%\text{Yb}^{3+}$, $2\%\text{Tm}^{3+}$, coated with octylamine-PAA or 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-PEG, when intravenously injected into the vena cava at a volume of 200 μL and concentration of 2 mg/mL (dose of 20 mg/kg), did not cause any evident histological damage or organ impairment. Biochemical, hematological, and histological analyses remained normal at a dose of 20 mg/kg. UCNP can be partially degraded or processed within macrophage cells without significant release [42].

In paper, spherical or hexagonal $\text{NaYF}_4:\text{Yb},\text{Er}$ UCNP with sizes of 25 nm (S-UCNP) and 120 nm (L-UCNP) were synthesized by high-temperature coprecipitation and subsequently modified with three kinds of polymers [29]. These included PEG and poly(*N,N*-dimethylacrylamide-co-2-aminoethylacrylamide) [P(DMA-AEA)] terminated with an alendronate anchoring group, and poly(methyl vinyl ether-co-maleic acid) (PMVEMA). To evaluate the biodistribution of the nanoparticles, 100 μL of UCNP@Ale-PEG or UCNP@Ale-(PDMA-AEA) dispersion (concentration 50 $\mu\text{g}/\text{mL}$ of mouse blood) was applied retroorbitally to the mice (each nanoparticle type to 5 mice). An *in vivo* study showed that both types of nanoparticles were eliminated from the body via the liver. While L-UCNP@Ale-PEG particles were almost completely eliminated from the liver 96 h after intravenous application, L-UCNP@Ale-(PDMA-AEA) particles remained in the liver in a significantly higher amount. The polymeric coating can influence the retention of nanoparticles in tissues to some extent. This can be advantageously used in various applications using cell labeling in tissues and organs.

To investigate the distribution and translocation of $\text{CitNaYb}_{0.98}\text{-xEr}_{0.02}\text{F}_4:5\%\text{Fe}^{2+}$ UCNP in the main organs of mice, two nude mice were intravenously injected with UCNP (2 mg/mL, 200 μL) and, after 1 or 24 h, dissected, and their main organs (heart, lung, liver, spleen, and kidney) were examined using fluorescence molecular tomography (FMT) imaging system [61]. At both 1 and 24 h post-injection, obvious fluorescence in lung, liver, and spleen is observed. However, the fluorescence intensity in the different organs changes dramatically with time. In a short period of time (1 h), the fluorescence distribution is relatively uniform, and the fluorescence intensity is high. After 24 h, the fluorescence intensity of lung, liver, and spleen decreases significantly, especially in the lung and spleen. It is worth noting that the heart also shows a faint fluorescence. These results suggest that UCNP may have been partially metabolized by the mice and that a small amount of UCNP is transferred to the heart. Therefore, $\text{Cit-NaYbF}_4:2\%\text{Er}^{3+}/5\%\text{Fe}^{2+}$ may provide visual diagnostic and therapeutic means for lung, liver, and spleen diseases.

To assess the *in vivo* biotoxicity, UCNP-CuMnO was used in the experimental group at a concentration of 100 μL and 5 mg/mL, while the control group received the same volume of saline solution [62]. The H&E staining analysis of five main organs (heart, liver, spleen, lung and kidney) showed that there is no distinct abnormality observed 7 days after injection. These results fully demonstrate the good biocompatibility of UCNP-CuMnO nanoprobe.

Additional decoration of the surface of nanoparticles with photodynamically active dyes has practically no effect on their toxicity.

In addition to detectability, the assessment of the new nanoparticles should also include their biocompatibility. Nanoparticles consisting of the miR-122 aptamer (Apt), UCNP, and Prussian blue nanoparticles (PBNP) (UCNP-Apt-PBNP) were studied in recent paper [63]. To assess the safety and potential toxicity of these nanoparticles *in vivo*, a dose of 20 mg/kg UCNP-Apt-PBNP was selected, based on the amount of UCNP. This dose was the highest used in previous studies by the authors, and it was injected into mice for 14 days. The nanoparticles did not cause nephrotoxic or hepatotoxic effects, indicating their biocompatibility and potential use for imaging the liver using miR-122 targeting.

PEG-coated monodisperse UCNP with a conjugated Flamma® fluorescent dye particles (UCNP@Ale-PEG-Flamma® particle) studied in the work [64]. Each group of mice was injected either intraperitoneally or intravenously (tail vein) with 100 μL of UCNP@Ale-PEG-Flamma® particle (0.15 mg) dispersion. After another 24 h, mice were euthanized and tumors and selected tissues (liver, kidney and muscle) were dissected. Intravenously administered UCNP@Ale-PEG-Flamma® particles accumulated strongly in the liver compared to intra peritoneal injection. The lower hepatic accumulation may be

explained by the reduced clearance rate by the mononuclear phagocytic system. In contrast, significantly increased nanoparticle accumulation in orthotopic tumors was observed in animals administered by the intraperitoneal route. The measured mean fluorescence intensity (MFI) 24 h after injection was more than 6-fold higher than that of intravenous administration. In addition, a slightly increased fluorescence signal from the kidneys of mice was observed after intravenous administration. Although insignificant, this increase may be explained by direct exposure of intravenously administered UCNPs to renal excretion mechanisms, whereas intraperitoneal administration leads to lower (and delayed) blood levels. Moreover, no differences in particle accumulation were found in skeletal muscle, which was chosen as a negative control to the remaining three tissues where accumulation was anticipated. However, this was an expected finding as there is no reason for particle accumulation in skeletal muscle. The observed differences in nanoparticle's distribution between intravenous and intraperitoneal administration may be attributed to distinct physiological barriers and clearance mechanisms involved. Intravenous administration exposes nanoparticles to immediate systemic circulation, leading to rapid clearance by organs such as the liver and kidneys. In contrast, intraperitoneal administration offers a localized delivery route that potentially bypasses some of the mechanisms of systemic clearance and allows for more targeted accumulation in peritoneal tumors. These differences in biodistribution underscore the importance of the administration route in achieving optimal therapeutic outcomes with nanoparticle-based drug delivery systems.

In the study, $\text{Ca}_{0.2}\text{Zn}_{0.9}\text{Mg}_{0.9}\text{Si}_2\text{O}_6$ particles doped with rare earth cations (Eu^{3+} , Dy^{3+} , Mn^{3+}) and coated with PEG at a concentration of 1 mg/mL were excited with UV light (254 nm) for 5 min [65]. Subsequently, 100 mL of a colloidal solution was introduced (1 mg of stably luminescent nanoparticles was resuspended in 1 mL of sterile 150 mM NaCl). It was found that the distribution of particles in tissues highly depends on the surface coating and the core size of the particle. Smaller particles (up to 80 nm) and star shaped particles remain in circulation for a longer time, whereas PEG coatings with a molecular weight of 5 to 20 kDa have only a minor effect on the biodistribution of silica coated nanoparticles *in vivo*.

A simple fluorescence resonance energy transfer (FRET) sensing platform (termed as USP), comprised of UCNPs as the energy donor and Cy5 as the energy acceptor, has been synthesized for cathepsin B (CTSB) activity detection *in vitro* and *in vivo* [66]. For *in vivo* toxicity analysis, 0.2 mL stroke-physiological saline solution (0.9 wt% NaCl solution) with or without USP ($10 \text{ mg} \cdot \text{kg}^{-1}$) were injected intravenously into the Kunming mice as experimental group and control group, respectively. Healthy mice were injected intravenously with USP and dissected at the 30th day

post-injection. The results of body weight, histological analysis, and blood routine examination show that there is no significant difference between USP-treated mouse group and control group, demonstrating the low long-term toxicity of USP.

In study, it was developed and characterized Arg-Gly-Asp (RGD)-decorated poly-lactic-acid-co-glycolic acid (PLGA) nanoparticles for simultaneous sustained and precise co-delivery of loaded cisplatin (CDDP) (therapeutic agent) and UCNPs (imaging agent) for lung cancer therapy [67]. Wistar rats were categorized ($n = 4$) in different groups and treated with CDDP loaded PLGA nanoparticles (CPF), CDDP and UCNPs loaded PLGA nanoparticles (CPFT), RGD decorated CDDP and UCNPs loaded PLGA nanoparticles (CPFT- RGD nanoparticles) formulations, Cisrest-50 injection (positive control) and normal saline solution (negative control), separately. The formulations at dose of CDDP (5 mg/kg) and normal saline were administered into rats via the tail vein with an administration rate 0.4 mL/min. After 15 days and 30 days of i.v. administration of CPF, CPFT, CPFT- RGD nanoparticles and Cisrest-50 injection to rats, the morphological characteristics of collected lung tissues were assessed by standard H&E staining. In the study, the alveoli and interstitium were observed normal in the control group treated with normal saline. No pathologic lesions (i.e., inflammation, alveolar damage, vascular and alveolar congestion) were observed in alveoli or bronchi of lung tissue. In the group treated with CPF, CPFT, CPFT-RGD nanoparticles, there was alveolar septal disruption (probably artificial) and fibrin thrombus in a few small vessels were found, with no significant alveolar hemorrhage or inflammation and bronchi are unremarkable as compared to control group. In Cisrest-50 treated group, lung parenchyma showed mild to chronic bronchial inflammation, as well as disruption of alveolar septa, small fibrin thrombi and hemorrhage. In all treated groups, no severe pathologic lesions (i.e., alveoli, bronchiole (lower right), capillaries and interstitium) were observed in the lung tissue as compared to Cisrest-50 treated groups. This result clearly indicates that CDDP loaded PLGA nanoparticles improved the biocompatibility of CDDP with lung tissue.

In tumor cells (for example, in acute lymphoblastic leukemia), asparagine is an essential amino acid; without it, the synthesis of proteins and nucleic acids is disrupted, leading to cell death.

The authors functionalized NaYF_4 UCNPs with isocyanate groups and then immobilized the enzyme pegylated L asparaginase (L-ASNase) (PEG-L-ASNase) via covalent bonding [68]. As noted above, pegylation of the surface reduces the toxicity of nanoparticles.

The nanoparticles injected into mice were irradiated with a 150 mW laser at 808 nm for 30 min at a distance of 1 cm. The histological pattern of heart and liver tissues did not change compared to the control. In kidney tissues, a slight increase in the density of degenerated cells in the renal medulla was observed in

the case of non pegylated UCNPs and after laser irradiation. In the same groups of mice, an increased density of degenerated cells was found in spleen tissues, and giant cells with nuclei were detected.

However, in the literature, similar tissues were histopathologically examined in *in vivo* studies with UCNPs containing various functional groups, and it was indicated that the toxic damage was minor [42, 69, 70].

At 48 h post-UCNP administration, high concentrations of Er^{3+} were observed in the spleen, while Er^{3+} levels remained quite low in the liver, heart, and kidneys.

Zhao and others reported that nanoparticles (NPs), which are considered foreign objects, primarily accumulate in the liver and spleen [71].

Since the biodistribution of nanomaterials is dictated by their surface chemistry – which affects the extent and

specificity of protein binding – modification with PEG-L-ASNase could facilitate the elimination of UCNPs from the body and prevent long term accumulation after intravenous administration [70]. Consequently, such particles can be safely used for *in vivo* administration.

When a suspension of $\text{NaGdF}_4:\text{Tm}^{3+}$, Er^{3+} , Yb^{3+} particles coated with oleic acid oxidized to azelaic acid was administered (volume: 150 μL ; concentration: 1 mg/mL), their high relaxivity of 5,60 $\text{s}^{-1}\text{mM}^{-1}$ was demonstrated [72].

These nanoparticles were successfully used as contrast agents for *in vivo* MRI [73].

The dependence of toxicity on such parameters as particle type, particle coating type, dose, concentration and method of particle administration, summarizing the reviewed articles, is presented in Table 1.

Table 1 Dependence of toxicity on such parameters as particle type, particle coating type, dose, concentration and method of particle administration.

UCNP type	Coating	Charge	Type of animal	Dose (mg/kg)	Concentration (mg/ml)	Route of administration	Main results (toxicity etc.)	Ref.
$\text{NaYF}_4:\text{Yb}^{3+}/\text{Er}^{3+}$	PEG	Not specified (neutral)	BALB/c mice	2.0 mg [Y]/kg (high), 0.2 mg [Y]/kg (low)	Not specified	Intravenous (tail vein)	No histopathological damage in heart, liver, spleen, lungs, kidneys at 24 and 72 h; no inflammation; complete elimination in 72 h via feces then urine	[50]
$\text{NaYF}_4:\text{Yb}^{3+}/\text{Tm}^{3+}$	SiO_2	Not specified	Rats	10	Not specified	Intravenous	No adverse health effects or behavioral changes for 7 days	[33,34]
$\text{NaYF}_4:\text{Yb}/\text{Er}$	SiO_2	Not specified	Mice	100	Not specified	Oral or intravenous	No significant toxicity; oral → accumulation in bones, stomach, intestines; i.v. → liver and spleen	[35]
$\text{NaYF}_4:\text{Yb}^{3+}, \text{Er}^{3+}$	PEI	Positive	Rats	10	0.0015, 0.01, 0.015, 0.025	Intravenous	No toxic effects on cells or mice	[36]
$\text{NaYF}_4:\text{Yb}, \text{Er}$	PEI	Positive	Mice	10	Not specified	Intravenous, intraperitoneal, intragastric	Intraperitoneal → spleen accumulation for 30 days; intragastric → ileum and cecum; i.v. → excretion within 30 days	[37,38]

Table 1 (Cont.)

NaYF ₄ :20% Yb ³⁺ , 1%Tm ³⁺	PAA	Negative	Mice	15	Not specified	Intravenous	No toxic effects; minor hyperplasia in splenic white pulp	[40]
NaYF ₄ :Yb ³⁺ /Tm ³⁺	PEG or PAA	Neutral (PEG) / Negative (PAA)	Mice	20	2	Intravenous	No noticeable toxic side effects up to 90 days; no histological damage	[42]
β-NaYF ₄ : 20%Yb ³⁺ , 2% Tm ³⁺	Octylamine-PAA or DSPE-PEG	Negative (PAA) / Neutral (PEG)	Mice	20	2	Intravenous (vena cava)	No histological damage or organ impairment; biochemical/hematological analyses normal	[42]
NaYF ₄ :Yb,Er (25 nm and 120 nm)	Ale-PEG or Ale-(PDMA-AEA)	Not specified	Mice	~5 (based on 50 µg/ml blood, 100 µl)	0.05 (in blood)	Retroorbital	Both eliminated via liver; L-UCNPs@Ale-PEG almost fully cleared at 96 h; L-UCNPs@Ale-(PDMA-AEA) remained longer	[29]
Cit-NaYbF ₄ :2% Er ³⁺ /5%Fe ²⁺	Citrate	Negative	Nude mice	~40 (2 mg/ml × 0.2 ml for ~20 g mouse)	2	Intravenous	Fluorescence in lung, liver, spleen at 1 and 24 h; decreased by 24 h; faint signal in heart	[61]
UCNPs-CuMnO	Not specified (likely PEG or similar)	Not specified	Mice	~25 (5 mg/ml × 0.1 ml for ~20 g mouse)	5	Not specified (likely intravenous)	No distinct abnormality in H&E staining of heart, liver, spleen, lung, kidney at 7 days	[62]
UCNP-Apt-PBNP (miR-122 aptamer)	Not specified (Apt-PBNP)	Not specified	Mice	20	Not specified	Not specified	No nephrotoxic or hepatotoxic effects after 14 days	[63]
UCNP@Ale-PEG-Flamma®	PEG + Flamma® dye	Not specified	Mice	~7.5 (0.15 mg per mouse, ~20 g)	Not specified	Intraperitoneal vs. intravenous	i.v. → high liver accumulation; i.p. → lower liver accumulation, 6× higher tumor accumulation	[64]

Table 1 (Cont.)

$\text{Ca}_{0.2}\text{Zn}_{0.9}\text{Mg}_{0.9}\text{Si}_2\text{O}_6\text{:Eu}^{3+},\text{Dy}^{3+},\text{Mn}^{3+}$	PEG	Not specified	Mice	~ 5 (1 mg/ml $\times$ 0.1 ml for ~ 20 g mouse)	1	Not specified (likely i.v. or i.p.)	Distribution depends on size and shape; smaller (<80 nm) and star-shaped remain longer in circulation	[65]
UCNP-Cy5 (USP, FRET sensor)	Not specified	Not specified	Kunming mice	10	Not specified	Intravenous	No significant difference in body weight, histology, blood routine at 30 days; low long-term toxicity	[66]
CDDP + UCNP in PLGA (CPFT-RGD)	PLGA + RGD	Not specified	Wistar rats	5 (as CDDP)	Not specified	Intravenous (0.4 ml/min)	No severe pathologic lesions in lung tissue; improved biocompatibility compared to free cisplatin	[67]
NaYF_4 with PEG-L-ASNase	PEG-L-asparaginase	Not specified	Mice	Not specified	Not specified	Not specified	No heart/liver changes; slight renal medulla degeneration in non-PEG group after laser; spleen giant cells in non-PEG + laser	[68]
$\text{NaGdF}_4\text{:Tm}^{3+},\text{Er}^{3+},\text{Yb}^{3+}$	Azelaic acid (from oleic acid oxidation)	Negative	Not specified (MRI study)	~ 0.15 (150 $\mu\text{l} \times 1$ mg/ml for ~ 1 kg assumed, but not specified)	1	Not specified	Used as MRI contrast agent; relaxivity $5.60\text{ s}^{-1}\text{mM}^{-1}$; no toxicity reported explicitly	[72,73]

Note: “Not specified” indicates the original article did not report the parameter.

Doses in mg/kg were calculated where possible based on typical mouse body weight (20 g) and injection volume; some values are approximate.

Charge assignments are based on known properties of coatings: PEI (+), PAA (–), PEG (neutral), PVP (neutral).

3 Conclusion

This review (Part I) has synthesized current knowledge on the *in vivo* dark toxicity of upconversion nanoparticles (UCNPs), with a focus on the physicochemical and physiological factors that determine their safety profile for biomedical applications. Based on comprehensive histological, biochemical, and hematological analyses reported in the literature, the following conclusions can be drawn.

UCNPs generally demonstrate low dark toxicity under tested conditions, provided that their surface is appropriately stabilized. The primary mechanism of dark toxicity is the dissolution of nanoparticles leading to the release of rare-earth ions (e.g., Y^{3+} , Yb^{3+} , Er^{3+}) and fluoride ions, which can accumulate in organs such as the liver, spleen, and bones. However, at recommended doses, no significant organ damage or adverse behavioral changes have been observed.

The toxicity of UCNPs is strongly modulated by their surface coating. Silica (SiO₂) coatings reduce solubility and fluoride ion release, showing low cytotoxicity at doses up to 100 mg/kg in mice. Polyethylene glycol (PEG) coatings provide excellent biocompatibility, reduce protein adsorption, and prolong circulation time without evident toxicity at 20 mg/kg. Polyacrylic acid (PAA) coatings (negatively charged) exhibit the lowest cytotoxicity among tested polymers, with no significant toxicity observed over 115 days at 15 mg/kg. Polyethyleneimine (PEI) coatings (positively charged) enhance cellular internalization but may induce membrane disruption and cytotoxicity at high molecular weights and concentrations; however, lower concentrations (10 mg/kg) have shown no toxic effects. Neutral coatings (PVP) demonstrate intermediate toxicity and penetration efficiency.

Positively charged UCNPs (e.g., PEI-coated) exhibit higher cellular uptake due to electrostatic attraction to negatively charged cell membranes, but also higher cytotoxicity. Negatively charged (PAA) and neutral (PVP) particles are less toxic, albeit with reduced penetration efficiency.

Based on a critical review of available evidence, the maximum permissible concentration of UCNPs is estimated at 2 mg/mL. The following dose regimens are recommended as safe for *in vivo* applications: for mice are 10 mg/kg for short-term exposure (12 h) and 50 mg/kg for long-term exposure (120 days); for rats is 5 mg/kg for both short- and long-term exposure. This choice is based on minimal or no changes in the internal organs based on the results of histology.

UCNPs primarily accumulate in the liver and spleen following intravenous administration, with partial clearance via fecal and urinary routes. Complete elimination has been reported within 72 h for PEG-coated particles and within 30 days for PEI-coated particles. Intraperitoneal administration may reduce hepatic accumulation compared to intravenous injection. Intravenous administration of PEI-coated UCNPs led to gradual excretion from the body within 30 days and decreasing accumulation in the spleen. In contrast, intraperitoneal administration resulted in significant particle accumulation in the spleen for up to 30 days. Following intragastric administration, UCNPs were

primarily detected in the ileum and cecum, with relatively low presence in other organs. Intravenous injection of PEGylated or PAA-coated nanoparticles caused no evident histological damage or organ impairment, with particles accumulating mainly in the liver and spleen.

Across numerous studies, H&E staining of major organs (heart, liver, spleen, lungs, kidneys) from UCNP-treated animals showed no significant inflammatory responses, necrosis, or morphological abnormalities compared to controls, except for minor, reversible changes such as transient hyperplasia in the splenic white pulp.

Smaller UCNPs (below 100 nm) are preferred for tissue penetration but may pose higher reactivity risks. Particle size and shape (e.g., spherical vs. hexagonal) influence biodistribution and retention time, with smaller particles remaining in circulation longer.

Decoration of UCNPs with targeting ligands (e.g., RGD peptides, aptamers) or photodynamically active dyes has been shown to have minimal additional toxic effects, supporting their use in multimodal theranostic platforms.

In summary, UCNPs represent a promising class of luminescent nanomaterials for *in vivo* imaging and therapy, with an acceptable safety profile at established concentrations and doses. However, the heterogeneity of experimental conditions across studies – particularly variations in coating materials, administration routes, and exposure durations – underscores the need for standardized toxicity assessment protocols. Part II of this review will address the phototoxicity of UCNPs, while Part III will examine the toxicity of multimodal upconversion particles.

Disclosures

All authors declare that there is no conflict of interests in this paper.

Acknowledgements

The study was supported by a grant Russian Science Foundation No. 25-22-00144, <https://rscf.ru/project/25-22-00144/>.

References

1. S. Hong, S. Rhee, and K. O. Jung, “*In vivo* molecular and single cell imaging,” BMB Reports 55(6), 267–274 (2022).
2. M. Woźniak, A. Płoska, A. Siekierzycka, L. W. Dobrucki, L. Kalinowski, and I. T. Dobrucki, “Molecular Imaging and Nanotechnology—Emerging Tools in Diagnostics and Therapy,” International Journal of Molecular Sciences 23(5), 2658 (2022).
3. J. C. L. Chow, “Nanomaterial-Based Molecular Imaging in Cancer: Advances in Simulation and AI Integration,” Biomolecules 15(3), 444 (2025).
4. N. W. Nkune, K. Moloudi, B. P. George, and H. Abrahamse, “An update on recent advances in fluorescent materials for fluorescence molecular imaging: a review,” RSC Advances 15(28), 22267–22284 (2025).
5. Y. Tang, Y. Li, C. He, Z. Wang, W. Huang, Q. Fan, and B. Liu, “NIR-II-excited off-on-off fluorescent nanoprobe for sensitive molecular imaging *in vivo*,” Nature Communications 16(1), 278 (2025).

6. K. Lu, R. Zhang, H. Wang, C. Li, Z. Yang, K. Xu, X. Cao, N. Wang, W. Cai, J. Zeng, and M. Gao, “[PEGylated Ultrasmall Iron Oxide Nanoparticles as MRI Contrast Agents for Vascular Imaging and Real-Time Monitoring](#),” *ACS Nano* 19(3), 3519–3530 (2025).
7. W. Zhang, X. Chen, D. Ding, G. Zhang, Z. Zhu, X. Yang, M. Li, L. Liang, X. Shi, T. Wang, and R. Gao, “[Real-time *in vivo* imaging reveals specific nanoparticle target binding in a syngeneic glioma mouse model](#),” *Nanoscale* 14(15), 5678–5688 (2022).
8. Y. Yang, Y. Chen, P. Pei, Y. Fan, S. Wang, H. Zhang, D. Zhao, B.-Z. Qian, and F. Zhang, “[Fluorescence-amplified nanocrystals in the second near-infrared window for *in vivo* real-time dynamic multiplexed imaging](#),” *Nature Nanotechnology* 18(10), 1195–1204 (2023).
9. S. Liu, W. Dong, H. Gao, Z. Song, and Z. Cheng, “[Near-Infrared-II Fluorescent Probes for Analytical Applications: From *In Vitro* Detection to *In Vivo* Imaging Monitoring](#),” *Accounts of Chemical Research* 58(4), 543–554 (2025).
10. R. Luo, C. Zhang, Z. Zhang, P. Ren, Z. Xu, and Y. Liu, “[NIR-II upconversion nanomaterials for biomedical applications](#),” *Nanoscale* 17(6), 2985–3002 (2025).
11. V. C. Thipe, A. R. Karikachery, P. Çakıl kaya, U. Farooq, H. H. Genedy, N. Kaekhamloed, D.-H. Phan, R. Rezwan, G. Tezcan, E. Roger, and K. V. Katti, “[Green nanotechnology—An innovative pathway towards biocompatible and medically relevant gold nanoparticles](#),” *Journal of Drug Delivery Science and Technology* 70, 103256 (2022).
12. A. Pavlou, A. Katsouras, M. Markou, A. Nega, S. Bellou, P. Koralli, A. Schiza, N. Chalmes, V. G. Gregoriou, A. Dimitrakopoulou-Strauss, T. Fotsis, A. Avgeropoulos, M. G. Siskos, C. Murphy, and C. L. Chochos, “[Rational Design of New Biocompatible Near-Infrared Conjugated Polymer Nanoparticles for Biomedical Applications](#),” *Macromolecules* 59(3), 1429–1442 (2026).
13. C.-Y. Zhong, L. Xiao, J. Zhou, Z. Chen, Y. Chen, Z.-Q. Liu, and J. Z. Zhang, “[Two-photon photoluminescence and bio-imaging application of monodispersed perovskite-in-silica nanocrystals with high biocompatibility](#),” *Chemical Engineering Journal* 431, 134110 (2022).
14. A. Schroter and T. Hirsch, “[Control of Luminescence and Interfacial Properties as Perspective for Upconversion Nanoparticles](#),” *Small* 20(14), 2306042 (2024).
15. E. M. Mettenbrink, W. Yang, and S. Wilhelm, “[Bioimaging with Upconversion Nanoparticles](#),” *Advanced Photonics Research* 3(12), 2200098 (2022).
16. Y. Yang, Y. Xie, and F. Zhang, “[Second near-infrared window fluorescence nanoprobe for deep-tissue *in vivo* multiplexed bioimaging](#),” *Advanced Drug Delivery Reviews* 193, 114697 (2023).
17. Z. Su, Y. Zhao, Y. Lian, Y. Li, B. Lou, C.-G. Ma, B. Xu, J. Dong, and G. Bai, “[Multifunctional lanthanide-doped nanoparticles enabled fiber probe with decoupled modules for tumor therapy and real-time temperature sensing](#),” *Science China Materials* 69(1), 139–148 (2026).
18. S. Gulati, R. Yadav, V. Kumari, S. Nair, C. Gupta, and M. Aishwari, “[Nanosensors in healthcare: transforming real-time monitoring and disease management with cutting-edge nanotechnology](#),” *RSC Pharmaceutics* 2(5), 1003–1018 (2025).
19. Y. Ma, M. Song, L. Li, X. Lao, M. Wong, and J. Hao, “[Advances in upconversion luminescence nanomaterial-based biosensor for virus diagnosis](#),” *Exploration* 2(6), 20210216 (2022).
20. J. Zhao, Z. Di, and L. Li, “[Spatiotemporally Selective Molecular Imaging via Upconversion Luminescence-Controlled, DNA-Based Biosensor Technology](#),” *Angewandte Chemie International Edition* 61(32), e202204277 (2022).
21. I. Yu. Yanina, E. K. Volkova, E. A. Sagaidachnaya, Ju. G. Konyukhova, V. I. Kochubey, and V. V. Tuchin, “[Comparison of temperature sensing of the luminescent upconversion and ZnCdS nanoparticles](#),” *Proceedings of SPIE* 10506, 1050616 (2018).
22. A. Molkenova, H. E. Choi, J. M. Park, J.-H. Lee, and K. S. Kim, “[Plasmon Modulated Upconversion Biosensors](#),” *Biosensors* 13(3), 306 (2023).
23. G. De Freitas Silva, G. De Lima Fernandes, J. H. Faleiro, T. K. De Lima Rezende, H. P. Barbosa, and J. L. Ferrari, “[Biosensing Based on Upconversion Nanoparticles](#),” in *Upconversion Nanoparticles (UCNPs) for Functional Applications*, V. Kumar, I. Ayoub, H. C. Swart, and R. Sehgal (Eds.), Springer Nature Singapore, 24, 369–403 (2023).
24. Y. Li, C. Chen, F. Liu, and J. Liu, “[Engineered lanthanide-doped upconversion nanoparticles for biosensing and bioimaging application](#),” *Microchimica Acta* 189(3), 109 (2022).
25. X. Wang, C. Jiang, Z. Wang, B. E. Cohen, E. M. Chan, and G. Chen, “[Triplet-Induced Singlet Oxygen Photobleaches Near-Infrared Dye-Sensitized Upconversion Nanosystems](#),” *Nano Letters* 23(15), 7001–7007 (2023).
26. N. Singh, R. Sen Gupta, and S. Bose, “[A comprehensive review on singlet oxygen generation in nanomaterials and conjugated polymers for photodynamic therapy in the treatment of cancer](#),” *Nanoscale* 16(7), 3243–3268 (2024).
27. M. Przygoda, D. Bartusik-Aebisher, K. Dynarowicz, G. Cieřlar, A. Kawczyk-Krupka, and D. Aebisher, “[Cellular Mechanisms of Singlet Oxygen in Photodynamic Therapy](#),” *International Journal of Molecular Sciences* 24(23), 16890 (2023).

28. I. A. Khan, T. Yu, M. Yang, J. Liu, and Z. Chen, “A Systematic Review of Toxicity, Biodistribution, and Biosafety in Upconversion Nanomaterials: Critical Insights into Toxicity Mitigation Strategies and Future Directions for Safe Applications,” *BME Frontiers* 6, 0120 (2025).
29. L. Machová Urdžíková, D. Mareková, T. Vasylyshyn, P. Matouš, V. Patsula, V. Oleksa, O. Shapoval, M. Vosmanská, D. Liebl, A. Benda, V. Herynek, D. Horák, and P. Jendelová, “Toxicity of Large and Small Surface-Engineered Upconverting Nanoparticles for *In Vitro* and *In Vivo* Bioapplications,” *International Journal of Molecular Sciences* 25(10), 5294 (2024).
30. S. Märkl, F. Przybilla, R. Rachel, T. Hirsch, M. Keller, R. Witzgall, Y. Mély, and J. Wegener, “Impact of surface chemistry of upconversion nanoparticles on time-dependent cytotoxicity in non-cancerous epithelial cells,” *Scientific Reports* 14(1), 30610 (2024).
31. S. Hofer, N. Hofstätter, B. Punz, I. Hasenkopf, L. Johnson, and M. Himly, “Immunotoxicity of nanomaterials in health and disease: Current challenges and emerging approaches for identifying immune modifiers in susceptible populations,” *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* 14(6), e1804 (2022).
32. M. M. Rahman, J. Uddin, A. M. Asiri, and M. R. Rahman (Eds.), *Toxicity of Nanoparticles-Recent Advances and New Perspectives: Recent Advances and New Perspectives*, IntechOpen, London, United Kingdom (2024). eBook ISBN: 978-1-83768-605-6.
33. R. Abdul Jalil, Y. Zhang, “Biocompatibility of silica coated NaYF₄ upconversion fluorescent nanocrystals,” *Biomaterials* 29(30), 4122–4128 (2008).
34. F. Wang, D. Banerjee, Y. Liu, X. Chen, and X. Liu, “Upconversion nanoparticles in biological labeling, imaging, and therapy,” *Analyst* 135(8), 1839 (2010).
35. M. Zhou, X. Ge, D.-M. Ke, H. Tang, J.-Z. Zhang, M. Calvaresi, B. Gao, L. Sun, Q. Su, and H. Wang, “The Bioavailability, Biodistribution, and Toxic Effects of Silica-Coated Upconversion Nanoparticles *in vivo*,” *Frontiers in Chemistry* 7, 218 (2019).
36. D. K. Chatterjee, A. J. Rufaihah, and Y. Zhang, “Upconversion fluorescence imaging of cells and small animals using lanthanide doped nanocrystals,” *Biomaterials* 29(7), 937–943 (2008).
37. J. Yu, W. Yin, T. Peng, Y. Chang, Y. Zu, J. Li, X. He, X. Ma, Z. Gu, and Y. Zhao, “Biodistribution, excretion, and toxicity of polyethyleneimine modified NaYF₄:Yb,Er upconversion nanoparticles in mice via different administration routes,” *Nanoscale* 9(13), 4497–4507 (2017).
38. N. Long, W.-T. Wong (Eds.), *The Chemistry of Molecular Imaging*, John Wiley & Sons (2014). ISBN: 978-1-118-09327-6
39. J. F. Jin, Y. J. Gu, C. W. Y. Man, J. P. Cheng, Z. H. Xu, Y. Zhang, H. S. Wang, V. H. Y. Lee, S. H. Cheng, and W. T. Wong, “Polymer-coated NaYF₄:Yb³⁺, Er³⁺ upconversion nanoparticles for charge-dependent cellular imaging,” *ACS Nano* 5(10), 7838–7847 (2011).
40. L. Xiong, T. Yang, Y. Yang, C. Xu, and F. Li, “Long-term *in vivo* biodistribution imaging and toxicity of polyacrylic acid-coated upconversion nanophosphors,” *Biomaterials* 31(27), 7078–7085 (2010).
41. X. Sun, J. Shi, X. Fu, Y. Yang, and H. Zhang, “Long-term *in vivo* biodistribution and toxicity study of functionalized near-infrared persistent luminescence nanoparticles,” *Scientific Reports* 8(1), 10595 (2018).
42. L. Cheng, K. Yang, M. Shao, X. Lu, and Z. Liu, “*In Vivo* Pharmacokinetics, Long-Term Biodistribution and Toxicology Study of Functionalized Upconversion Nanoparticles in Mice,” *Nanomedicine* 6(8), 1327–1340 (2011).
43. J. Li, X. Chang, X. Chen, Z. Gu, F. Zhao, Z. Chai, and Y. Zhao, “Toxicity of inorganic nanomaterials in biomedical imaging,” *Biotechnology Advances* 32(4), 727–743 (2014).
44. M. Z. Ahmad, B. A. Abdel-Wahab, A. Alam, S. Zafar, J. Ahmad, F. J. Ahmad, P. Midoux, C. Pichon, and S. Akhter, “Toxicity of Inorganic Nanoparticles Used in Targeted Drug Delivery and Other Biomedical Application: An Updated Account on Concern of Biomedical Nanotoxicology,” *Journal of Nanoscience and Nanotechnology* 16(8), 7873–7897 (2016).
45. A. Gnach, T. Lipinski, A. Bednarkiewicz, J. Rybka, and J. A. Capobianco, “Upconverting nanoparticles: assessing the toxicity,” *Chemical Society Reviews* 44(6), 1561–1584 (2015).
46. C. Dai, S. Chen, C. Wang, L. Zhang, K. Ge, and J. Zhang, “Ytterbium ion promotes apoptosis of primary mouse bone marrow stromal cells?,” *Journal of Rare Earths* 33(4), 445–452 (2015).
47. K. A. Gschneidner, J.-C. G. Bünzli, and V. K. Pecharsky, “Preface,” in *Handbook on the Physics and Chemistry of Rare Earths*, K. A. Gschneidner, J.-C. G. Bünzli, and V. K. Pecharsky (Eds.), 37, Elsevier, v–x (2007).
48. B. Tian, Q. Wang, Q. Su, W. Feng, and F. Li, “*In vivo* biodistribution and toxicity assessment of triplet-triplet annihilation-based upconversion nanocapsules,” *Biomaterials* 112, 10–19 (2017).
49. R. Yang, L. Cai, A. Zheng, and J. Guan, “Inhalation of upconversion nanoparticles does transient damage to the mouse liver and lung,” *Chinese Journal of Tissue Engineering Research* 20 (52), 7865–7871 (2016).
50. S. Wang, H. Liang, Z. Yang, Z. Wang, B. Yang, and C. Lu, “Direct large-scale synthesis of water-soluble and biocompatible upconversion nanoparticles for *in vivo* imaging,” *RSC Advances* 14(25), 17350–17354 (2024).
51. B. Fubini, A. Hubbard, “Reactive oxygen species (ROS) and reactive nitrogen species (RNS) generation by silica in inflammation and fibrosis,” *Free Radical Biology and Medicine* 34(12), 1507–1516 (2003).
52. D. B. Warheit, “Nanoparticles: health impacts?,” *Materials Today* 7(2), 32–35 (2004).

53. V. Bastos, P. Oskoei, E. Andresen, M. I. Saleh, B. Rühle, U. Resch-Genger, and H. Oliveira, “Stability, dissolution, and cytotoxicity of NaYF₄-upconversion nanoparticles with different coatings,” *Scientific Reports* 12(1), 3770 (2022).
54. F. Chen, W. Bu, W. Cai, and J. Shi, “Functionalized Upconversion Nanoparticles: Versatile Nanoplatforms for Translational Research,” *Current Molecular Medicine* 13(10), 1613–1632 (2013).
55. Z. Yu, Y. Xia, J. Xing, Z. Li, J. Zhen, Y. Jin, Y. Tian, C. Liu, Z. Jiang, J. Li, and A. Wu, “Y1 -receptor–ligand-functionalized ultrasmall upconversion nanoparticles for tumor-targeted trimodality imaging and photodynamic therapy with low toxicity,” *Nanoscale* 10(36), 17038–17052 (2018).
56. L. Sun, Z. Wei, H. Chen, J. Liu, J. Guo, M. Cao, T. Wen, and L. Shi, “Folic acid-functionalized up-conversion nanoparticles: toxicity studies *in vivo* and *in vitro* and targeted imaging applications,” *Nanoscale* 6(15), 8878–8883 (2014).
57. J. Zhang, F. Liu, T. Li, X. He, and Z. Wang, “Surface charge effect on the cellular interaction and cytotoxicity of NaYF₄:Yb³⁺, Er³⁺@SiO₂ nanoparticles,” *RSC Advances* 5(10), 7773–7780 (2015).
58. A. E. Guller, A. N. Generalova, E. V. Petersen, A. V. Nechaev, I. A. Trusova, N. N. Landyshev, A. Nadort, E. A. Grebenik, S. M. Deyev, A. B. Shekhter, and A. V. Zvyagin, “Cytotoxicity and non-specific cellular uptake of bare and surface-modified upconversion nanoparticles in human skin cells,” *Nano Research* 8(5), 1546–1562 (2015).
59. S. Maiolino, A. Russo, V. Pagliara, C. Conte, F. Ungaro, G. Russo, and F. Quaglia, “Biodegradable nanoparticles sequentially decorated with Polyethyleneimine and Hyaluronan for the targeted delivery of docetaxel to airway cancer cells,” *Journal of Nanobiotechnology* 13(1), 29 (2015).
60. H. Yang, Y. Li, T. Li, M. Xu, Y. Chen, C. Wu, X. Dang, and Y. Liu, “Multifunctional Core/Shell Nanoparticles Cross-linked Polyetherimide-folic Acid as Efficient Notch-1 siRNA Carrier for Targeted Killing of Breast Cancer,” *Scientific Reports* 4(1), 7072 (2014).
61. X. Luo, Q. Chen, H. Guo, H. Zhang, X. He, and W. Zhao, “One-step hydrothermal synthesis of Cit-NaYbF₄:Er³⁺ nanocrystals with enhanced red upconversion emission for *in vivo* fluorescence molecular tomography,” *Journal of Rare Earths* 42(1), 36–45 (2024).
62. S. Yan, G. Xing, X. Yuan, E. Cui, K. Ji, X. Yang, J. Su, D. Mara, J. Tang, Y. Zhao, J. Hu, and J. Liu, “Upconversion nanoparticles-CuMnO₂ nanoassemblies for NIR-excited imaging of reactive oxygen species *in vivo*,” *Journal of Colloid and Interface Science* 677, 666–674 (2025).
63. Q.-L. Wang, L.-C. Meng, Z. Zhao, J.-F. Du, P. Li, Y. Jiang, and H.-J. Li, “Ultrasensitive upconverting nanoprobe for *in situ* imaging of drug-induced liver injury using miR-122 as the biomarker,” *Talanta* 274, 126108 (2024).
64. T. Vasylyshyn, V. Patsula, D. Větvicka, O. Shapoval, J. Pankrác, M. Kabešová, J. Beneš, and D. Horák, “Intraperitoneal versus intravenous administration of Flamma®-conjugated PEG-alendronate-coated upconversion nanoparticles in a mouse pancreatic cancer model,” *Nanoscale Advances* 7(1), 144–154 (2025).
65. T. Maldiney, C. Richard, J. Seguin, N. Wattier, M. Bessodes, and D. Scherman, “Effect of Core Diameter, Surface Coating, and PEG Chain Length on the Biodistribution of Persistent Luminescence Nanoparticles in Mice,” *ACS Nano* 5(2), 854–862 (2011).
66. X. Li, M. Zhang, H. Zhang, Z. Wang, and H. Zhang, “Upconversion nanoparticle-based fluorescence resonance energy transfer sensing platform for the detection of cathepsin B activity *in vitro* and *in vivo*,” *Microchimica Acta* 190(5), 181 (2023).
67. B. Yadav, M. Chauhan, S. Shekhar, A. Kumar, A. K. Mehata, A. K. Nayak, R. Dutt, V. Garg, V. Kailashiya, M. S. Muthu, Sonali, and R. P. Singh, “RGD-decorated PLGA nanoparticles improved effectiveness and safety of cisplatin for lung cancer therapy,” *International Journal of Pharmaceutics* 633, 122587 (2023).
68. B. Ateş, A. Ulu, M. Asiltürk, S. A. A. Noma, S. D. Topel, G. Dik, O. Özhan, B. Bakar, A. Yıldız, N. Vardı, and H. Parlakpınar, “Enhancement of enzyme activity by laser-induced energy propulsion of upconverting nanoparticles under near-infrared light: A comprehensive methodology for *in vitro* and *in vivo* applications,” *International Journal of Biological Macromolecules* 260, 129343 (2024).
69. Y. Bai, Y. Li, R. Wang, and Y. Li, “Low Toxicity, High Resolution, and Red Tissue Imaging in the Vivo of Yb/Tm/GZO@SiO₂ Core–Shell Upconversion Nanoparticles,” *ACS Omega* 5(10), 5346–5355 (2020).
70. H. Zhao, W. Hu, H. Ma, R. Jiang, Y. Tang, Y. Ji, X. Lu, B. Hou, W. Deng, W. Huang, and Q. Fan, “Photo-Induced Charge-Variable Conjugated Polyelectrolyte Brushes Encapsulating Upconversion Nanoparticles for Promoted siRNA Release and Collaborative Photodynamic Therapy under NIR Light Irradiation,” *Advanced Functional Materials* 27(44), 1702592 (2017).
71. Y. Zhao, D. Sultan, and Y. Liu, “Biodistribution, Excretion, and Toxicity of Nanoparticles,” Chapter 2 in *Theranostic Bionanomaterials*, W. Cui, X. Zhao (Eds.), Elsevier, 27–53 (2019).
72. N. Ghafoor, K. Javed, U. E. Ummara, N. Anees, T. Mehar, U. Khursheed, R. Shahid, and A. Ghafoor, “Nanoparticles in diagnostic imaging: advancements and challenges,” in *Complementary and Alternative Medicine: Nanotechnology I*, V. G. G. Rubio, A. Khan, S. Altaf, Z. Saeed, and W. Qamar (Eds.), Unique Scientific Publishers, Faisalabad, Pakistan, 276–284 (2024).

73. J. Zhou, Y. Sun, X. Du, L. Xiong, H. Hu, and F. Li, “[Dual-modality *in vivo* imaging using rare-earth nanocrystals with near-infrared to near-infrared \(NIR-to-NIR\) upconversion luminescence and magnetic resonance properties](#),” *Biomaterials* 31(12), 3287–3295 (2010).

In Vivo Toxicity of Upconversion Nanoparticles. Overview. Part II

Irina Yu. Yanina* and Vyacheslav I. Kochubey

Institute of Physics, Saratov State University, 83 Astrakhanskaya str., Saratov 410012, Russian Federation

*e-mail: irina-yanina@list.ru

Abstract. This is the second section of review synthesizes current research on the *in vivo* toxicity of upconversion nanoparticles (UCNPs) for biomedical applications, including photoinduced drug delivery, photodynamic therapy (PDT), and multimodal therapies. The aim is to assess the safety of UCNPs under near-infrared (NIR) irradiation and in combination with photosensitizers, chemotherapeutics, and other therapeutic modalities. Analysis of numerous *in vivo* studies in mouse and rat models demonstrates that UCNPs exhibit low dark toxicity and targeted phototoxicity, with no significant damage to major organs (heart, liver, spleen, lungs, kidneys) at therapeutic doses. For mice, safe doses are 10 mg/kg for short-term exposure (48 h and for long-term exposure 40 days). Surface coatings (PEG, silica, MVEMA, cell membranes) enhance biocompatibility and reduce off-target effects. Multimodal UCNP platforms (PDT combined with chemotherapy, photothermal therapy, chemodynamic therapy, immunotherapy, or radiotherapy) demonstrate effective tumor suppression without substantial systemic toxicity. These findings support the continued development of UCNPs as safe and effective theranostic agents for precision oncology.

Keywords: phototoxicity, upconversion nanoparticles, dose, photodynamic therapy.

Paper #9596 received 17 Apr 2026; revised manuscript received 21 May 2026; accepted for publication 26 May 2026; published online 28 Sep 2026. doi: [10.18287/JBPE26.12.030202](https://doi.org/10.18287/JBPE26.12.030202).

1 Phototoxicity of Upconversion Nanoparticles

1.1 Drug Delivery and Photo Induced Release

One of the potential applications of upconversion nanoparticles (UCNPs) combined with a photosensitizer (PS) is targeted drug delivery and photo induced drug release. For instance, in recent paper, UCNPs coated with a PS, a stopper agent, and modified mesoporous silica (designated as UCNPs@mSiO[DEX] Py/β CD/FITC, or USDPFs) were developed for synergistic combination therapy of lung ischemia reperfusion injury (LIRI), enabling precise release of dexamethasone (DEX) [1].

DEX is a synthetic glucocorticoid that helps reduce inflammatory responses, suppress the immune response, and improve the tolerability of anticancer therapy.

UCNPs were developed by coating an inert YOF:Yb shell onto a YOF:Yb, Tm core to achieve high intensity blue and red upconversion emission upon irradiation with a NIR laser.

To evaluate the *in vivo* safety of USDPFs, healthy adult C57BL/6 mice were administered 5 mg kg⁻¹ of USDPFs via intratracheal instillation. After two weeks, mice in the USDPF group showed no infiltration of inflammatory cells or tissue necrosis in the heart, lungs, liver, spleen, and kidneys, as compared to the control group. This indicates that intratracheal administration of the nanoparticles does not damage lung tissue, and that USDPFs do not exhibit significant toxicity to several vital organs in mice, demonstrating good *in vivo* safety.

To evaluate the safety and function of USDPFs (UCNPs for DEX delivery), the authors investigated several conditions in mice: USDPFs alone, USDPFs with NIR irradiation (980 nm, 3 × 10 min, 1 W/cm²),

ischemia/reperfusion (I/R) injury, I/R + USDPFs, and I/R + USDPFs + NIR [1–4]. Mice received a single dose of 5 mg/kg USDPFs via intratracheal or intravenous administration.

Following intravenous injection, USDPFs remained detectable in the bloodstream for approximately 10 hours, confirming prolonged circulation compatible with the enhanced permeability and retention (EPR) effect. Importantly, neither USDPFs alone nor USDPFs + NIR caused any lung tissue damage, edema, or inflammatory cell infiltration compared to healthy controls. In contrast, the I/R and I/R + USDPFs groups showed severe lung injury, characterized by thickened alveolar walls, edema, and elevated levels of IL-1 β , IL-6, and TNF α .

Remarkably, the I/R + USDPFs + NIR group exhibited significantly reduced lung damage, pulmonary edema (lower wet-to-dry weight ratio), and lower inflammatory cytokine levels compared to both I/R-only and I/R + USDPFs groups. These results demonstrate that NIR irradiation triggers DEX release from USDPFs, effectively alleviating inflammation and edema without intrinsic nanoparticle toxicity. Overall, USDPFs show a favorable biosafety profile.

Complex heterostructures comprising UCNPs (DUCNPs) and manganese based metal–organic frameworks (Mn MOFs) can be used for drug delivery [5].

The authors developed a nanocarrier designated DUCNP@Mn MOF, which demonstrated good loading efficiency and delivery capability for the cytotoxic antitumor agent 3 F 10 OH evodiamine (FOE).

The combined properties of Mn MOF and DUCNPs endowed the DUCNP@Mn MOF/FOE system with synergistic chemodynamic and chemotherapeutic effects. The developed nanosystem exhibited high tumor targeting selectivity.

In a mouse model of breast cancer, the nanocarrier effectively suppressed tumor growth without causing significant toxicity.

Wang Y., et al. developed NaYF₄:Yb,Tm-based upconversion nanoparticles conjugated with the chemotherapeutic agent 9-aminocamptothecin (ACPT) via photolabile linkers. The nanoparticles were coated with a 4T1 tumor cell membrane functionalized with melittin to enable homologous tumor targeting and enhance drug penetration [6].

In a 4T1 breast cancer mouse model, intravenous injections were administered on days 0, 3, and 6 (melittin dose: 1 mg kg⁻¹), followed by 808 nm NIR irradiation (0.8 W cm⁻², 20 min) on days 2, 5, and 8. The treatment inhibited tumor growth and suppressed lung metastasis. No obvious pathological changes were observed in the heart, liver, spleen, lungs, or kidneys.

Serum inflammatory factors such as tumor necrosis factor α (TNF α) and interferon γ (IFN γ), which are secreted by immune cells, can inhibit tumor cell growth [7].

Therefore, the concentrations of TNF α and IFN γ were analyzed. The results showed that treatment with U

ACPT@MM + NIR significantly increased the serum levels of both TNF α and IFN γ .

In addition, no obvious pathological changes were observed in the tissues of the heart, liver, spleen, lungs, and kidneys.

1.2 Photodynamic Therapy

In photodynamic therapy (PDT), the use of UCNPs involves loading them with PSs and exciting the PSs via the luminescence of the nanoparticles. Since the excitation of UCNPs is achieved using infrared light – which penetrates deeply into biological tissues – this approach enables effective treatment of deeper seated tumors without causing photodamage to the overlying tissues.

When excited, the PS undergoes specific reactions, initiating a chain of processes that lead to the formation of various reactive oxygen species (ROS). These ROS exert a damaging effect on cellular structures, ultimately causing the death of tumor cells.

More than a hundred substances with photosensitizing properties are known. Among natural compounds, PSs include chlorophylls, phycobilins, porphyrins and their synthetic intermediates, several antibiotics, quinine, riboflavin, and others.

Most commonly in PDT, several types of PS are used – mainly various derivatives of porphyrins. In addition to porphyrins themselves, related derivatives are also employed: chlorin based compounds (chlorins, bacteriochlorins), as well as phthalocyanines and naphthalocyanines [8].

Besides the PS, the nanoparticle surface must be coated with functional groups to ensure targeted accumulation of nanoparticles within the tumor.

Analysis of nanoparticle use in PDT should also include investigation of pathological morphological changes in major organs (liver, heart, kidneys, spleen, and lungs), both under nanoparticle irradiation and without it. The absence of such changes indicates no dark toxicity and no secondary effects arising from irradiation when using the studied nanoparticles.

However, nanoparticles can also be used without sensitizers.

The authors synthesized NaGdF₄:25 % Yb³⁺, 1 % Nd³⁺, 0.5 % Tm³⁺ nanoparticles, whose luminescence could be excited both at 808 nm (Nd³⁺ ions) and at 980 nm (Yb³⁺ ions) [9]. The emission of these nanoparticles in the 450–500 nm range (Tm³⁺ ions) is effectively absorbed by all organic PSs. When using nanoparticles coated with thiochromenocarbazole imide (TCI), the authors noted low dark toxicity and the possibility of *in vivo* imaging. However, the authors did not observe a significant effect on the growth of HeLa stimulated tumors.

In the work, UCNPs of the composition TiO₂/(Ca, Y)F₂:Tm, Yb were obtained by a hydrothermal method; these nanoparticles respond to near infrared light and are intended for combating tumors [10].

Additionally, a femtosecond laser (FL) was used to reduce the size and aggregation of

TiO₂/(Ca, Y)F₂:Tm, Yb, which allowed improving its dispersity. The authors evaluated the phototherapeutic effect of the nanoparticles on CFPAC 1 pancreatic tumors.

The experiments showed that in the absence of TiO₂/(Ca, Y)F₂:Tm, Yb FL, the growth rate and size of tumors in nude mice were significantly larger than in the treatment case. This indicates that TiO₂/(Ca, Y)F₂:Tm, Yb FL nanoparticles play a certain role in inhibiting tumor growth under near infrared light, but due to the complex and variable environment *in vivo*, the hydroxyl radicals released from TiO₂/(Ca, Y)F₂:Tm, Yb FL did not achieve the effect of destroying tumor cells over a large area. Instead, they only led to inhibition of tumor growth in nude mice without reducing the tumor area or killing all tumor cells.

This may be associated with the hypoxic microenvironment of tumor tissue, which causes tumor cells to activate signaling pathways leading to increased synthesis or expression of certain genes and proteins (oxygen regulated protein (ORP), vascular endothelial growth factor (VEGF), erythropoietin (EPO), etc.), resulting in structural and functional abnormalities in the blood vessels of tumor tissue. Consequently, their sensitivity to the oxygen concentration in the surrounding environment is lower than that of normal tissues [11].

To enhance nanoparticle penetration into tumor tissue, boost PDT efficacy, and reduce off target toxicity, special surface coatings are required.

In their study, Ling B., et al. developed nucleus penetrating UCNPs (UCNP@Glu DMMA) with charge reversal properties [12].

The design relies on two key components:

- glucose, which confers tumor cell tropism;
- 2,3 dimethylmaleic anhydride (DMMA), which imparts a negative surface charge to the nanoparticles.

This configuration ensures blood circulation stability and accumulation in the tumor microenvironment (TME). DMMA and its derivatives allow the nanoparticle surface charge to be modulated based on the tumor's acidity. Exposure to the acidic TME and γ glutamyl transpeptidase (GGT) triggers a charge reversal from negative to positive.

Positively charged nanoparticles are more readily taken up by cells, thereby amplifying the PDT effect. Direct delivery to the cell nucleus enables localized release of ROS, leading to sustained photo stimulated DNA damage.

In the experimental phase, UCNPs@Glu DMMA was injected into nude mice bearing HeLa tumors. The animals then underwent laser irradiation at 980 nm for 15 min (30 s of irradiation followed by a 1 min break to prevent overheating) at a power density of 0.8 W cm⁻².

The results showed a significant reduction in tumor size. Irradiation induced substantial apoptosis in tumor tissues and effectively suppressed cellular proliferation. Toxicity assessment revealed minimal signs of tissue damage or inflammation in major organs.

Currently, a wide range of phototherapeutic agents based on PS conjugated UCNPs has been tested for NIR

induced PDT, primarily aiming to demonstrate cellular uptake and cell death following NIR irradiation [13–15].

The first *in vivo* application was performed by intratumoural injection of NaYF₄:Yb³⁺, Er³⁺@poly(ethylene glycol) particles non covalently loaded with chlorin e6 (Ce6) into mice bearing 4T1 breast tumors [16]. PDT with 980 nm irradiation (0.5 W cm⁻², 30 min) significantly reduced tumor growth rates in mice.

A similar effect after NIR irradiation (5 min, 0.6 W cm⁻²) was observed in nude mice bearing U87MG tumors that received intravenous injections of NaYF₄:Yb³⁺, Er³⁺@NaGdF@poly(ethylene glycol) phospholipid particles loaded with Ce6 [17].

In mice bearing S180 sarcoma tumors treated with NaYF₄:Yb³⁺, Er³⁺ particles loaded with zinc phthalocyanine and chitosan – and irradiated twice (3 h and 24 h post injection) with a 980 nm laser (0.4 W cm⁻², 15 min) – only a slight increase in tumor volume was observed, along with high survival rates [18].

Injection of NaYF₄:Yb³⁺, Er³⁺@CaF₄@poly(acrylic acid)hydrazide particles covalently functionalized with 5 aminolevulinic acid into mice significantly reduced tumor size after 980 nm laser irradiation (0.5 W cm⁻²) for 40 min [19].

An *in vivo* PDT study in mice with CT 26wt tumors showed reduced tumor growth following intratumoural injection of NaYF₄:Yb³⁺, Er³⁺@phosphatidylcholine folic acid (FA) particles loaded with tetraphenylporphyrin and irradiation at 980 nm (1 W cm⁻², 30 min) [20].

Similarly, many organic PSs – such as methylene blue [21], Rose Bengal [22], and pyropheophorbide a methyl ester [23] – when integrated with UCNPs for NIR activated PDT (0.15–1 W cm⁻²; up to 40 min), significantly suppressed tumour growth.

The use of UCNPs with protoporphyrin (PPIX) for PDT was investigated in paper [24]. Abundant amino groups of a branched PEI with a spatial structure were employed to enhance the loading capacity of protoporphyrin (PPIX), which ultimately improved the rate of skin tumor removal.

The authors achieved specific tumour treatment through the use of FA targeting and active enrichment of PPIX. When nanoparticle solutions were injected intravenously over a period of 4 weeks, no significant pathological morphological changes were observed in major organs (liver, heart, kidneys, spleen, and lungs), both with and without nanoparticle irradiation. This indicates good biosafety of the treatment method.

In the group treated with nanoparticles without irradiation, the tumor cell structure remained ordered, and the nuclear shape was complete and well defined. In contrast, after irradiation, the tumor tissue was severely damaged, and the nucleus was largely destroyed or even disappeared. These results demonstrate effective anticancer photoactivity while maintaining safety for other organs.

One of the widely used PSs is a porphyrin analogue – chlorins (porphyrins hydrogenated at

C17 = C18). Among chlorins, chlorin e6 is one of the most commonly used in PDT.

Chlorin e6 selectively accumulates in pathological cells. Upon local excitation with light at a wavelength of 660–670 nm, chlorin e6 becomes activated as a PS within pathological cells, initiating a photochemical reaction that produces singlet oxygen and other highly reactive free radicals, which induce apoptosis.

Since tumor development often follows a hypoxic mechanism, an additional requirement for the effective use of chlorin e6 is ensuring sufficient oxygen supply.

To investigate the impact of hypoxia and achieve an antitumor effect by switching the phenotype of tumor associated macrophages (TAMs) from the tumor promoting M2 type to the tumor suppressing M1 type, the authors fabricated $\text{NaErF}_4/\text{NaLuF}_4$ nanoparticles decorated with mannose and coated with a macrophage membrane (RAW), covered by a silica dioxide layer doped with an oxygen carrier – perfluorocarbon (PFC) – and chlorin e6 (Ce6), and loaded with the antitumor drug paclitaxel (PTX) ($\text{UCNP@mSiO}_2 \text{ PFC/Ce6@RAW Man/PTX}$; ~61 nm; 11.6 mV) [25].

The primary UCNPs consisted of lanthanide elements (erbium and lutetium) in a core@shell structure and emitted light at 660 nm in response to near infrared laser irradiation at 808 nm. These nanoparticles had the following key properties:

- efficient generation of singlet oxygen, enabled by oxygen supply from the PFC component;
- high targeting specificity for tumor associated macrophages (TAMs), predominantly of the M2 phenotype;
- capacity to induce M1 type macrophage polarization.
- M1 macrophages release pro inflammatory cytokines and suppress breast cancer growth. The design of these nanoparticles thus combined:
 - oxygen delivery (via PFC) to counteract tumor hypoxia;
 - PS activity (Ce6) for PDT under NIR irradiation;
 - macrophage membrane coating (RAW) for improved biocompatibility and targeting;
 - mannose decoration for enhanced recognition by TAMs;
 - chemotherapy (PTX) for synergistic antitumor action.

This multifunctional approach aimed to reprogram TAMs from a tumor supportive (M2) to a tumor suppressive (M1) phenotype, thereby enhancing the overall antitumor immune response in hypoxic breast cancer.

4T1 cells ($100 \mu\text{L}$, 2×10^6 cells) were subcutaneously injected into the flanks of mice. When tumor volume reached approximately 150 mm^3 , the mice were randomly divided into five groups: G1 – phosphate buffer solution (PBS) (no treatment); G2 – $\text{UCNPs@mSiO}_2 \text{ PFC/Ce6}$; G3 – $\text{UCNPs@mSiO}_2 \text{ PFC/Ce6} + 808 \text{ nm}$ irradiation (1.5 W cm^{-2} for 20 min); G4 – $\text{UCNPs@mSiO}_2 \text{ PFC/Ce6@RAW}$

Man/PTX ; G5 – $\text{UCNPs@mSiO}_2 \text{ PFC/Ce6@RAW Man/PTX} + 808 \text{ nm}$ irradiation.

The nanoparticle dose was set at 6.5 mg kg^{-1} ($130 \mu\text{g}$ total nanoparticle amount), with 0.4, 0.6, and 0.5 mg kg^{-1} for PFC, Ce6, and PTX, respectively. Laser irradiation at 808 nm (1.5 W cm^{-2}) was applied to each mouse's tumor for 20 min.

Both PBS and hypoxic $\text{UCNPs@mSiO}_2 \text{ PFC/Ce6@RAW Man/PTX}$ (without PFC doping and purged with N_2) generated very little oxygen (total sO_2 : 19–20%), even 12 and 24 h after intravenous administration. However, oxygenated $\text{UCNPs@mSiO}_2 \text{ PFC/Ce6@RAW Man/PTX}$, with ultrasound stimulated oxygenation, significantly increased tumor oxygen levels (total $\text{sO}_2 > 40\%$) 24 h after intravenous injection.

These results indicate that intravenous administration of the nanoparticles successfully delivers a substantial amount of oxygen, effectively alleviating hypoxia in solid tumors.

Tumors in G5 mice (treated with $\text{UCNPs@mSiO}_2 \text{ PFC/Ce6@RAW Man/PTX}$ plus 808 nm laser irradiation) were the most significantly suppressed. This is presumably due to the potent antitumor effect arising from: polarization of TAMs into tumor suppressive M1 type macrophages; ROS generation; the combined action of paclitaxel (PTX).

The antitumor effect can also be achieved via intratumoural injection of microgels containing UCNPs [26].

Hb μGels consist of haemoglobin (Hb), bovine serum albumin (BSA), chlorin e6 (Ce6), and erbium@lutetium doped UCNPs (~35 nm), which efficiently convert 808 nm irradiation into 660 nm visible light.

Tumors were established by subcutaneously injecting 4T1 cells ($100 \mu\text{L}$, 2×10^6 cells) into the flanks of mice.

As an optical probe for hypoxia, the authors used pimonidazole (2 nitroimidazole), which is specifically reduced and binds to intracellular macromolecules in hypoxic regions of tumors [27, 28].

The authors demonstrated that laser irradiation at 808 nm alone, or the use of nanoparticles without Ce6, did not produce significant hyperthermic or photodynamic effects on 4T1 tumor xenografts.

Cytotoxicity was achieved through ROS mediated apoptosis, as well as through intense and sudden thermal stress induced by laser heating and additional heat delivery to temperatures exceeding 48°C . Enhanced tumor oxygenation was accomplished by:

- oxygen saturation of the formulations;
- the use of Hb $\mu\text{Gels}(+\text{UCNPs/Ce6})$ nanoparticles loaded with ascorbic acid;
- ultrasound stimulated oxygenation.

These results indicate that Hb(6) μGels can deliver oxygen to hypoxic tumor tissues, thereby helping to suppress hypoxic tumor growth.

Moreover, the use of these gels resulted not only in minimal accumulation in the five major tissue organs (liver, kidneys, spleen, lungs, and heart), but also in negligible changes to tissue histology. This demonstrates that the hemoglobin based μGel incorporating UCNPs is

biocompatible and safe, making it a promising injectable microgel system for therapeutic applications.

To further alleviate tumor hypoxia, biocompatible and environmentally friendly zinc oxide nanoparticles (ZnO NPs) with upconversion properties and catalase like nanozyme activity were developed [29].

As nanozymes, the ZnO NPs can catalyse the decomposition of hydrogen peroxide (H_2O_2) within tumours, thereby generating oxygen to support photodynamic action and relieve hypoxia.

In experiments with 4T1 tumors, growth was significantly suppressed after implantation of:

- Ce6 nanoparticles irradiated with a 660 nm laser;
- ZnO Ce6 nanoparticles irradiated with an 808 nm laser.

However, the antitumour activity of 808 nm laser irradiated ZnO Ce6 NPs was notably higher than that of 660 nm laser irradiated Ce6 NPs.

Remarkably, tumors were completely eradicated in three out of five mice bearing 4T1 tumors following treatment with 808 nm laser irradiated ZnO Ce6 NPs.

The corresponding tumor growth suppression rates were:

- 90.0% for ZnO Ce6 NPs with 808 nm laser irradiation;
- 80.2% for Ce6 NPs with 660 nm laser irradiation – compared to the control group.

These findings highlight the potential of ZnO NPs with upconversion and nanozyme properties as a multifunctional platform for overcoming tumor hypoxia and enhancing photodynamic therapy efficacy.

UCNPs are also used in combination with the antitumor drug doxorubicin (DOX) [30, 31]. For effective phototargeting, FA was employed as a photoactive targeting molecule. In chemotherapy, the anticancer drug DOX was thiolated on the surface of UCNPs, forming a disulfide bond that can be cleaved by lysosomal enzymes inside cells [31].

A study demonstrated that long term *in vivo* toxicity tests in mice receiving intravenous injections of 10 mg/kg $\text{BaGdF}_5:\text{Yb}^{3+}/\text{Tm}^{3+}$ with DOX did not reveal significant adverse health effects, and survival was observed for 40 days [30].

One derivative of chlorin is meso tetrahydroxyphenylchlorin (Temoporfin, THPC, m THPC), also known as Foscan. It is a photosensitising agent used in photodynamic therapy for the treatment of head and neck squamous cell carcinoma. However, there are several studies exploring the use of UCNPs with THPC for treating other types of cancer.

The efficacy of UCNP@PMVEMA THPC 2 conjugates for PDT following NIR light irradiation at 980 nm was tested *in vivo* in a study using human pancreatic adenocarcinoma (Capan 2) grown subcutaneously in athymic nude mice.

One of the key requirements for the particles is surface modification – for instance, using poly(methyl vinyl ether alt maleic acid) (PMVEMA) and the PS THPC – to ensure their colloidal and chemical stability

in aqueous media and to enable singlet oxygen generation upon NIR irradiation, respectively [32].

A coating of alendronate terminated poly[N,N dimethylacrylamide co 2 aminoethyl acrylamide] graft poly(ethylene glycol) [P(DMA AEM)PEG Ale] was also used to ensure chemical and colloidal stability of the particles in aqueous physiological fluids, which additionally enhances therapeutic efficacy [33].

Co doping of Fe^{2+} , Yb^{3+} , and Er^{3+} ions into the NaYF_4 host triggered upconversion emission in the red region, which is dominant for directly exciting THPC. It was found that such monodisperse UCNPs – $\text{NaYF}_4:\text{Yb}^{3+}, \text{Er}^{3+}, \text{Fe}^{2+}$ coated with PMVEMA and chemically linked to THPC – efficiently transfer energy and generate singlet oxygen. The authors investigated the effect of these particles on Capan 2 pancreatic adenocarcinoma cells.

The mice were administered intratumourally with PBS (control group), UCNP@PMVEMA particles, or UCNP@PMVEMA THPC 2 particles, followed by laser irradiation at a wavelength of 980 nm for 3 min at a power density of 0.5 W cm^{-2} (except for the negative control group).

It should be noted that a power density of $< 0.7 \text{ W cm}^{-2}$ is the conservative limit established for safe exposure of human skin to 980 nm light [34].

No significant differences in body weight were observed between the groups, indicating the absence of adverse effects during the treatment period. The same was observed for UCNPs containing chlorin e6 [35].

On the first or second day after irradiation, extensive necrosis developed in all mice in the group treated with UCNP@PMVEMA THPC 2 particles, whereas no necrosis was observed in the control group or in the group containing UCNP@PMVEMA.

However, the UCNP@PMVEMA THPC 2 group exhibited a statistically significant suppression of tumor growth rate, and the tumor volume on day 30 was 0.35 cm^3 .

Thus, the potential of THPC conjugated UCNPs coated with PMVEMA in the treatment of pancreatic adenocarcinoma – with sufficiently low toxicity – has been confirmed.

The use of THPC and UCNPs *in vivo* for glioblastoma treatment has also been investigated [36].

Encapsulation of the dyes THPC and IR 780 into $\text{NaYF}_4:\text{Yb}^{3+}, \text{Er}^{3+}$ @PEG particles enabled a promising approach to glioblastoma therapy under combined laser irradiation at 980 nm (0.8 W cm^{-2}) and 808 nm (0.36 W cm^{-2}) for 5 min.

In all studies involving THPC, no nanoparticle toxicity or tumor changes were observed in the absence of laser irradiation.

Phthalocyanines are actively used in diagnostics and PDT.

In study an energy segmentation strategy was employed to synthesise core multi shell UCNPs capable of emitting three distinct colors (red, green, and blue) when irradiated with different NIR lights (1550 nm, 808 nm, and 980 nm) [37]. By combining these UCNPs

with PSs and nitric oxide (NO) donors, a “smart” PDT nanopatform was developed.

4T1 tumor cells were subcutaneously injected into mice. Although the internalization of UCNPs in living organisms depends on their size and surface ligands, UCNPs within the 50–250 nm size range are internalized via endocytosis.

The authors investigated PDT using laser irradiation at various wavelengths. When tumors reached approximately 100 mm³, the mice were randomly divided into eight groups (five mice per group):

- Control Group 1: PBS solution;
- Control Group 2: UCNPs@mSiO₂ ZnPc RBS;
- Group 3: UCNPs@mSiO₂ ZnPc RBS with 808 nm irradiation;
- Group 4: UCNPs@mSiO₂ ZnPc RBS with 980 nm irradiation;
- Group 5: UCNPs@mSiO₂ ZnPc RBS with 1550 nm irradiation;
- Group 6: UCNPs@mSiO₂ ZnPc RBS with 1550 nm + 980 nm irradiation;
- Group 7: UCNPs@mSiO₂ ZnPc RBS with 1550 nm/980 nm irradiation;
- Group 8: UCNPs@mSiO₂ ZnPc RBS with 980 nm/1550 nm irradiation.

The laser power density was set at 1 W cm⁻², and the irradiation time for each group was 20 min.

Tumor growth was rapid in Control Group 1, Control Group 2, Group 3, whereas tumor growth in Groups 4, 5, 6, and 7 was somewhat slowed. In contrast, the tumor size in Group 8 was the smallest, indicating that tumor growth was significantly suppressed by programmed phototherapy – specifically, by the sequential release of NO followed by ROS.

No significant decrease or increase in body weight was observed, indicating good biocompatibility of the developed nanocomposites.

Histological analysis of tumors and major organs showed that the most pronounced apoptosis or necrosis and nuclear shrinkage occurred in the 980 nm/1550 nm group.

No obvious damage to the major organs (heart, liver, spleen, lungs, and kidneys) was detected, suggesting that the UCNPs@mSiO₂ ZnPc RBS do not exhibit short term toxicity and demonstrate great potential for clinical application.

The simultaneous action of two photosensitive agents – MC540 (merocyanine 540) and FePc (iron phthalocyanine) – was investigated in the study [38].

The authors developed UCNPs NaErF₄@NaYbF₄@NaYF₄ coated with mesoporous silica loaded with MC540 and FePc.

Nude mice were subcutaneously injected with HeLa tumor cells. When the tumor size reached approximately 100 mm³, the mice were randomly divided into five groups (three mice per group):

- Group 1: PBS injection only (control group);
- Group 2: UCNPs@mSiO₂ FePc MC540 injection with 980 nm laser irradiation;

- Group 3: UCNPs@mSiO₂ MC540 injection with 1550 nm laser irradiation;

- Group 4: UCNPs@mSiO₂ FePc injection with 1550 nm laser irradiation;

- Group 5: UCNPs@mSiO₂ FePc MC540 injection with 1550 nm laser irradiation.

The laser power density was 235 mW cm⁻², and the irradiation time for each group was 10 min.

The group treated with UCNPs@mSiO₂ FePc MC540 plus 1550 nm light showed the greatest cell reduction and apoptosis.

Moreover, no significant damage or lesions were observed in the major organs (heart, liver, spleen, lungs, and kidneys) across the different groups. This also indicates that the developed nanoagents UCNPs@mSiO₂ FePc MC540 do not exhibit pronounced toxicity in mice.

One of the mechanisms of PDT may be the stimulation of iron dependent lipid peroxidation – ferroptosis.

The authors developed a nanozyme that responds in the NIR region. It features a uniform dispersion of Fe atoms on a g C₃N₄ doping network within a mesoporous upconversion structure (UmFe), followed by coating with a hybrid platelet membrane (hPM) and oleanolic acid (OA) to obtain UmFe OA@hPM [39].

To establish a primary tumor model, 100 µL of a 4T1 cell suspension (10⁶ cells per mouse) was subcutaneously injected into the right flank of 6 week old male BALB/c mice. When the tumor volume approached 150 mm³, the therapeutic agents were administered intravenously to the mice five times, with a 4 day interval between doses (UmFe dose = 10 mg kg⁻¹; OA dose = 0.5 mg kg⁻¹).

Laser irradiation (1 W cm⁻², 15 min) was applied 24 h after the injection. In addition to excellent antitumor efficacy, UmFe OA treatment exhibited virtually no serious toxic effects on mice and normal tissues, as evidenced by the mice's weight chart and H&E staining of major organs.

The highest nanoparticle accumulation was observed 24–48 h after intravenous administration. Samples coated with hPM demonstrated a higher tumor targeting capability while reducing accumulation in the liver compared to liposomal NPs.

UmFe OA showed high antitumor efficacy – 74% among all therapeutic groups. This can be attributed to the synergistic enhancement of ferroptosis upon NIR treatment, compared to induction solely via the Fenton reaction. A similar trend was observed without laser irradiation, but the overall antitumour efficacy was lower due to the slower rate of the Fenton reaction.

In samples (G5) and (G6) with UmFe OA@hPM, tumor cell density was significantly lower compared to samples from other groups. These two laser treated groups also exhibited the highest level of TUNEL fluorescence.

Besides its significant antitumor efficacy, UmFe OA(+) treatment showed virtually no serious toxicity to mice and normal tissues.

The effect of PDT using UCNPs on ferroptosis was investigated in the study [40].

The authors developed a nanoplatform (ApoE UMSNs GOx/SRF) based on UMSNPs (UCNPs@mSiO₂). A layer of mesoporous silica and a lipid bilayer were sequentially deposited onto the UCNPs, with glucose oxidase (GOx) and the multikinase inhibitor sorafenib (SRF) loaded onto them, respectively. Subsequently, the apolipoprotein E (ApoE) peptide was attached.

The UCNPs used convert NIR radiation into UV light. This enables the reduction of iron ions (Fe³⁺) to Fe²⁺, which is critical for initiating ferroptosis.

The antitumor efficacy of ApoE UMSNs GOx/SRF was studied *in vivo* using an orthotopic glioma G422 model. The antitumor efficacy of ApoE UMSNs GOx/SRF was evaluated with and without NIR light treatment.

Tumor bearing mice were randomly assigned to six groups to evaluate antitumor efficacy as follows: (1) control; (2) UMSNs Gox; (3) UMSNs SRF; (4) UMSNs Gox/SRF; (5) ApoE UMSNs GOx/SRF; and (6) ApoE UMSNs GOx/SRF + NIR (doses: SRF: 20 mg/kg and GOx: 40 µg/kg; *n* = 6 per group).

The tumors were exposed to laser irradiation at 808 nm (0.5 W/cm², 5 min).

The UMSNs GOx/SRF nanoparticles demonstrated enhanced tumor growth inhibition (TGI), which may be attributed to an improved Fenton reaction induced by GOx – resulting in additional H₂O₂ production through catalysed glucose oxidation.

The use of ApoE UMSNs GOx/SRF further reduced the tumor growth rate, achieving a TGI of 90.3%, which is significantly higher than that of UMSNs GOx/SRF (81%).

The antitumor effect of ApoE UMSNs GOx/SRF was enhanced by laser irradiation at 808 nm (0.25 W/cm², 5 min) administered 6 h after ApoE UMSNs GOx/SRF injection. In this group, the TGI reached 97.2%, and glioma growth was successfully suppressed. These results indicate that light mediated Fe²⁺ regeneration induced by ApoE UMSNs GOx/SRF – combined with H₂O₂ generated by GOx – can effectively enhance the antitumor effect.

Treatment with ApoE UMSNs GOx/SRF caused a significant decrease in ferroptosis related protein (xCT) expression, from 36.5% to 16.9%. NIR irradiation further reduced xCT expression to 8.98%, which can be explained by NIR induced enhancement of ferroptosis mediated by UCNPs. These findings suggest that ApoE UMSNs GOx/SRF inhibits GBM through NIR enhanced ferroptosis.

Throughout the study, no significant body weight loss was observed. Serum biochemistry and histological analyses of major organs showed all parameters within normal ranges, indicating minimal toxicity of the developed formulations.

Study demonstrated the phototoxicity of UCNPs supplemented with flavin mononucleotide (Pharmstandart®) in PDT [41]. In tumors treated with UCNPs followed by irradiation at 975 nm, a sharp

decrease in nucleated cells was observed. Additionally, some tumor areas were filled with eosinophilic material or empty spaces devoid of cells. PDT caused vascular rupture with extensive haemorrhages in tumour tissue.

A nanohybrid combining UCNPs and metal – organic frameworks (MOFs) exhibits a unique heterogeneous structure that facilitates efficient energy transfer from UCNPs to MOFs, thereby enhancing NIR activated photodynamic therapy [42].

In the study, mice bearing HepG2 tumors were administered nanoparticles intravenously at a concentration of 4 mg/mL (100 µL). Tumors were irradiated 8 h post injection using a 980 nm laser (1.5 W cm⁻¹, 10 min total irradiation time with a 2 min break after the first 2 min of irradiation). Additional injections and treatments using the same parameters were performed every two days. After 14 days of treatment, the mice were euthanized.

Post treatment analysis revealed suppressed proliferation and apoptosis of tumor cells. The nanoparticles's ability to alleviate hypoxia in solid tumors was also confirmed. This effect is primarily attributed to the catalytic activity of the nanoparticles, which decomposed intracellular H₂O₂ and produced O₂ within the tumor microenvironment. This process enhanced the efficacy of oxygen dependent type II PDT *in vivo*.

Experiments showed that 6 h after injection, the nanoparticles distributed to all organs (brain, liver, kidneys, spleen, lungs, and tumors), with most of them detected in the kidneys, spleen, liver, and tumors. After 48 h, they were almost completely metabolised in all organs except tumor tissues. Immunohistochemical experiments revealed no significant changes in the morphology of each tissue across all treatment groups. Hematological parameters showed no noticeable differences from the control group.

Pharmacological activity has been described with intravenous administration of UCNP to animals with tumors, but safety has been described with intratracheal administration of UCNP to healthy animals. This can be explained by the fact that, along with reducing systemic toxicity, the inhalation method can cause complications directly upon administration [43, 44].

The dependence of toxicity on such parameters as particle type, PS, dose, irradiation, model and result, summarizing the reviewed articles, is presented in Table 1.

It has been shown previously that the most non-toxic concentration identified across studies is 2 mg/mL (see Part I of this Overview). For mice, the maximum permissible dose for short-term exposure (12 h) is 10 mg/kg, and for long-term exposure (120 days) is 50 mg/kg. For rats, a dose of 5 mg/kg is safe for both short- and long-term studies of dark toxicity. Dark toxicity occurs at higher doses than phototoxicity (e.g., dark: 50 mg/kg for 120 days; photo: 10 mg/kg for 40 days).

Table 1 The dependence of toxicity on such parameters as particle type, PS, dose, irradiation, model and result.

Particle type / coating	Photosensitizer (PS) / drug	Dose (mice)	Irradiation (NIR)	Model	Toxicity & efficacy result	Ref.
YOF:Yb,Tm@YOF:Yb + mesoporous silica + DEX (USDPPFs)	Dexamethasone (DEX)	5 mg/kg (intratracheal or i.v.)	980 nm, 3×10 min, 1 W/cm ²	C57BL/6 mice (lung I/R injury)	No inflammatory cell infiltration or necrosis in heart, lung, liver, spleen, kidney; NIR-triggered DEX release reduced lung I/R injury	[1]
DUCNP@Mn-MOF	FOE (evodiamine derivative)	not specified	not specified (chemodynamic + chemo)	Breast cancer mouse model	High tumor targeting, tumor suppression, no significant toxicity	[5]
NaYF ₄ :Yb,Tm + 4T1 cancer cell membrane + melittin (U ACPT@MM)	9-aminocamptothecin (ACPT)	Melittin 1 mg/kg (i.v., days 0,3,6)	808 nm, 0.8 W/cm ² , 20 min (days 2,5,8)	4T1 breast cancer	Tumor growth inhibition, lung metastasis suppression; increased TNFα/IFNγ; no organ damage (heart, liver, spleen, lung, kidney)	[6]
TiO ₂ /(Ca,Y)F ₂ :Tm,Yb (femtosecond laser treated)	none (ROS generation)	not specified	NIR (980 nm)	CFPAC-1 pancreatic tumor (nude mice)	Tumor growth inhibition (not eradication); no large-area tumor cell killing due to hypoxia	[10]
NaGdF ₄ :Yb,Nd,Tm + TCI	thiochromenocarbazole imide (TCI)	not specified	808 or 980 nm	HeLa tumor	Low dark toxicity, imaging possible, but no significant tumor growth inhibition	[9]
UCNP@Glu-DMMA (charge reversal, glucose targeting)	none (PDT via ROS)	not specified	980 nm, 0.8 W/cm ² , 15 min (30 s on/60 s off)	HeLa tumor (nude mice)	Significant tumor reduction, minimal organ damage/inflammation	[12]
NaYF ₄ :Yb,Er@PEG	Chlorin e6 (Ce6)	not specified (intratumoral)	980 nm, 0.5 W/cm ² , 30 min	4T1 breast cancer	Significant reduction in tumor growth rate	[16]
NaYF ₄ :Yb,Er@NaGdF ₄ @PEG-phospholipid	Ce6	not specified (i.v.)	980 nm, 0.6 W/cm ² , 5 min	U87MG glioblastoma	Tumor growth reduction	[17]
NaYF ₄ :Yb,Er + chitosan	Zinc phthalocyanine	not specified	980 nm, 0.4 W/cm ² , 2×15 min (3 h & 24 h post inj.)	S180 sarcoma	Slight tumor volume increase, high survival	[18]
NaYF ₄ :Yb,Er@CaF ₄ @PAA-hydrazide	5-aminolevulinic acid	not specified	980 nm, 0.5 W/cm ² , 40 min	not specified	Significant tumor size reduction	[19]

Table 1 (Cont.)

NaYF ₄ :Yb,Er@phosphatidylcholine-FA	Tetraphenylporphyrin	not specified (intratumoral)	980 nm, 1 W/cm ² , 30 min	CT26wt colon cancer	Reduced tumor growth	[20]
UCNP + branched PEI + FA	Protoporphyrin (PPIX)	not specified (i.v., 4 weeks)	not specified	Skin tumor	No organ damage (liver, heart, kidney, spleen, lung); irradiation caused severe tumor destruction	[24]
NaErF ₄ @NaLuF ₄ + macrophage membrane + mannose + SiO ₂ (PFC) + Ce6 + PTX	Ce6 + paclitaxel (PTX)	6.5 mg/kg UCNP (0.4/0.6/0.5 mg/kg for PFC/Ce6/PTX)	808 nm, 1.5 W/cm ² , 20 min	4T1 breast cancer	Most significant tumor suppression; increased tumor O ₂ (>40% vs 19–20% controls); no toxicity reported	[25]
Hb/BSA/Ce6/Er@Lu microgels (Hb μ Gels)	Ce6	not specified (intratumoral)	808 nm	4T1 breast cancer	ROS + thermal stress (>48°C); hypoxia relief; minimal organ accumulation, no histology changes	[26]
ZnO NPs (catalase-like, upconversion)	Ce6	not specified	808 nm (ZnO-Ce6) vs 660 nm (Ce6 alone)	4T1 breast cancer	90% suppression (808 nm ZnO-Ce6) vs 80.2% (660 nm Ce6); 3/5 mice cured; hypoxia relieved	[29]
BaGdF ₅ :Yb,Tm	Doxorubicin (DOX)	10 mg/kg (i.v.)	not specified	not specified	No adverse effects; survival 40 days	[30]
NaYF ₄ :Yb,Er,Fe@PM VEMA	THPC (Foscan)	not specified (intratumoral)	980 nm, 0.5 W/cm ² , 3 min	Capan-2 pancreatic adenocarcinoma (nude mice)	Day 30 tumor volume 0.35 cm ³ ; extensive necrosis by day 1–2; no body weight loss	[32]
NaYF ₄ :Yb,Er@PEG	THPC + IR780	not specified	980 nm (0.8) + 808 nm (0.36 W/cm ²), 5 min	Glioblastoma	No toxicity without laser; effective therapy	[36]
Core-multishell UCNPs (1550/808/980 nm) + mSiO ₂ + ZnPc + RBS (NO donor)	Zn-phthalocyanine (ZnPc) + NO donor	not specified	1 W/cm ² , 20 min, various sequences	4T1 breast cancer	Sequential 980→1550 nm best suppression; no organ damage, no weight change	[37]
NaErF ₄ @NaYbF ₄ @NaYF ₄ @mSiO ₂	MC540 + FePc	not specified	1550 nm, 235 mW/cm ² , 10 min	HeLa (nude mice)	Best cell reduction with dual PS; no organ lesions	[38]
UmFe (Fe/g-C ₃ N ₄ on mesoporous UCNP) + platelet membrane (hPM) + oleanolic acid (OA)	OA (ferroptosis)	UmFe 10 mg/kg, OA 0.5 mg/kg (i.v., 5 doses, 4-day interval)	1 W/cm ² , 15 min (24 h post inj.)	4T1 breast cancer	74% antitumor efficacy; no serious toxicity; highest TUNEL in laser groups	[39]

Table 1 (Cont.)

UCNP@mSiO ₂ + lipid bilayer + ApoE (ApoE UMSNs GOx/SRF)	Glucose oxidase (GOx) + sorafenib (SRF)	SRF 20 mg/kg, GOx 40 µg/kg	808 nm, 0.5 W/cm ² , 5 min	Orthotopic G422 glioma	TGI 97.2% (laser) vs 90.3% (no laser); reduced xCT expression; no weight loss, normal serum/organ histology	[40]
UCNPs (unspecified)	Flavin mononucleotide	not specified	975 nm	Tumor model	Sharp decrease in nucleated cells, vascular rupture, hemorrhages	[41]
UCNP + MOF (heterostructure)	none (O ₂ generation from H ₂ O ₂)	4 mg/mL, 100 mL (i.v.)	980 nm, 1.5 W/cm ² , 10 min (2 min on/2 min off), every 2 days	HepG2 tumor	Hypoxia relief; nanoparticle cleared by 48 h except tumor; no organ morphology changes	[42]

2 Conclusion

This review examined the *in vivo* phototoxicity of upconversion nanoparticles (UCNPs) in photoinduced drug delivery, photodynamic therapy (PDT), and multimodal therapies.

UCNPs activated by NIR irradiation (808/980 nm, 0.15–1.5 W/cm², 3–40 min) demonstrate targeted phototoxicity confined to tumors. No significant damage to major organs (heart, liver, spleen, lungs, kidneys) was observed across numerous studies.

The doses of UCNP in the range of 5–10 mg/kg (single or repeated administration) did not cause histological changes in internal organs and did not lead to death of animals for up to 40 days of observation. From the point of view of clinical translation, it is important to know what happens to experimental animals in the long term up to 40 days, but unfortunately, for the specified types of particles, this information was not found in the articles.

Significant tumor suppression was achieved in breast cancer, pancreatic adenocarcinoma, cervical carcinoma, glioblastoma, colorectal cancer, and others. Complete eradication occurred in some cases (e.g., ZnO-Ce6 NPs: 90% suppression, 3/5 mice cured).

Photosensitizers are safe. Ce6, THPC, ZnPc, PPIX, methylene blue, and others showed no additional systemic toxicity while enabling efficient singlet oxygen generation.

Hypoxia alleviation improves efficacy. Oxygen carriers (perfluorocarbons, hemoglobin, catalase-

mimicking ZnO) increased tumor O₂ levels (>40% vs. 19–20% controls), enhancing PDT outcomes.

Multimodal combinations are safe. UCNPs combined with chemotherapy, PTT, CDT, immunotherapy, radiotherapy, or cryotherapy showed enhanced efficacy without increased toxicity (e.g., TGI up to 97.2% in glioblastoma).

PEG, silica, PMVEMA, and cell membranes (cancer, erythrocyte, platelet, macrophage) ensured low toxicity. Charge-reversal coatings improved tumor targeting.

UCNPs cleared within 48 h to 40 days via fecal/urinary routes, with minimal long-term retention.

Studies up to 40 days showed no weight loss, behavioral changes, or organ damage at recommended doses.

The safety profile and efficacy support further development of UCNPs for deep-tissue cancer therapy and theranostics.

UCNPs exhibit low phototoxicity at established doses (mice 5–50 mg/kg). Multimodal platforms enhance efficacy without compromising safety, positioning UCNPs as promising agents for precision oncology.

Disclosures

All authors declare that there is no conflict of interests in this paper.

Acknowledgements

The study was supported by a grant Russian Science Foundation No. 25-22-00144, <https://rscf.ru/project/25-22-00144/>.

References

1. X. He, Z. Li, M. Ye, C. Zhao, S. Wu, Y. Qin, Y. Guo, L. Zhang, and F. Lin, “Near-infrared laser-irradiated upconversion nanoparticles with dexamethasone precise released for alleviating lung ischemia-reperfusion injury,” *Frontiers in Bioengineering and Biotechnology* 11, 1176369 (2023).

2. X. Hu, Y. Tang, Y. Hu, F. Lu, X. Lu, Y. Wang, J. Li, Y. Li, Y. Ji, W. Wang, D. Ye, Q. Fan, and W. Huang, "Gadolinium-Chelated Conjugated Polymer-Based Nanotheranostics for Photoacoustic/Magnetic Resonance/NIR-II Fluorescence Imaging-Guided Cancer Photothermal Therapy," *Theranostics* 9(14), 4168–4181 (2019).
3. X. Hu, Z. Chen, A. J. Jin, Z. Yang, D. Gan, A. Wu, H. Ao, W. Huang, and Q. Fan, "Rational Design of All-Organic Nanoplatform for Highly Efficient MR/NIR-II Imaging-Guided Cancer Phototheranostics," *Small* 17(12), 2007566 (2021).
4. Y. Lu, Q. Luo, X. Jia, J. P. Tam, H. Yang, Y. Shen, and X. Li, "Multidisciplinary strategies to enhance therapeutic effects of flavonoids from *Epimedium Folium*: Integration of herbal medicine, enzyme engineering, and nanotechnology," *Journal of Pharmaceutical Analysis* 13(3), 239–254 (2023).
5. X. Zhao, S. He, B. Li, B. Liu, Y. Shi, W. Cong, F. Gao, J. Li, F. Wang, K. Liu, C. Sheng, J. Su, and H.-G. Hu, "DUCNP@Mn–MOF/FOE as a Highly Selective and Bioavailable Drug Delivery System for Synergistic Combination Cancer Therapy," *Nano Letters* 23(3), 863–871 (2023).
6. Y. Wang, Y. Qiu, S. Chen, J. Huang, X. Hu, J. Chen, S. Wang, X. Yang, Y. Zhang, and Y. Zhu, "Functionalized Tumor Cell Membrane-Camouflaged Photo-Activatable Nanoparticle for Spatiotemporal Antitumor Therapy," *Chemical Engineering Journal* 474, 145676 (2023).
7. J. R. Groom, J. Richmond, T. T. Murooka, E. W. Sorensen, J. H. Sung, K. Bankert, U. H. von Andrian, J. J. Moon, T. R. Mempel, and A. D. Luster, "CXCR3 Chemokine Receptor-Ligand Interactions in the Lymph Node Optimize CD4⁺ T Helper 1 Cell Differentiation," *Immunity* 37(6), 1091–1103 (2012).
8. M. Z. El-Sadek, M. K. A. El-Aziz, A. H. Shaaban, S. A. Mostafa, and A.-H. S. Wadan, "Advancements and emerging trends in photodynamic therapy: innovations in cancer treatment and beyond," *Photochemical & Photobiological Sciences* 24(8), 1489–1511 (2025).
9. A. Nsubuga, K. Morice, N. Fayad, F. Pini, V. Jossierand, X. Le Guével, A. Alhabi, M. Henry, D. Puchán Sánchez, N. Plassais, P. Josse, J. Boixel, P. Blanchard, C. Cabanetos, and N. Hildebrandt, "Sub 20 nm Upconversion Photosensitizers for Near-Infrared Photodynamic Theranostics," *Adv Funct Materials* 35(1), 2410077 (2025).
10. H. Zhou, J. Wang, B. Yang, M. Sun, C. Liang, L. Yang, X. Zhao, and H. Wang, "Preparation of near-infrared light responsive upconversion nanoparticles TiO₂/(Ca,Y)F₂:Tm,Yb for photodynamic antitumor therapy," *Journal of Materials Science* 58(15), 6743–6756 (2023).
11. E. Bugyik, F. Renyi-Vamos, V. Szabo, K. Dezso, N. Ecker, A. Rokusz, P. Nagy, B. Dome, and S. Paku, "Mechanisms of vascularization in murine models of primary and metastatic tumor growth," *Cancer Communications* 35(1), 19 (2016).
12. B. Ling, L. Yang, C. Wang, L. Dong, Y. Yang, L. Wang, J. Zhang, and Y. Yuan, "Cascade-Responsive Upconversion Nanoplatform for Efficient Cell Nucleus Targeting and Boosted Photodynamic Tumor Therapy," *ACS Materials Letters* 6(12), 5256–5265 (2024).
13. C. A. Del Valle, T. Hirsch, and M. J. Marín, "Recent advances in near infrared upconverting nanomaterials for targeted photodynamic therapy of cancer," *Methods and Applications in Fluorescence* 10(3), 034003 (2022).
14. M. R. Hamblin, "Photodynamic Therapy for Cancer: What's Past is Prologue," *Photochemistry and Photobiology* 96(3), 506–516 (2020).
15. Y. Li, G. Chen, "Upconversion Nanoparticles for Cancer Therapy," *Advanced NanoBiomed Research* 2(12), 2200092 (2022).
16. C. Wang, H. Tao, L. Cheng, and Z. Liu, "Near-infrared light induced *in vivo* photodynamic therapy of cancer based on upconversion nanoparticles," *Biomaterials* 32(26), 6145–6154 (2011).
17. Y. I. Park, H. M. Kim, J. H. Kim, K. C. Moon, B. Yoo, K. T. Lee, N. Lee, Y. Choi, W. Park, D. Ling, K. Na, W. K. Moon, S. H. Choi, H. S. Park, S. Yoon, Y. D. Suh, S. H. Lee, and T. Hyeon, "Theranostic Probe Based on Lanthanide-Doped Nanoparticles for Simultaneous *In Vivo* Dual-Modal Imaging and Photodynamic Therapy," *Advanced Materials* 24(42), 5755–5761 (2012).
18. W. Cai, X. Chen, Q. Pan, S. Zhang, L. Tan, X. Sun, R. Huang, and A. Xia, "Antitumor activity of yulansan polysaccharides in mice bearing S180 sarcoma tumors," *Molecular and Clinical Oncology* 7(4), 716–720 (2017).
19. S. Cui, H. Chen, H. Zhu, J. Tian, X. Chi, Z. Qian, S. Achilefu, and Y. Gu, "Amphiphilic chitosan modified upconversion nanoparticles for *in vivo* photodynamic therapy induced by near-infrared light," *Journal of Materials Chemistry* 22(11), 4861 (2012).
20. P. Thanasekaran, C.-H. Chu, S.-B. Wang, K.-Y. Chen, H.-D. Gao, M. M. Lee, S.-S. Sun, J.-P. Li, J.-Y. Chen, J.-K. Chen, Y.-H. Chang, and H.-M. Lee, "Lipid-Wrapped Upconversion Nanoconstruct/Photosensitizer Complex for Near-Infrared Light-Mediated Photodynamic Therapy," *ACS Applied Materials & Interfaces* 11(1), 84–95 (2019).
21. C.-W. Chen, Y.-C. Chan, M. Hsiao, and R.-S. Liu, "Plasmon-Enhanced Photodynamic Cancer Therapy by Upconversion Nanoparticles Conjugated with Au Nanorods," *ACS Applied Materials & Interfaces* 8(47), 32108–32119 (2016).
22. Y. Li, X. Zhang, Y. Zhang, Y. Zhang, Y. He, Y. Liu, and H. Ju, "Activatable Photodynamic Therapy with Therapeutic Effect Prediction Based on a Self-correction Upconversion Nanoprobe," *ACS Applied Materials & Interfaces* 12(17), 19313–19323 (2020).

23. N. Zhao, B. Wu, X. Hu, and D. Xing, “NIR-triggered high-efficient photodynamic and chemo-cascade therapy using caspase-3 responsive functionalized upconversion nanoparticles,” *Biomaterials* 141, 40–49 (2017).
24. H. Shan, X. Wang, Q. Wei, H. Dai, and X. Chen, “Enriched photosensitizer for deep-seated-tumor photodynamic therapy,” *Photonics Research* 12(5), 1024 (2024).
25. J. Yoon, X. T. Le, J. Kim, H. Lee, N. T. Nguyen, W. T. Lee, E. S. Lee, K. T. Oh, H.-G. Choi, and Y. S. Youn, “Macrophage-reprogramming upconverting nanoparticles for enhanced TAM-mediated antitumor therapy of hypoxic breast cancer,” *Journal of Controlled Release* 360, 482–495 (2023).
26. H. Kim, J. Yoon, H. K. Kim, W. T. Lee, N. T. Nguyen, X. T. Le, E.-H. Lee, E. S. Lee, K. T. Oh, H.-G. Choi, and Y. S. Youn, “Upconverting nanoparticle-containing erythrocyte-sized hemoglobin microgels that generate heat, oxygen and reactive oxygen species for suppressing hypoxic tumors,” *Bioactive Materials* 22, 112–126 (2023).
27. J. A. Raleigh, S.-C. Chou, G. E. Arteel, and M. R. Horsman, “Comparisons among Pimonidazole Binding, Oxygen Electrode Measurements, and Radiation Response in C3H Mouse Tumors,” *Radiation Research* 151(5), 580 (1999).
28. X.-F. Li, S. Carlin, M. Urano, J. Russell, C. C. Ling, and J. A. O’Donoghue, “Visualization of Hypoxia in Microscopic Tumors by Immunofluorescent Microscopy,” *Cancer Research* 67(16), 7646–7653 (2007).
29. J. Wang, M. Liu, J. Wang, Z. Li, Z. Feng, M. Xu, H. Wang, H. Li, Z. Li, J. Yu, J. Liu, Q. Wei, S. Zhang, and X. Zhang, “Zinc oxide nanoparticles with catalase-like nanozyme activity and near-infrared light response: A combination of effective photodynamic therapy, autophagy, ferroptosis, and antitumor immunity,” *Acta Pharmaceutica Sinica B* 14(10), 4493–4508 (2024).
30. D. Yang, Y. Dai, J. Liu, Y. Zhou, Y. Chen, C. Li, P. Ma, and J. Lin, “Ultra-small BaGdF₅-based upconversion nanoparticles as drug carriers and multimodal imaging probes,” *Biomaterials* 35(6), 2011–2023 (2014).
31. Y.-H. Chien, Y.-L. Chou, S.-W. Wang, S.-T. Hung, M.-C. Liao, Y.-J. Chao, C.-H. Su, and C.-S. Yeh, “Near-Infrared Light Photocontrolled Targeting, Bioimaging, and Chemotherapy with Caged Upconversion Nanoparticles *in Vitro* and *in Vivo*,” *ACS Nano* 7(10), 8516–8528 (2013).
32. O. Shapoval, D. Větvicka, V. Patsula, H. Engstová, O. Kočková, M. Konefał, M. Kabešová, and D. Horák, “Temoporfin-Conjugated Upconversion Nanoparticles for NIR-Induced Photodynamic Therapy: Studies with Pancreatic Adenocarcinoma Cells *In Vitro* and *In Vivo*,” *Pharmaceutics* 15(12), 2694 (2023).
33. O. Shapoval, V. Patsula, D. Větvicka, H. Engstová, V. Oleksa, M. Kabešová, T. Vasylyshyn, P. Poučková, and D. Horák, “Temoporfin-Conjugated PEGylated Poly(N , N -dimethylacrylamide)-Coated Upconversion Colloid for NIR-Induced Photodynamic Therapy of Pancreatic Cancer,” *Biomacromolecules* 25(9), 5771–5785 (2024).
34. N. M. Idris, M. K. Gnanasammandhan, J. Zhang, P. C. Ho, R. Mahendran, and Y. Zhang, “*In vivo* photodynamic therapy using upconversion nanoparticles as remote-controlled nanotransducers,” *Nature Medicine* 18(10), 1580–1585 (2012).
35. S. Yan, X. Zeng, Y. Tang, B. Liu, Y. Wang, and X. Liu, “Activating Antitumor Immunity and Antimetastatic Effect Through Polydopamine-Encapsulated Core-Shell Upconversion Nanoparticles,” *Advanced Materials* 31(46), 1905825 (2019).
36. Y.-C. Tsai, P. Vijayaraghavan, W.-H. Chiang, H.-H. Chen, T.-I. Liu, M.-Y. Shen, A. Omoto, M. Kamimura, K. Soga, and H.-C. Chiu, “Targeted Delivery of Functionalized Upconversion Nanoparticles for Externally Triggered Photothermal/Photodynamic Therapies of Brain Glioblastoma,” *Theranostics* 8(5), 1435–1448 (2018).
37. S. Zheng, H. Zhang, T. Sheng, Y. Xiang, J. Wang, Y. Tang, and Y. Zhang, “Photoswitchable upconversion nanoparticles with excitation dependent emission for programmed stepwise NIR phototherapy,” *iScience* 26(10), 107859 (2023).
38. T. Zhao, D. Song, J. Wang, H. Chen, W. Chen, and X. Zhu, “Nanoshell-mediated color tuning of erbium-sensitized upconversion nanoparticles for advanced phototherapy above 1500 nm,” *Journal of Photochemistry and Photobiology A: Chemistry* 463, 116285 (2025).
39. X. T. Le, N. T. Nguyen, W. T. Lee, Y. Yang, H. Choi, and Y. S. Youn, “Peroxidase-Mimicking Iron-Based Single-Atom Upconversion Photocatalyst for Enhancing Chemodynamic Therapy,” *Advanced Functional Materials* 34(34), 2401893 (2024).
40. K. Xue, R. Yang, Y. An, Y. Ding, S. Li, F. Miao, D. Liu, D. Chen, and Q. Tang, “NIR-promoted ferrous ion regeneration enhances ferroptosis for glioblastoma treatment,” *Journal of Controlled Release* 368, 595–606 (2024).
41. E. V. Khaydukov, K. E. Mironova, V. A. Semchishen, A. N. Generalova, A. V. Nechaev, D. A. Khochenkov, E. V. Stepanova, O. I. Lebedev, A. V. Zvyagin, S. M. Deyev, and V. Ya. Panchenko, “Riboflavin photoactivation by upconversion nanoparticles for cancer treatment,” *Scientific Reports* 6(1), 35103 (2016).
42. X. Zhang, J. Cui, J. Liu, X. Chen, M. Chen, and J. Wang, “Dual ligand-assisted assembly of metal–organic frameworks on upconversion nanoparticles for NIR photodynamic therapy against hypoxic tumors,” *Journal of Materials Chemistry B* 11(39), 9516–9524 (2023).
43. T. Okuda, H. Okamoto, “Present Situation and Future Progress of Inhaled Lung Cancer Therapy: Necessity of Inhaled Formulations with Drug Delivery Functions,” *Chemical and Pharmaceutical Bulletin* 68(7), 589–602 (2020).
44. Y. Han, Y. Yang, Q. Sun, B. Li, C. Yue, Y. Liu, J. M. De La Fuente, and D. Cui, “Dual-targeted lung cancer therapy via inhalation delivery of UCNP-siRNA-AS1411 nanocages,” *Cancer Biology & Medicine* 19(7), 1047–1060 (2021).

Photocrosslinking of Tarsal Collagen for Treating Eyelid Laxity: Applying the Exposure Reciprocity Law to Establish an Optimal Irradiant Profile

Traian V. Chirila^{1,2,3,4*}, Shuko Suzuki¹, Nadja E. Pop^{3,4,5}, Claire Cattelot⁶, and Alexandra I. Manta^{1,3,4,5,7}

¹ Queensland Eye Institute, 87 Ipswich Road, Woolloongabba, Queensland 4102, Australia

² Australian Institute of Bioengineering and Nanotechnology, University of Queensland, St Lucia, Queensland 4072, Australia

³ Centre for Advanced Medical and Pharmaceutical Research (CCAMF), George E. Palade University of Medicine, Pharmacy, Sciences and Technology, 540139 Targu Mures, Romania

⁴ Faculty of General Medicine, George E. Palade University of Medicine, Pharmacy, Sciences and Technology, 540139 Targu Mures, Romania

⁵ Doctoral School, George E. Palade University of Medicine, Pharmacy, Sciences and Technology, 540139 Targu Mures, Romania

⁶ Aix-Marseille University, Polytech Marseille, 13284 Marseille Cedex 07, France

⁷ Faculty of Health and Behavioural Sciences, University of Queensland, St Lucia, Queensland 4072, Australia

*e-mail: traian.chirila@qei.org.au

Abstract. The aim of this study was to develop an irradiant exposure strategy able to control validity of the exposure reciprocity law upon exposing tarso-conjunctival tissue to ultraviolet A (UVA) radiation as a treatment for palpebral laxity, and thus to establish a research-based criterion for implementing optimal irradiation levels in clinical settings. Tarso-conjunctival specimens were harvested postmortem from ovine and caprine eyelids and exposed to UVA radiation (365 nm) in specified conditions. Thermographic infrared analysis was applied to monitor the photothermal response. By performing irradiation in the presence of a photosensitizer (riboflavin) a photochemomechanical response was also determined based on changes in mechanical properties as evaluated in a mechanical tester. A graphical method was employed to define the failure zones of reciprocity law when applied to radiation-induced photoresponses. Discontinuities in law's validity were found and confirmed by descending values of Schwarzschild's p -coefficients. The photothermal response showed that law's failure zone centered on an irradiance of 30 mW/cm² in both sheep and goat tissues, for radiant exposures between 6 and 18 J/cm². The photochemomechanical response indicated a failure onset around 60 mW/cm². Anticipating the concept of safest application route, an irradiance of 30 mW/cm² appeared an appropriate upper limit for the validity zone of reciprocity law, as indicated by the photothermal response, at exposure durations assuring a radiant exposure between 5 and 20 J/cm².

Keywords: eyelid laxity; collagen; ultraviolet A radiation; photochemical crosslinking; Bunsen-Roscoe law; Schwarzschild's coefficients.

Paper #9545 received 2 Mar 2026; revised manuscript received 18 May 2025; accepted for publication 18 May 2026; published online 13 Jul 2026. [doi: 10.18287/JBPE26.12.030301](https://doi.org/10.18287/JBPE26.12.030301).

1 Introduction

A method has been developed in our laboratories aimed at reinforcing mechanically the tarsal plate in the eyes afflicted by palpebral laxity, which is associated with a range of conditions such as floppy eyelid syndrome, entropion, ectropion and ptosis, and may affect the performance of meibomian glands. There is evidence [1–6] that tarsal laxity is linked to deficient biosynthesis of elastin and collagen, and to disarrangement of their networks.

Surgical procedures are frequently employed for treating eyelid laxity. According to the principle of our proposed method, the exposure of the tissue to ultraviolet radiation A (henceforth, UVA) in the presence of photosensitizers, such as riboflavin, can lead to an increase in both tensile strength and stiffness of the tarsus. Such mechanical enhancement is a consequence of the photochemical crosslinking of tarsal collagen, and could result in reversing the excessive palpebral laxity or in forestalling its post-surgical recurrence. We have confirmed augmentative mechanical effects both in *ex vivo* animal tissue (sheep) [7, 8] and in human cadaveric tissue [9], and other investigators have confirmed our findings in *ex vivo* tarsal plate specimens of animal (rat) or human origin [10–13]. The biological safety of tarsal collagen crosslinking procedure (henceforth, TXL) is critical for its therapeutic applications pertaining introduction to clinical practice. UVA radiation (conventionally between the wavelengths 315 and 400 nm) is less energetic than other UV regions, and is regarded as a non-ionizing and non-thermal radiation; however, it can still be deleterious to cells and tissues. Effects of UV radiation on skin [14–18] and ocular tissues [19–22] have been extensively investigated, but the scope was confined mainly to natural radiation as generated by the sun.

The effect of artificially sourced UVA radiation in non-cutaneous medical applications has been sparingly reported. In fact, most information on the subject comes from ophthalmology, where a therapeutic application of UVA-induced corneal stromal collagen crosslinking (henceforth, CXL) proved to be a successful treatment for keratoconus, an ectatic corneal disorder associated with severe visual disability. Using established UVA/riboflavin crosslinking protocols, progression of keratoconus can be effectively arrested due to mechanical reinforcement of the cornea that is induced through the photocrosslinking of stromal collagen [23–30]. Based on its additional germicidal and anti-inflammatory effects, UVA radiation was subsequently applied to treat other cornea disorders including pellucid marginal degeneration, postsurgical ectasia, refractive myopia, infectious keratitis, and bullous keratopathy [23, 31, 32] with some improvements in visual acuity and keratometric parameters.

When applying therapeutically the UVA radiation to various segments of the eye, it is essential to avoid injury to the target tissue and adjacent tissues caused by exposure to radiation, a process entirely dependent on the

radiant exposure (i.e., the energy of radiation delivered and absorbed by the target), which in turn is determined by both irradiance (i.e., the power of radiation delivered to the target) and exposure time, in agreement with an inverse proportionality relationship, as discussed later in depth. While for CXL a number of protocols have been developed which proved to be safe when applied in the clinic, aspects of the biological safety of TXL have been much less investigated [7–13, 33, 34]. Our study aims primarily at providing a scientific foundation for a rational selection and implementation of safe irradiation boundaries.

2 Methods and Materials

2.1 Materials

Eyelids were dissected postmortem and collected on site from domestic sheep (*Ovis aries*) and goat (*Capra hircus*) cadavers supplied by Brisbane Valley Meats Pty Ltd, Esk, Queensland, Australia. The supplier is an abattoir unit operating under the rules of the Australian Code of Practice of the Animal Welfare Standards for Livestock Processing Establishments (2001), the Australian Animal Welfare Standards and Guidelines for Sheep (2016), and the Australian Industry Welfare Standards and Guidelines for Goats (v. 03, 2020). The animals were slaughtered solely for commercial purposes, and the eyelid tissue would have been discarded if not acquired by us for the present study. Animals were between 1 and 2 years old at the time of sacrifice.

Riboflavin 5'-phosphate monosodium salt (RF-5P) was supplied by Cayman Chemical Co. (Ann Arbor, MI, USA). A solution of RF-5P in phosphate-buffered saline (PBS), at a concentration of 0.25% w/v, was applied only to the tissue specimens designated to undergo photocrosslinking.

2.2 Infrared Thermographic Analysis

Fifty full-thickness eyelid specimens were excised postmortem from the ocular adnexa of sheep and goat eyes that were refrigerated for less than 24 h. Eyelashes were trimmed, and the specimens were soaked for 30 min at room temperature in PBS solution. Each sample was then removed from solution and placed on a Petri dish with the conjunctival side facing upwards. Any excess of liquid was carefully blotted out with Kimwipe® paper tissues.

A thermal camera model FLIR C5 (FLIR Systems Australia Pty Ltd, Mulgrave, Victoria, Australia) with a thermal sensitivity of less than 0.07 °C was used for thermographic analysis following instrument calibration. As a source of radiation, the UV Curing System OmniCure® 1500 (Excelitas Technologies Corp., Waltham, MA, USA) was employed. To determine the level of irradiance at the exposure site, a radiometer Dymax ACCU-CAL 50 (Dymax Corp., Torrington, CT, USA) was used for the pre-irradiation adjustment of the distance between the UV source and tissue surface.

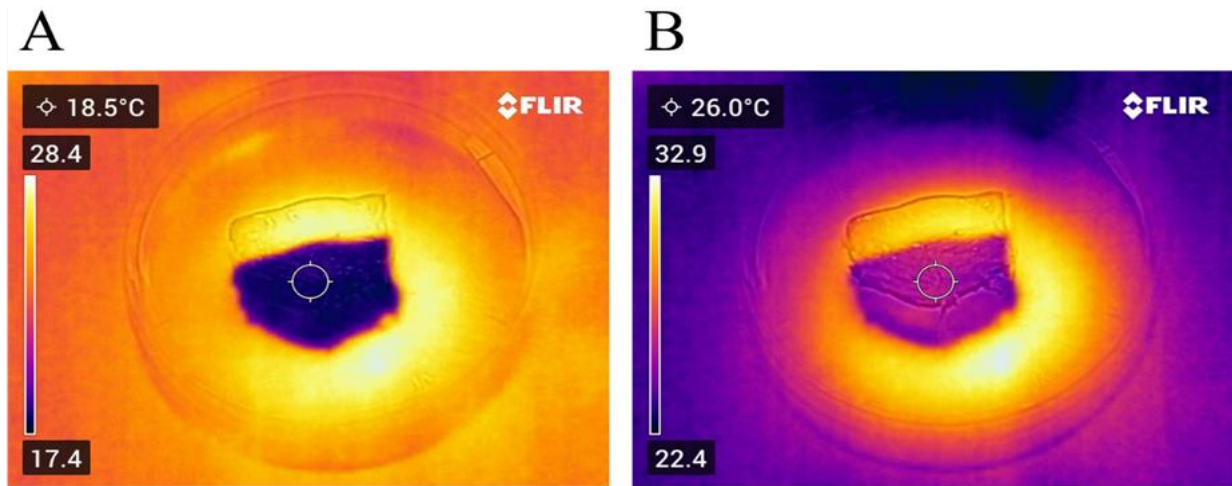


Fig. 1 Thermographic images recorded (A) before and (B) after the exposure to UVA radiation (365 nm) of a tarso-conjunctival specimen (goat eyelid) at an irradiance of 75 mW/cm^2 for an exposure time of 240 s. Total radiant exposure: 18 J/cm^2 .

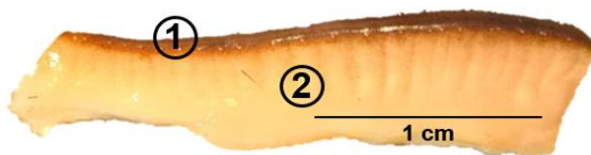


Fig. 2 Photographic image of a tarso-conjunctival specimen excised from goat eyelid: 1 – lid margin; 2 – palpebral conjunctiva.

The samples were each exposed for variable durations to UVA radiation (365-nm wavelength), at 10 different irradiances from 5 to 150 mW/cm^2 , assuring radiant exposures of 6, 13.5, and 18 J/cm^2 , respectively. Five specimens were used at each irradiance. The data were presented as mean values $\pm$ s.e.m.

The infrared camera mounted on a tripod was placed at a distance of *ca* 20 cm from the sample. The measurement spot (the ‘spotmeter’) on the camera screen was positioned onto the surface of the conjunctiva as centrally as possible. During each irradiation episode, the thermal images were recorded at the temperature on the target spot for a period corresponding to the target radiant exposure, and the baseline measurement prior to irradiation was also included. During experiments, the ambient temperature in laboratory was between 21 and 23°C . The baseline temperature, i.e., the temperature of the conjunctival surface prior to irradiation, was comparable in all experiments ($20.0 \pm 1.4^\circ\text{C}$). Fig. 1 shows a typical thermographic image recorded during measurements before and after exposure at a radiant exposure of 18 J/cm^2 , resulting from an irradiance of 75 mW/cm^2 applied for 240 s.

2.3 Photocrosslinking of Tarso-Conjunctival Tissue

Eighty tarso-conjunctival specimens were harvested postmortem from the upper eyelids of goats following a

surgical technique we have previously described in detail [9] for the dissection of equivalent human cadaveric tissue. Fig. 2 shows a tarso-conjunctival specimen as excised from the eyelid of a goat before its exposure to UVA radiation.

Each specimen was soaked in the RF-5P solution (0.25%) for 30 min, and then exposed to UVA (365-nm wavelength) radiation on the conjunctival side of the eyelid, using the protocol described in the previous section. After irradiation, all samples were stored in PBS until further processing and evaluation.

2.4 Biomechanical Analysis

Each tissue specimen was subjected to mechanical testing, both prior and after crosslinking. The Instron® Materials Testing System Model #5943 (Instron, Norwood, MA, USA), equipped with a 50-N load cell, was used to obtain stress-strain plots under uniaxial loading. Tarsal strips were secured using pneumatic grips set to a gauge length of 14 mm and submerged in PBS buffer at $21.5 \pm 1^\circ\text{C}$ in a BioPuls™ unit for the duration of testing. A tensile preloading of 30 mN was applied at the elongation rate of 0.01 mm/s. Subsequently, 15 cyclic tensile loading-unloading tests were performed to 10% displacement (strain) at a rate of 1%/s. The final cycle (#15) was used to determine the stress (σ) at 10% strain and to compute Young’s modulus (MY). The specimens were then removed from the mechanical tester, exposed to irradiation in order to accomplish crosslinking, and re-evaluated in the tester after soaking in PBS for at least 30 min.

3 Results

Fig. 3 shows the plots of the maximum temperature increase (ΔT) at the conjunctival surface as a function of irradiance (E_e), at three different levels of radiant exposure (H_e) in two different animal eyelids, sheep (Fig. 3A) and goat (Fig. 3B).

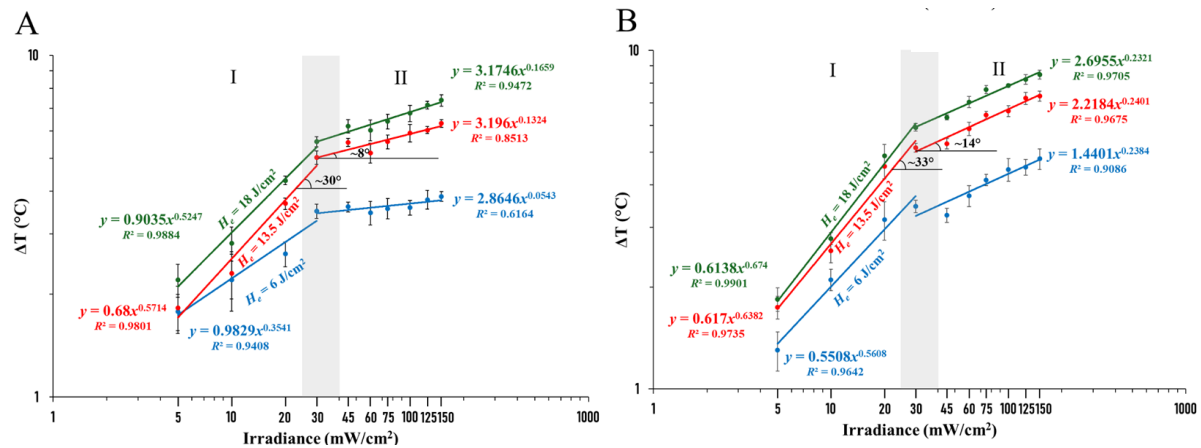


Fig. 3 Graphical method to assess the validity of Bunsen-Roscoe law and to quantify the value of Schwarzschild's p -coefficients based on the photothermal response (ΔT) in irradiated tarso-conjunctival specimens excised from (A) sheep and (B) goat eyelids, and exposed to UVA radiation (365 nm). Thermographic analysis was conducted at three different levels of radiant exposure (6, 13.5, and 18 J/cm^2). Each data point value is an average of 5 measurements $\pm$ s.e.m. Zone I – validity zone; zone II – failure zone. The inclination angle (θ) is indicated for each plotted line.

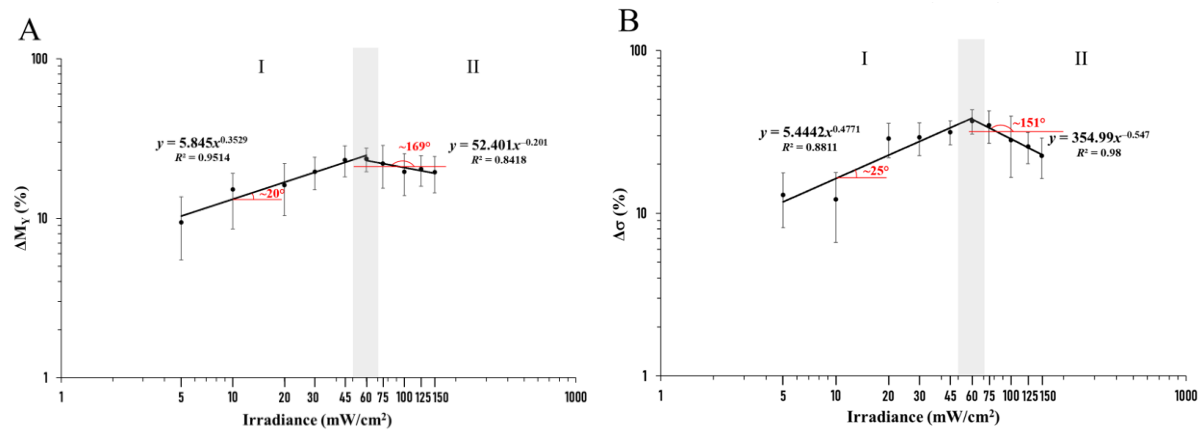


Fig. 4 Graphical method to assess the validity of Bunsen-Roscoe law and to quantify the value of Schwarzschild's p -coefficients based on the photochemomechanical responses as defined by increases of (A) stiffness (ΔM_Y) and (B) strength ($\Delta \sigma$) in the irradiated tarso-conjunctival specimens exposed to UVA radiation (365 nm). Tissue specimens were excised from goat eyelid. The measurements were completed at a radiant exposure of 13.5 J/cm^2 . Each data point value is an average of 8 measurements $\pm$ s.e.m. Zone I – validity zone; zone II – failure zone. The inclination angle (θ) is indicated for each plotted line.

When Young's modulus (M_Y) and stress (σ), measured in MPa units, were selected as the target photoresponses, the resulting plots are shown in Fig. 4, where the calculated percent increase of either M_Y (Fig. 4A) or σ (Fig. 4B) were plotted against irradiance at a radiant exposure of 13.5 J/cm^2 .

4 Discussion

To deal with a rather confusing amalgam of terms in current use for the quantities employed in radiometry, photometry, and astrophotography, we followed here the radiometric terminology and symbols recommended by the International Standard ISO 80000–7:2019 [35]. Terms such as 'fluence' and 'energy dose' for *radiant exposure* (H_e , J/m^2), or 'intensity', 'flux', 'flux density',

'fluence rate' and 'illuminance' for *irradiance* (E_e , W/m^2) are no longer recommended and should be avoided or used only in specified circumstances. In fairness, when perusing photochemical or photobiological publications, it seems unlikely that those terms will be abandoned soon [36].

4.1 Significance of Bunsen-Roscoe/Schwarzschild Law

Radiation-induced photochemical and photobiological processes and their physicochemical effects on matter are due entirely to the total amount of electromagnetic radiant energy absorbed per unit area by the material target, i.e., the radiant exposure [37–39]. A photoresponse (symbolized by $\mathcal{P}$) is dependent on the

product of the two individual quantities that actually define the radiant exposure itself, namely the irradiance (E_e) and the duration of exposure (t). In principle, this can be expressed by Eq. (1).

$$\mathcal{P} = f(xy), \text{ where } x = E_e \text{ and } y = t. \quad (1)$$

Furthermore, the same response will be observed when the target receives the same total radiant exposure, irrespective of how the energy was delivered; it is the total absorbed energy that actually matters, not the method of exposure. By either applying a higher irradiance for a shorter duration, or a lower irradiance for a longer duration, the same effect of the irradiation shall be achieved. The law can be expressed mathematically through any of the inverse variation relationships in Eq. (2a) or (2b). In other words, when targeting a certain photoresponse in an irradiated system, one variable quantity, either E_e or t , can be expeditiously changed while modifying the other one in order to maintain constant the product of their multiplication.

$$E_e \propto 1/t, \quad (2a)$$

$$E_{ei} \times t_i = \text{const}, \forall i \in \{1, 2, \dots, n\}. \quad (2b)$$

In essence, as long as the multiplication product of irradiance and exposure time is the same, the photoresponse shall be the same. This law is known as Bunsen-Roscoe law (henceforth, BRL), acknowledging the two 19th century scientists who demonstrated it experimentally using photographic films [40, 41], and is alternatively known as the ‘exposure reciprocity law’. The law also stipulates that a certain radiant exposure will remain the same regardless whether the duration of exposure is continuous or interrupted within the same experiment [42, 43]. While BRL is seldom regarded as one of the fundamental laws of photochemistry [44, 45], alongside Grotthuss-Draper law and photochemical equivalence law (known also as Stark-Einstein law) [37, 39], it can, however, be interpreted as a quantitative expression of the former [38]. BRL is also applied in vision science for investigating how the eye perceives very short flashes of light, where is known as Bloch’s law, named after the scientist who was the first to relate the concept to vision [46], apparently without being aware of the existence of BRL. The role and limitations of BRL and Bloch’s law in eye perception have been investigated to some extent [47–51].

The above statements, as well as the relationships in Eqs. (2a) and (2b) themselves, are only valid as long as the experimental data comply with the law. However, failures of BRL have been noted in the past [52–62], and are still reported [43, 63–67]. In general, the reciprocity law holds for photochemical processes, with deviations noted usually at extreme levels of irradiance. In irradiated biological systems, the failure of BRL is expected to be more significant, considering that the responses of cells and tissues involve complex interacting mechanisms,

photoreactivation repair processes, and involvement of photosensitizing biomolecules present in tissues [37, 65].

To rationalize the failure of BRL, Schwarzschild developed an empirical model based on his work in astrophotography [53]. The model was further improved [54–60, 62] to match closer certain experimental findings. In its new form, known also as Schwarzschild law, BRL can be generally expressed by the Eqs. (3), where p or q are known as Schwarzschild’s coefficients (or exponents), and $p \times q = 1$.

$$E_e \times t^p = \text{const}, \quad (3a)$$

$$E_e^q \times t = \text{const}. \quad (3b)$$

If p (or q) = 1, the two laws are identical, and therefore Schwarzschild law can be considered a generalization of BRL, and it is adequate for monitoring the reciprocity. Other proposed models display additional mathematical complexity and most of them are related to the photographic field [43, 60, 63]. Closer to our times, the interest in BRL failure has expanded to other fields such as dermatology, biology, photomedicine, photocatalysis, and photodegradation of materials. In a seminal review [43], a survey of the literature on BRL failure, including estimated values for Schwarzschild’s coefficients (where available), indicated that the reported cases of failure in biology and medicine were actually less than those in other fields. The quantum band theory [68, 69] has been episodically employed as a theoretical model able to explain the variation of Schwarzschild’s p -coefficients in certain photoprocesses, but is limited mostly to photoconductance and photography studies, and is based entirely on photophysical rather than photochemical processes [43, 60, 70–74].

In the biomedical applications of electromagnetic radiation, it is important to understand the limits to validity of BRL, which will allow suitable irradiant strategies. In practice, formulae such as in Eqs. (4) are frequently employed in order to calculate the radiant exposure that results from a certain irradiance applied over a given exposure time; nevertheless, they yield workable values only if used within the confines of BRL’s validity, i.e., outside the failure zones.

$$H_e \left[\frac{J}{m^2} \right] = E_e \left[\frac{W}{m^2} \right] \times t[s], \quad (4a)$$

$$H_e \left[\frac{J}{cm^2} \right] = E_e \left[\frac{mW}{cm^2} \right] \times t[s] \times 10^{-3}, \quad (4b)$$

$$H_e \left[\frac{J}{cm^2} \right] = E_e \left[\frac{mW}{cm^2} \right] \times t[\text{min}] \times 0.06. \quad (4c)$$

When selecting a preferred irradiation regime in a clinical setting, the ophthalmic surgeon may choose to apply irradiances (E_e) and/or exposure times (t) falling outside the law’s validity zone, either on purpose or unintentionally. While such a choice does not necessarily put at risk the safety of the procedure, values of the radiant exposure (H_e) estimated with Eqs. (4) will be

different from the actual values, therefore outside surgeon's cognizance or control.

In proposing our procedure as a potential treatment for eyelid laxity conditions, we took as a model the UVA-induced CXL, an established treatment for keratoconus that has been the subject of over 2000 studies published by the year 2020 [75]. Foreseeably, CXL was the application that aroused an interest in BRL's role in ophthalmology. That was mainly a consequence of certain innovative developments in the procedure, including accelerated protocols [25, 32, 76–84], the use of high radiant exposures (fluences) [25, 77, 78, 85–88], and supplementation with oxygen during procedure [25, 89–93]. Regarding the radiant exposure, 20 J/cm² was the highest reported value in *ex vivo* animal corneas [87], and 12 J/cm² for treating human patients [85]. The most complete study on the relation between irradiation parameters and validity range of BRL has been performed [78] on excised porcine corneas, where the tensile elasticity modulus (or Young's modulus, a quantitative measure of stiffness) was monitored as a photoresponse. Ten different pairs of E_e and t , each generating the same radiant exposure ($H_e = 5.4 \text{ J/cm}^2$), were applied starting with $E_e = 3 \text{ mW/cm}^2$ and $t = 30 \text{ min}$ and ending with $E_e = 90 \text{ mW/cm}^2$ and $t = 1 \text{ min}$. It was found [78] that BRL maintains validity only for irradiances around 45 mW/cm² and lower, and at exposure durations longer than 2 min, as indicated by measuring the changes in stiffness. Regarding the role of oxygen during irradiation, based on the mechanism and kinetics of UVA/riboflavin-induced CXL [89], its presence appears to be beneficial to the photochemical reactions leading to formation of crosslinks. While hyperoxic conditions resulted indeed in more efficient crosslinking in *ex vivo* animal corneas [90, 92, 93], the absence of any effect has been also reported [91]. Predictably, the role of oxygen shall be relevant in CXL procedures, and its participation 'may be the missing piece to understand the invalidity of Bunsen-Roscoe's law' [93].

4.2 Validation of BRL Law Based on Photoresponses

Failure of reciprocity law has been extensively investigated in the photographic field, where graphical methods were developed to identify the failure region [43, 55–58, 60, 63]. Such methods are generally based on monitoring the irradiance (E_e) and its product with the duration of exposure ($E_e \times t$) for a fixed photoresponse, i.e., for a given state of transformation induced by irradiation. Experimentally, this is easy to achieve in photographic research, as it implies no more than a facile darkening of photosensitive films or plates. Outside the photographic field, however, it is rather difficult to find a photoresponse amenable to being maintained at a fixed value while the other parameters (irradiance, time) shall be varied and monitored until they attain the values needed to generate that particular photoresponse. In our present study, we have used a graphical method whereby a measurable photoresponse

was plotted against irradiance for a chosen radiant exposure. The plot is on a logarithmic scale, allowing data points to fall on the same line at each exposure value [43]. Although less complicated, the method has been only episodically employed to investigate BRL failure, for instance in studies involving target photoresponses such as intensity of photocurrents generated in dyes by exposure to light [94], photovoltage generated by exposure to UV radiation of titanium dioxide films [95], or rate of isopropanol UV-induced photooxidation [96].

According to this method, a change in the slope of the logarithmically linearized plot $\log \mathcal{P} = f(\log E_e)$ indicates a discontinuity in the validity of BRL, i.e., where the law breaks down. In the linearized power function $y = ax^m$, the exponent m represents the slope. Furthermore, the slope is the equivalent of Schwarzschild's p -coefficient [43], and the closer its value is to 1 (for an inclination angle $\theta = 45^\circ$), the lower is the level of failure.

The target tissue in our study was the tarsal conjunctiva of the eyelid in two bovid species (sheep and goat), which we have exposed to UVA radiation (365 nm). Two different photoresponses were considered upon exposure to selected levels of irradiance: (a) the increase of the local temperature on the conjunctiva surface at the site whereupon radiation impinges and its energy is transferred to the target; and (b) the irradiation-induced increase of stiffness (expressed through the values of Young's modulus) or increase of strength (expressed by the measured stress), both recorded during experimental stretching of the tarso-conjunctival substrate, prior to and after crosslinking.

Fig. 3 shows that at any of the tested levels of exposure based on the photothermal response, a drop in the slopes (i.e., the Schwarzschild's p -coefficients) occurred within an irradiance region approximately between 25 and 40 mW/cm² (shaded areas) in each of the two animal tissues, with 30 mW/cm² being a suitable and safe estimation for a practicable limiting value. This region divided the plot into two zones, validity zone (I) and failure zone (II). A variation in the values of p -coefficients shall be interpreted as a change in the law's validity, in this case due to thermal effects, and can be regarded as the commencement of law's failure in zone II. Table 1 summarizes the actual values of p -coefficients in each tissue at three different radiant exposures, as extracted from the equations plotted in Fig. 3. The difference between coefficients in the two zones is compelling, clearly suggesting breakdowns in BRL's validity.

Fig. 4 shows the plots of crosslinking-induced mechanical augmentation as reflected by increases of Young's modulus and stress, i.e., stiffness and strength, respectively. Since the results involving three different exposure levels of two different tissues showed very similar patterns (as seen in Fig. 3), we selected only one level ($H_e = 13.5 \text{ J/cm}^2$) for investigating the functional relationship between E_e and M_Y or between E_e and σ , and investigated only the goat tarso-conjunctival tissue.

Table 1 Values of Schwarzschild's p -coefficients for the photothermal responses.

Eyelid Tissue	Sheep		Goat	
Zone-specific plots ^a	I	II	I	II
Radiant exposure (H_e , J/cm ²)	p -Coefficients ^b			
6	0.3541	0.0543	0.5608	0.2384
13.5	0.5714	0.1324	0.6382	0.2401
18	0.5247	0.1659	0.6740	0.2321

^a Linear zones characterized by two different slopes as shown in Fig. 3A and 3B. Zone II represents the law's failure zone.

^b Schwarzschild's p -coefficients are identical with the slopes m in the linearized power functions ($y=ax^m$) plotted in Fig. 3A and 3B.

A dramatic decrease of the p -coefficients occurs in the region approximately between 50 and 70 mW/cm², furthermore attaining negative values in zone II. An acceptable limiting value of 60 mW/cm² for the validity zone seems appropriate. It can be seen that the effects of UVA on mechanical properties of the tarsus led to higher irradiance levels that marked the BRL failure as compared to those indicated by thermal response.

A question arises as to why the discontinuity in the slope, marking the failure of BRL, was more pronounced when the monitored photoresponse was the enhancement of either the mechanical stiffness (ΔM_Y) or of the built-up force ($\Delta\sigma$) during stretching of the tissue, i.e., a photochemomechanical response, as compared to the irradiation-induced temperature increase at the surface of the tissue, i.e., a photothermal response. To attempt an explanation we should briefly address what is known about the effects of exposing materials to UVA radiation, which is considered a non-ionizing and non-thermal radiation. The photoresponses in multiparticulate systems are the result of time-directed, cause-and-effect processes, where the UVA radiation is the causal agent. Like with any electromagnetic radiation, such processes induce photophysical and photochemical changes. The former are single-particle events that include photon absorption, localized excitation, as well as generation and migration of charge carriers (e.g., excited electrons), and they are always the same irrespective of the nature of irradiated material substrate. Importantly, the photophysical processes are precursors to photochemical changes. The latter are intricate and unpredictable, as they depend on both chemical structure and homogeneity of the irradiated substrates. Upon their absorption by the target, the UVA photons possess sufficient energy to excite molecules, which are then able to trigger chemical processes causing significant transformations in

biological matter including lipid oxidation, structural alteration of nucleic acids, and crosslinking of native proteins. The excited electrons that did not contribute to such changes can release their energy through luminescence phenomena (fluorescence, phosphorescence), or through production of heat. Notably, the amount of heat generated by UVA radiation is generally low. Although some of the released heat can also be a result of exothermic chemical reactions activated by the radiation itself, such contribution was minor in our experiments because no photosensitizer was present to initiate chemical reactions during evaluation of the photothermal response. That was clearly illustrated by our measured ΔT values indicating a weak response. Moreover, we have found in a previous study [34] that even in the presence of a photosensitizer the thermal effect of UVA on tarso-conjunctival tissue was minor at irradiances lower than 75 mW/cm². If we further consider a baseline temperature higher than that employed in the present study, such as that reported for human tarsal conjunctiva (*ca* 33°C) [97], it is still unlikely that UVA radiation could induce a surface terminal temperature above 40°C, the well-known safe limit for the thermal stability of biological macromolecules [98–101].

The photochemomechanical response, ultimately manifested in changes of modulus M_Y and stress σ , was the result of a photochemical reaction (crosslinking of collagen) concomitantly followed by a process of mechanical augmentation of the tarsal tissue. Crosslinking-induced enhancement of collagen's mechanical properties is a complex process. Molecular models have been developed [102–105] in order to understand the relation between crosslink density and deformation mechanisms. A crucial role of crosslinks is to assure interconnectivity between fibrillary collagen molecules, essential for enhancing the network's strength and toughness. Insufficient crosslinks cause intermolecular sliding resulting in weak structural networks. Chemistry of the covalent crosslinkages is also relevant: for the same crosslink density, the networks based on trivalent crosslinks are stronger than those based on divalent crosslinks, due to different interconnectivity levels. The fully crosslinked collagen networks undergo deformation in synergy with relatively minor intermolecular sliding and macromolecular uncoiling but can stand up to external forces that otherwise would cause the mechanical failure of uncrosslinked networks. At a certain degree of deformation, however, the crosslinks may start to break down; in principle, additional crosslinking may mitigate the process if given the opportunity.

We suggest that due to an intrinsic complexity of the underlying mechanisms leading to reinforcement of the tissue, the photochemomechanical response is delayed, leading to an apparent extension of BRL's validity zone, at the same time causing a dramatically more significant discontinuity in the slope (Fig. 4) than that seen when a localized surface-level temperature increase was the intended photoresponse (Fig. 3).

It remains to consider which of the two investigated photoresponses shall be selected as a criterion for defining the BRL's failure zone and for recommending practical values for the irradiation parameters aimed at preventing the use of an irradiant profile different from the actual profile. Although the photochemomechanical response is associated with major structural changes in the biological substrate and would therefore appear to be more relevant as a marker for the onset of BRL's failure, we propose the photothermal response as a criterion, notwithstanding that the thermal effects induced by UVA radiation on the tarsal conjunctiva are minor. Our reasons are mainly of a biological safety nature. Thus, the principle of the safest application route was a reason for recommending limiting irradiances around 30 mW/cm² instead 60 mW/cm², as less power delivered to the tissue means less potential damage. Any accidental thermal damage will be more devastating and faster than that produced by the photochemical crosslinking. Another reason is based on the vast published experience with CXL, which clearly indicates that there is no need for irradiances higher than 45 mW/cm², while maintaining fluences under 20 J/cm², to be applied for crosslinking the corneal stromal collagen. In reality, even lower values are currently used in most clinics around the world. Moreover, all UVA devices for CXL currently on the market (a 1.2-billion-dollar market) are limited to deliver a maximum irradiance of 45 mW/cm². For purely safe practicality, we believe that such values are also suitable for the therapeutic irradiation of the tarsal conjunctiva, and there is no reason to apply irradiances near 60 mW/cm² as suggested by the experimental photochemomechanical response in our study.

5 Conclusion

By applying a graphical method based on the relation between irradiance and certain photoresponses, the validity interval of Bunsen-Roscoe law can be established and employed to select suitable conditions for exposing the tarso-conjunctival tissue to UVA radiation for therapeutic purposes. When the local radiation-induced temperature increase at the conjunctival surface was chosen as a photothermal response, the method indicated failure of the law occurring at an irradiance of approximately 30 mW/cm² for a radiant exposure (fluence) between 6 and 18 J/cm². When changes in the mechanical properties of the tissue (stiffness, strength) were investigated as photochemomechanical responses, the law broke down at around 60 mW/cm². However, following the safest application route principle, a biologically safe performance of the procedure should be prioritized over other considerations. Therefore, based on the law's validity limit disclosed by the analysis of photothermal response, an irradiance less than 30 mW/cm², associated with exposure times able to assure a radiant exposure within an estimative range of 5 to 20 J/cm², are recommended for crosslinking the tarsal collagen.

Symbols and Abbreviations

ΔT – temperature increase;
 E_e – irradiance;
 H_e – radiant exposure;
 m – slope ($\tan \theta$);
 M_Y – Young's modulus;
 p – Schwarzschild's coefficient;
 $\mathcal{P}$ – photoresponse;
 R^2 – coefficient of determination (for linear regression);
 σ – stress;
 t – time;
 θ – inclination angle;
 BRL – Bunsen-Roscoe law;
 CXL – corneal collagen crosslinking;
 PBS – phosphate-buffered solution;
 RF-5P – riboflavin 5'-phosphate monosodium salt;
 s.e.m. – standard error of the mean;
 TXL – tarsal collagen crosslinking;
 UV – ultraviolet radiation;
 UVA – region A of ultraviolet radiation (wavelength from 315 to 400 nm).

Author Contributions

T.V.C.: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Supervision, Project administration; S.S.: Methodology, Investigation, Resources, Data curation, Writing – review & editing; N.E.P.: Methodology, Investigation, Resources; C.C.: Methodology, Investigation; A.I.M.: Methodology, Investigation, Writing – review & editing. All authors contributed to the experiments, checked the final draft, and approved its submission. The authors assume full responsibility for the results reported in this article.

Ethics Approvals, Consent for Publication

None applicable.

Data Sharing and Availability

The data supporting the findings of this study are provided within the article. Additional information is available upon reasonable request addressed to the corresponding author.

Funding

This research did not receive any grant from funding agencies in the public or commercial sector, or from sponsorships.

Disclosures

All authors declare that there is no conflict of interest in this paper.

References

1. P. A. Netland, S. P. Sugrue, D. M. Albert, and J. W. Shore, "Histopathologic features of the floppy eyelid syndrome," *Ophthalmology* 101(1), 174–181 (1994).
2. U. Schlotzer-Schrehardt, M. Stojkovic, C. Hofmann-Rummelt, C. Cursiefen, F. Kruse, and L. Holbach, "The pathogenesis of floppy eyelid syndrome involvement of matrix metalloproteinases in elastic fiber degradation," *Ophthalmology* 112(4), 694–704 (2005).
3. A. M. Fowler, J. J. Dutton, "Floppy eyelid syndrome as a subset of lax eyelid conditions: relationships and clinical relevance (an ASOPRS thesis)," *Ophthalmic Plastic & Reconstructive Surgery* 26(3), 195–204 (2010).
4. D. G. Ezra, M. Beaconsfield, M. Sira, C. Bunce, R. Wormald, and R. Collin, "The associations of floppy eyelid syndrome: A case control study," *Ophthalmology* 117(4), 831–838 (2010).
5. D. G. Ezra, J. S. Ellis, C. Gaughan, M. Beaconsfield, R. Collin, C. Bunce, M. Bailly, and P. Luthert, "Changes in tarsal plate fibrillar collagens and elastic fibre phenotype in floppy eyelid syndrome: Changes suggest mechanical adaptation," *Clinical & Experimental Ophthalmology* 39(6), 564–571 (2011).
6. R. W. Damasceno, M. H. Osaki, P. E. C. Dantas, and R. Belfort, "Involitional ectropion and entropion: clinicopathologic correlation between horizontal eyelid laxity and eyelid extracellular matrix," *Ophthalmic Plastic & Reconstructive Surgery* 27(5), 321–326 (2011).
7. T. M. Smith, S. Suzuki, B. G. Cronin, M. Haghighatpanah, E. B. Petcu, C. J. Philippa, and T. V. Chirila, "Photochemically induced crosslinking of tarsal collagen as a treatment for eyelid laxity: Assessing potentiality in animal tissue," *Ophthalmic Plastic & Reconstructive Surgery* 34(5), 477–482 (2018).
8. T. M. Smith, S. Suzuki, N. Sabat, C. L. Rayner, D. G. Harkin, and T. V. Chirila, "Further investigations on the crosslinking of tarsal collagen as a treatment for eyelid laxity: optimizing the procedure in animal tissue," *Ophthalmic Plastic & Reconstructive Surgery* 35(6), 600–603 (2019).
9. A. I. Manta, N. E. Pop, R. G. Tripon, F. Vultur, S. Suzuki, B. A. Cordos, C. C. Radu, T. Hoge, C. Carasca, K. U. Horvath, G. A. Muntean, V. C. Siserman, O. S. Cotoi, M. H. B. Radford, and T. V. Chirila, "Photochemical crosslinking of tarsal collagen as a treatment for eyelid laxity: evaluation in *ex vivo* human tissue," *Ophthalmic Plastic & Reconstructive Surgery* 41(1), 28–35 (2025).
10. S. S. Akella, J. Liu, Y. Miao, R. S. Chuck, A. Barmettler, and C. Zhang, "Collagen structural changes in rat tarsus after crosslinking," *Translational Vision Science & Technology* 10(5), 3 (2021).
11. S. Ugradar, A. Le, M. Lesgart, R. A. Goldberg, D. Rootman, and J. L. Demer, "Biomechanical and morphologic effects of collagen cross-linking in human tarsus," *Translational Vision Science & Technology* 8(6), 25 (2019).
12. S. Ugradar, J. Karlin, A. Le, J. Park, and R. A. Goldberg, "Photochemical collagen cross-linking reverses elastase-induced mechanical degradation of upper eyelid tarsus," *Ophthalmic Plastic & Reconstructive Surgery* 36(6), 562–565 (2020).
13. A. M. Kocer, E. M. Sen, M. Caydere, S. Yenigun, and S. Hucumenoglu, "The histopathological findings in excised upper eyelids of patients with dermatochalasis following collagen cross-linking treatment," *Graefe's Archive for Clinical and Experimental Ophthalmology* 260, 2737–2743 (2022).
14. R. M. Lavker, G. F. Gerberick, D. Veres, C. J. Irwin, and K. H. Kaidbey, "Cumulative effects from repeated exposures to suberythral doses of UVB and UVA in human skin," *Journal of the American Academy of Dermatology* 32(1), 53–62 (1995).
15. C. Battie, S. Jitsukawa, F. Bernerd, S. Del Bino, C. Marionnet, and M. Verschoore, "New insights in photoaging, UVA induced damage and skin types," *Experimental Dermatology* 23(s1), 7–12 (2014).
16. K. J. Gromkowska-Kępa, A. Puścion-Jakubik, R. Markiewicz-Żukowska, and K. Socha, "The impact of ultraviolet radiation on skin photoaging — review of *in vitro* studies," *Journal of Cosmetic Dermatology* 20(11), 3427–3431 (2021).
17. F. Bernerd, T. Passeron, I. Castiel, and C. Marionnet, "The damaging effects of long UVA (UVA1) rays: A major challenge to preserve skin health and integrity," *International Journal of Molecular Sciences* 23(15), 8243 (2022).
18. J.-W. Yang, G.-B. Fan, F. Tan, H.-M. Kong, Q. Liu, Y. Zou, and Y.-M. Tan, "The role and safety of UVA and UVB in UV-induced skin erythema," *Frontiers in Medicine* 10, 1163697 (2023).
19. C. Remé, J. Reinboth, M. Clausen, and F. Hafezi, "Light damage revisited: converging evidence, diverging views?," *Graefe's Archive for Clinical and Experimental Ophthalmology* 234(1), 2–11 (1996).
20. R. D. Glickman, "Phototoxicity to the Retina: Mechanisms of damage," *International Journal of Toxicology* 21(6), 473–490 (2002).
21. M. S. Oliva, H. Taylor, "Ultraviolet radiation and the eye," *International Ophthalmology Clinics* 45(1), 1–17 (2005).
22. I. V. Ivanov, T. Mappes, P. Schaupp, C. Lappe, and S. Wahl, "Ultraviolet radiation oxidative stress affects eye health," *Journal of Biophotonics* 11(7), e201700377 (2018).
23. F. Raiskup, E. Spoerl, "Corneal crosslinking with riboflavin and ultraviolet A. I. Principles," *The Ocular Surface* 11(2), 65–74 (2013).

24. R. S. Rubinfeld, C. Caruso, and C. Ostacolo, “[Corneal cross-linking: The science beyond the myths and misconceptions](#),” *Cornea* 38(6), 780–790 (2019).
25. B. J. Blackburn, A. M. Rollins, and W. J. Dupps, “[Biomechanics of ophthalmic crosslinking](#),” *Translational Vision Science & Technology* 10(5), 8 (2021).
26. M. R. Santhiago, J. B. Randleman, “[The biology of corneal cross-linking derived from ultraviolet light and riboflavin](#),” *Experimental Eye Research* 202, 108355 (2021).
27. D. Wu, D. K.-A. Lim, B. X. H. Lim, N. Wong, F. Hafezi, R. Manotosh, and C. H. L. Lim, “[Corneal cross-linking: The evolution of treatment for corneal diseases](#),” *Frontiers in Pharmacology* 12, 686630 (2021).
28. L. Angelo, A. Gokul, C. McGhee, and M. Ziaei, “[Corneal crosslinking: Present and future](#),” *Asia-Pacific Journal of Ophthalmology* 11(5), 441–452 (2022).
29. N. Papachristoforou, A. Ueno, K. Ledwos, J. Bartuś, A. Nowińska, and I. Karska-Basta, “[A review of keratoconus cross-linking treatment methods](#),” *Journal of Clinical Medicine* 14(5), 1702 (2025).
30. F. Hafezi, S. Kling, N. L. Hafezi, M. E. Aydemir, N.-J. Lu, M. Hillen, B. Knyazer, S. Awwad, C. Mazzotta, L. Kollros, and E. A. Torres-Netto, “[Corneal cross-linking](#),” *Progress in Retinal and Eye Research* 104, 101322 (2025).
31. N. Sorkin, D. Varssano, “[Corneal collagen crosslinking: A systematic review](#),” *Ophthalmologica* 232(1), 10–27 (2014).
32. G. S. Sachdev, M. Sachdev, “[Recent advances in corneal collagen cross-linking](#),” *Indian Journal of Ophthalmology* 65(9), 787–796 (2017).
33. N. E. Pop, S. Suzuki, R. G. Tripon, A. I. Manta, K. M. Lee, M. Parisot, A. G. Cocisiu, M. H. B. Radford, D. G. Harkin, and T. V. Chirila, “[Photochemical crosslinking of eyelid tissue by trans-conjunctival exposure to UV-A radiation: implications for biological safety](#),” *Journal of Biophotonics* 19(1), e202500458 (2026).
34. S. Suzuki, N. E. Pop, A. I. Manta, and T. V. Chirila, “[Ex vivo thermographic analysis of the eyelid conjunctival surface during ultraviolet A-mediated photochemical crosslinking of tarsal collagen: A step in translational safety evaluation](#),” *Global Translational Medicine* 5(1), 87 (2026).
35. International Standard ISO 80000–7:2019 (Quantities and units – Part 7: Light and radiation), 2nd Ed., BSI Standards Ltd, London (2019).
36. J. R. Bolton, I. Mayor-Smith, and K. G. Linden, “[Rethinking the concepts of fluence \(UV Dose\) and fluence rate: the importance of photon-based units – a systemic review](#),” *Photochemistry and Photobiology* 91(6), 1252–1262 (2015).
37. K. C. Smith, “Basic photochemistry,” *Photobiological Sciences Online* (accessed 3 November 2025). [www.photobiology.info/Photochem.html].
38. A. Albini, “[The framework of photochemistry: The laws](#),” *Photochemistry: Past, Present and Future*, Springer-Verlag, Berlin, 9–40 (2016).
39. H. E. Bonfield, T. Knauber, F. Lévesque, E. G. Moschetta, F. Susanne, and L. J. Edwards, “[Photons as a 21st century reagent](#),” *Nature Communications* 11(1), 804 (2020).
40. R. Bunsen, H. Roscoe, “[Photochemische untersuchungen. Fünfte Abhandlung](#),” *Annalen der Physik* 184(10), 193–273 (1859).
41. R. W. Bunsen, H. E. Roscoe, “[Photochemical researches.—Part V. On the measurement of the chemical action of direct and diffuse sunlight](#),” *Proceedings of the Royal Society of London* (12), 306–312 (1863).
42. P. D. Forbes, R. E. Davies, and F. Urbach, “[Aging, environmental influences, and photocarcinogenesis](#),” *Journal of Investigative Dermatology* 73(1), 131–134 (1979).
43. J. W. Martin, J. W. Chin, and T. Nguyen, “[Reciprocity law experiments in polymeric photodegradation: a critical review](#),” *Progress in Organic Coatings* 47(3–4), 292–311 (2003).
44. A. Albini, “[Some remarks on the first law of photochemistry](#),” *Photochemical & Photobiological Sciences* 15(3), 319–324 (2016).
45. Y. Markushyna, A. Savateev, “[Light as a tool in organic photocatalysis: multi-photon excitation and chromoselective reactions](#),” *European Journal of Organic Chemistry* 2022(24), e202200026 (2022).
46. A.-M. Bloch, “Expériences sur la vision,” *Comptes Rendus de la Société de Biologie* 37, 493–495 (1885).
47. E. D. Adrian, R. Matthews, “[The action of light on the eye: Part II. The processes involved in retinal excitation](#),” *The Journal of Physiology* 64(3), 279–301 (1927).
48. G. S. Brindley, “[The Bunsen-Roscoe law for the human eye at very short durations](#),” *The Journal of Physiology* 118(1), 135–139 (1952).
49. R. A. Rosenblum, “[Electroretinographic evaluation of the Bunsen-Roscoe law for the human eye at high energy levels](#),” *Investigative Ophthalmology* 10(11), 904–910 (1971).
50. E. Greene, R. T. Ogden, “[Shapes displayed with durations in the microsecond range do not obey bloch’s law of temporal summation](#),” *i-Perception* 4(6), 429–436 (2013).
51. E. Greene, “[Violation of bloch’s law that specifies reciprocity of intensity and duration with brief light flashes](#),” *i-Perception* 4(8), 543–550 (2013).

52. W. D. W. Abney, “IX. On a failure of the law in photography that when the products of the intensity of the light acting and of the time of exposure are equal, equal amounts of chemical action will be produced,” *Proceedings of the Royal Society of London* 54(326–330), 143–147 (1894).
53. K. Schwarzschild, “On the deviations from the law of reciprocity for bromide of silver gelatine,” *The Astrophysical Journal* 11, 89 (1900).
54. S. E. Sheppard, C. E. K. Mees, *Investigations on the Theory of the Photographic Process*, Longmans Green & Co, London, 214–224 (1907).
55. E. Kron, “Über das Schwärzungsgesetz photographischer Trockenplatten,” *Annalen der Physik* 346(9), 751–758 (1913).
56. J. Halm, “On the determination of fundamental photographic magnitudes,” *Monthly Notices of the Royal Astronomical Society* 75(3), 150–177 (1915).
57. L. A. Jones, “Photographic spectrophotometry in the ultra-violet region,” *Eos, Transactions, American Geophysical Union* 8, 109–123 (1927).
58. H. Arens, J. Eggert, “Über die Schwärzungsfläche photographischer Schichten,” *Zeitschrift für Physikalische Chemie* 131U(1), 297–309 (1927).
59. J. H. Webb, “The effect of temperature upon reciprocity law failure in photographic exposure,” *Journal of the Optical Society of America A* 25(1), 4 (1935).
60. C. E. K. Mees, *The Theory of the Photographic Process*, Macmillan Co, New York, 234–254 (1942).
61. R. H. Morgan, “Reciprocity law failure in X-ray films,” *Radiology* 42(5), 471–479 (1944).
62. J. H. Webb, “Low intensity reciprocity-law failure in photographic exposure: Energy depth of electron traps in latent-image formation; Number of quanta required to form the stable sublatent image,” *Journal of the Optical Society of America A* 40(1), 3 (1950).
63. J. Chin, T. Nguyen, E. Byrd, and J. Martin, “Validation of the reciprocity law for coating photodegradation,” *Journal of Coatings Technology and Research* 2(7), 499–508 (2005).
64. A. Schindl, B. Rosado-Schlosser, and F. Trautinger, “Die reziprozitätsregel in der photobiologie,” *Der Hautarzt* 52(9), 779–785 (2001).
65. A. Schindl, B. Rosado-Schlosser, and F. Trautinger, “The reciprocity rule in photobiology – a review,” *Internet Photochemistry & Photobiology*, Emedia Science Ltd (accessed 30 September 2025). [<http://www.photobiology.com/reviews/bunsen/index.htm>].
66. J. G. Leprince, M. Hadis, A. C. Shortall, J. L. Ferracane, J. Devaux, G. Leloup, and W. M. Palin, “Photoinitiator type and applicability of exposure reciprocity law in filled and unfilled photoactive resins,” *Dental Materials* 27(2), 157–164 (2011).
67. L. Feng, B. I. Suh, “Exposure reciprocity law in photopolymerization of multi-functional acrylates and methacrylates,” *Macromolecular Chemistry & Physics* 208(3), 295–306 (2007).
68. D. Halliday, *Fundamentals of Physics*, vol. 2, 11th Ed., Wiley, Hoboken, NJ, 1252–1275 (2019).
69. L. Colombo, “The band theory”, Chapter 8 in *Solid State Physics: A Primer*, IOP Publishing, London (2021).
70. J. T. Randall, M. H. F. Wilkins, “The phosphorescence of various solids,” *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 184(999), 347–364 (1945).
71. J. T. Randall, M. H. F. Wilkins, “Phosphorescence and electron traps - I. The study of trap distributions,” *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 184(999), 365–389 (1945).
72. A. Rose, “An outline of some photoconductive processes,” *RCA Review* 12, 362–414 (1951).
73. J. F. Fowler, “X-ray induced conductivity in insulating materials,” *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 236(1207), 464–480 (1956).
74. W. Arnold, H. K. Sherwood, “Are chloroplasts semiconductors?,” *Proceedings of the National Academy of Sciences* 43, 105–114 (1957).
75. K. Yang, L. Xu, S. Wang, M. Zhu, Q. Fan, Y. Gu, Y. Wang, Q. Wang, D. Zhao, C. Pang, and S. Ren, “A bibliometric analysis of 100 most-cited articles on corneal cross-linking,” *Frontiers in Medicine* 9, 904077 (2022).
76. S. Schumacher, L. Oeftiger, and M. Mrochen, “Equivalence of biomechanical changes induced by rapid and standard corneal cross-linking, using riboflavin and ultraviolet radiation,” *Investigative Ophthalmology & Visual Science* 52(12), 9048 (2011).
77. E. Lanchares, M. A. Del Buey, J. A. Cristóbal, L. Lavilla, and B. Calvo, “Biomechanical property analysis after corneal collagen cross-linking in relation to ultraviolet A irradiation time,” *Graefes’s Archive for Clinical and Experimental Ophthalmology* 249(8), 1223–1227 (2011).
78. J. Wernli, S. Schumacher, E. Spoerl, and M. Mrochen, “The efficacy of corneal cross-linking shows a sudden decrease with very high intensity UV light and short treatment time,” *Investigative Ophthalmology & Visual Science* 54(2), 1176 (2013).
79. I. M. Beshtawi, R. Akhtar, M. C. Hillarby, C. O’Donnell, X. Zhao, A. Brahma, F. Carley, B. Derby, and H. Radhakrishnan, “Biomechanical properties of human corneas following low- and high-intensity collagen cross-linking determined with scanning acoustic microscopy,” *Investigative Ophthalmology & Visual Science* 54(8), 5273 (2013).

80. A. Hammer, O. Richoz, S. A. Mosquera, D. Tabibian, F. Hoogewoud, and F. Hafezi, “[Corneal biomechanical properties at different corneal cross-linking \(CXL\) irradiances](#),” *Investigative Ophthalmology & Visual Science* 55(5), 2881 (2014).
81. E. B. Ozgurhan, B. I. Sezgin Akcay, Y. Yildirim, G. Karatas, T. Kurt, and A. Demirok, “[Evaluation of corneal stromal demarcation line after two different protocols of accelerated corneal collagen cross-linking procedures using anterior segment optical coherence tomography and confocal microscopy](#),” *Journal of Ophthalmology* 2014, 1–5 (2014).
82. J. H. Woo, J. V. Iyer, L. Lim, M. H. Hla, J. S. Mehta, C. M. Chan, and D. T. Tan, “[Conventional versus accelerated collagen cross-linking for keratoconus: A comparison of visual, refractive, Topographic and Biomechanical Outcomes](#),” *The Open Ophthalmology Journal* 11(1), 262–272 (2017).
83. M.-S. Dina, M.-M. Constantin, M.-C. Marinescu, C.-G. Corbu, C.-I. Tataru, and C.-P. Tataru, “[Standard and accelerated crosslinking protocols in keratoconus – differences and evolution at one year](#),” *Romanian Journal of Ophthalmology* 69, 175–183 (2025).
84. R. E. T. Ang, A. J. A. Guloy, E. M. Cruz, and J. J. G. C. Lo, “[Evaluation of refractive, tomographic and biomechanical changes after customized accelerated corneal collagen cross-linking in keratoconus patients: a retrospective observational study](#),” *BMC Ophthalmology* 25(1), 378 (2025).
85. A. J. Kanellopoulos, G. Asimellis, “[Hyperopic correction: clinical validation with epithelium-on and epithelium-off protocols, using variable fluence and topographically customized collagen corneal crosslinking](#),” *Clinical Ophthalmology* 8, 2425–2433 (2014).
86. C. Mazzotta, A. Pandolfi, and M. Ferrise, “[Progressive high-fluence epithelium-on accelerated corneal crosslinking: a novel corneal photodynamic therapy for early progressive keratoconus](#),” *Frontiers in Medicine* 10, 1198246 (2023).
87. I. Fischinger, S. A. Reifeltshammer, T. G. Seiler, M. H. Nambiar, M. A. Komninou, P. Büchler, J. Wendelstein, A. Langenbucher, and M. Bolz, “[Analysis of biomechanical response after corneal crosslinking with different fluence levels in porcine corneas](#),” *Current Eye Research* 48(8), 719–723 (2023).
88. N. L. Hafezi, M. E. Aydemir, N.-J. Lu, E. A. Torres-Netto, M. Hillen, and C. Koppen, “[Effect of accelerated high-fluence riboflavin and rose bengal-mediated corneal cross-linking on resistance to enzymatic digestion](#),” *BMC Ophthalmology* 24(1), 37 (2024).
89. P. Kamaev, M. D. Friedman, E. Sherr, and D. Muller, “[Photochemical kinetics of corneal cross-linking with riboflavin](#),” *Investigative Ophthalmology & Visual Science* 53(4), 2360–2367 (2012).
90. O. Richoz, A. Hammer, D. Tabibian, Z. Gatzoufas, and F. Hafezi, “[The biomechanical effect of corneal collagen cross-linking \(CXL\) with riboflavin and UV-A is oxygen dependent](#),” *Translational Vision Science & Technology* 2(7), 6 (2013).
91. V. F. Diakonis, N. Y. Likht, N. Yesilirmak, D. Delgado, A. E. Karatapanis, Y. Yesilirmak, C. Fraker, S. H. Yoo, and N. M. Ziebarth, “[Corneal elasticity after oxygen enriched high intensity corneal cross linking assessed using atomic force microscopy](#),” *Experimental Eye Research* 153, 51–55 (2016).
92. J. Hill, C. Liu, P. Deardorff, B. Tavakol, W. Eddington, V. Thompson, D. Gore, M. Raizman, and D. C. Adler, “[Optimization of oxygen dynamics, UV-A delivery, and drug formulation for accelerated epi-on corneal crosslinking](#),” *Current Eye Research* 45(4), 450–458 (2020).
93. T. G. Seiler, M. A. Komninou, M. H. Nambiar, K. Schuerch, B. E. Frueh, and P. Büchler, “[Oxygen kinetics during corneal cross-linking with and without supplementary oxygen](#),” *American Journal of Ophthalmology* 223, 368–376 (2021).
94. G. Tollin, D. R. Kearns, and M. Calvin, “[Electrical properties of organic solids. I. Kinetics and mechanism of conductivity of metal-free phthalocyanine](#),” *The Journal of Chemical Physics* 32(4), 1013–1019 (1960).
95. P. Vohl, “[Surface photovoltage effects in titanium dioxide films](#),” *Photographic Science and Engineering* 13(3), 120–126 (1969).
96. T. A. Egerton, C. J. King, “[The influence of light intensity on photoactivity in TiO₂ pigmented systems](#),” *Journal of the Oil and Colour Chemists Association* 62, 386–391 (1979).
97. R. Arita, N. Morishige, R. Shirakawa, Y. Sato, and S. Amano, “[Effects of eyelid warming devices on tear film parameters in normal subjects and patients with meibomian gland dysfunction](#),” *The Ocular Surface* 13(4), 321–330 (2015).
98. C. A. Miles, A. J. Bailey, “[Thermal denaturation of collagen revisited](#),” *The Proceedings of the Indian Academy of Sciences* 111, 71–80 (1999).
99. J. C. Bischof, X. He, “[Thermal stability of proteins](#),” *Annals of the New York Academy of Sciences* 1066(1), 12–33 (2006).
100. X. He, “[Thermostability of biological systems: fundamentals, challenges, and quantification](#),” *The Open Biomedical Engineering Journal* 5(1), 47–73 (2011).
101. F. Pucci, M. Rومان, “[Physical and molecular bases of protein thermal stability and cold adaptation](#),” *Current Opinion in Structural Biology* 42, 117–128 (2017).
102. M. J. Buehler, “[Nature designs tough collagen: Explaining the nanostructure of collagen fibrils](#),” *Proceedings of the National Academy of Sciences* 103(33), 12285–12290 (2006).

- 103.M. J. Buehler, "[Nanomechanics of collagen fibrils under varying cross-link densities: Atomistic and continuum studies](#)," Journal of the Mechanical Behavior of Biomedical Materials 1(1), 59–67 (2008).
- 104.S. G. M. Uzel, M. J. Buehler, "[Molecular structure, mechanical behavior and failure mechanism of the C-terminal cross-link domain in type I collagen](#)," Journal of the Mechanical Behavior of Biomedical Materials 4(2), 153–161 (2011).
- 105.B. Depalle, Z. Qin, S. J. Shefelbine, and M. J. Buehler, "[Influence of cross-link structure, density and mechanical properties in the mesoscale deformation mechanisms of collagen fibrils](#)," Journal of the Mechanical Behavior of Biomedical Materials 52, 1–13 (2015).

Polarized Time-Resolved Fluorescence Reveals Enzyme-Specific Anisotropy Dynamics of Bound NADPH

Ioanna A. Gorbunova^{*1}, Denis A. Volkov², Maxim E. Sasin², Oleg S. Vasyutinskii²

¹Institute for Regenerative Medicine, I.M. Sechenov First Moscow State Medical University, 8 Trubetskaya str., Moscow 119991, Russian Federation

²Ioffe Institute, Russian Academy of Sciences, 26 Polytekhnicheskaya, Saint Petersburg 194021, Russian Federation

*e-mail: i.gorbunova@mail.ioffe.ru

Abstract. We used time-resolved polarized fluorescence spectroscopy for the study of nicotinamide adenine dinucleotide phosphate (NADPH) bound to isocitrate dehydrogenase (IDH) and alcohol dehydrogenase (ADH). It was shown that the fluorescence lifetimes of bound NADPH are strongly enzyme-dependent and are practically identical to those of bound NADH, indicating that fluorescence lifetime imaging microscopy (FLIM) analysis cannot universally distinguish between these two coenzymes. However, in the polarization-sensitive experiments we observed sub-nanosecond fluorescence anisotropy decay components of NADPH bound to IDH and ADH that we found to be significantly enzyme-specific. This result suggests that combined lifetime and anisotropy analysis provides increased molecular specificity for NAD(P)H.

Keywords: fluorescence anisotropy; NAD(P)H; TCSPC; fluorescence lifetime; laser spectroscopy.

Paper #9501 received 30 Jan 2026; revised manuscript received 28 May 2026; accepted for publication 4 Jun 2026; published online 1 Sep 2026. [doi: 10.18287/JBPE26.12.030302](https://doi.org/10.18287/JBPE26.12.030302).

1 Introduction

Nicotinamide adenine dinucleotide (NAD) and its phosphate form (NADP) play a crucial role in the regulation of cellular metabolic pathways. Their reduced forms, NADH and NADPH, are endogenous fluorophores absorbing in the UV spectral range (330–380 nm) and emitting fluorescence with a maximum near 460 nm. These properties make NAD(P)H an attractive target for non-invasive, label-free optical monitoring of metabolic activity in living cells.

Fluorescence lifetime imaging microscopy (FLIM) of NADH and NADPH is widely used for metabolic imaging cells and tissues. Within FLIM, fluorescence decay curves are recorded pixel by pixel using fast single-photon detectors and a time-correlated single-photon counting system, enabling the determination of distributions of fluorescence lifetimes τ_i and their relative contributions a_i within cells. Fluorescence lifetimes of NAD(P)H have been shown to be highly sensitive to the local molecular environment, binding state, and conformational dynamics, which underlies the application of FLIM for monitoring cellular metabolism [1–12].

As is known [10, 13], free NADH and NADPH have identical spectral and time-resolved photophysical properties, and their combined signal is commonly referred to as NAD(P)H fluorescence. In living cells, NAD(P)H fluorescence decay is typically described by a bi-exponential model with a short lifetime of $\tau_1 \sim 400$ ps, attributed to free NADH, and a longer lifetime of $\tau_2 \sim 1.1$ –6.0 ns, associated with enzyme-bound NAD(P)H. Changes in these parameters are widely used to access metabolic shifts, redox balance, and energy production [10, 14–16]. However, NADH and NADPH participate in different metabolic pathways: catabolic and biosynthetic/antioxidant processes, respectively. Since both cofactors exhibit nearly identical photophysical properties in their free form [10, 13], separating the contributions of NADH and NADPH, and therefore distinguishing NADH- and NADPH-dependent metabolic processes by FLIM, remains challenging [2, 4, 6].

Recent studies focused on the fluorescence decay of enzyme-bound NADH and NADPH, suggested that differences between them may be detected. Several studies reported that enzyme-bound NADPH may exhibit longer fluorescence lifetimes than bound

NADH, particularly under conditions of oxidative stress or enhanced NADPH oxidase activity [2, 4, 6, 15, 17, 18]. These observations suggested that fluorescence lifetime could be used to separate bound NADPH from bound NADH. However, more recent works demonstrated that the fluorescence lifetime of enzyme-bound NADH and NADPH strongly depends on a specific enzyme binding site and can vary over a wide range [10, 16]. Moreover, most results on the fluorescence decay of enzyme-bound NADH and NADPH were obtained in living cells without systematic validation in simple solution models, which limits the universality of lifetime-based approaches for separating bound NADH and NADPH. An additional challenge arises from the heterogeneity of NAD(P)H fluorescence decay. Both free and protein-bound NAD(P)H exhibit multiple decay components due to conformational diversity, local mobility, and excited-state interactions [9, 10, 19]. Consequently, simplified mono- or bi-exponential analysis of fluorescence decay in mixtures of free and protein-bound NAD(P)H may lead to a substantial loss of molecular specificity.

It has been suggested that the separation of NADH and NADPH in FLIM can be achieved by means of fluorescence anisotropy measurements. Fluorescence anisotropy provides complementary information on excited-state dynamics and molecular mobility that is not accessible from polarization-insensitive (isotropic) fluorescence decay measurements. In particular, time-resolved anisotropy analysis enables the separation of free and protein-bound NADH based on rotational diffusion times and allows estimation of local coenzyme mobility within enzyme binding sites [1, 3, 10]. Although fluorescence anisotropy of NAD(P)H has already been studied in solutions and enzyme complexes [9, 10, 20], its systematic application to enhance molecular specificity in FLIM remained limited.

In this work, we applied time-resolved polarized fluorescence spectroscopy to study the photophysical properties of NADPH bound to two different enzymes, isocitrate dehydrogenase (IDH) and alcohol dehydrogenase (ADH). It was shown that global analysis of polarized fluorescence decay enables simultaneous determination of fluorescence lifetimes and anisotropy parameters and significantly increases molecular specificity for endogenous NAD(P)H fluorescence. We demonstrated that fluorescence anisotropy decay times τ_{ani} and fundamental anisotropy r_0 of NADPH bound to IDH and ADH differ significantly. These differences reflect distinct NADPH mobility depending on the enzyme binding site and enable improved differentiation of enzyme-bound states. Anisotropic fluorescence decay parameters of NADH and NADPH bound to ADH were compared and clear differences between them were obtained. These results highlight polarization-sensitive measurements as a powerful extension of conventional lifetime analysis for resolving metabolic heterogeneity and separating NADH and NADPH contributions in biological systems.

2 Methods

2.1 Materials

The β -NADH (disodium salt, purity >99%, DIA-M, Russia), β -NADPH (disodium salt, purity >98%, Macklin, lot C804424), malate dehydrogenase from porcine heart (>600 units/mg protein, biuret method, ammonium sulfate suspension, lot M1567, Sigma-Aldrich), isocitrate dehydrogenase from porcine heart (Type IV, buffered aqueous glycerol solution, 3–20 units/mg protein, lot I2002, Sigma-Aldrich), and alcohol dehydrogenase (lyophilized powder, recombinant from *E. coli*, 5.0–15.0 U/mg, lot 49854, Sigma-Aldrich) were used in this study. The free NADPH was dissolved in buffer solution in the concentration of 60 μ M and then IDH or ADH was added to the solution. The IDH and ADH concentration was at about 130 μ M. All solutions were prepared in phosphate buffer (PBS) at pH = 7.2. The units U/mg and U/mL denote enzyme activity normalized to the mass or volume of the reagent, respectively. Here, “U” refers to enzyme units, which represent the amount of active enzyme required to catalyze a defined reaction under specified conditions.

2.2 Experimental Procedure

The procedure for experimental detection of polarized fluorescence decay signals was the same as described in our previous publication [8]. The molecular fluorescence was two-photon excited at 720 nm with a femtosecond Ti:sapphire laser (Mai Tai HP, Spectra Physics, USA). The pulse duration was of 100 fs and a repetition rate of 80 MHz. The average laser power at the molecular sample position was 500 mW.

Polarized fluorescence was collected in a direction perpendicular to the excitation beam propagation and detected in the range of 440–480 nm. Two orthogonal polarization components of fluorescence, parallel ($I_{\parallel}$) and perpendicular ($I_{\perp}$) to the polarization excitation beam, were separated using a Glan prism. The two polarization components were then detected simultaneously and independently by two ultrafast avalanche photodiodes (APD-050-CTC, MPD, Germany) operating in the photon-counting mode. The detected signals were processed using a time-correlated single-photon counting (TCSPC) system (PicoHarp 300, PicoQuant, Germany) with a time resolution of 4 ps per channel. The instrumental response function $IRF(t)$ was determined experimentally by detecting scattered laser radiation at wavelengths corresponding to the fluorescence detection range of NADPH.

2.3 Experimental Signal Processing

For analysis of the orthogonally polarized fluorescence signals of NADPH-IDH and NADPH-ADH complexes in solution the following model was used [9]:

$$I_{\parallel}(t) = G \sum_k \left[I_{iso}^k(t) (1 + 2r^k(t)) \right] * IRF(t) + y, (1)$$

$$I_{\perp}(t) = \sum_k \left[I_{iso}^k(t) (1 - r^k(t)) \right] * \text{IRF}(t) + y, \quad (2)$$

where $I_{iso}^k(t)$ is the isotropic (polarization-insensitive) fluorescence decay, $r^k(t)$ is the fluorescence anisotropy decay, y is a background, and $\text{IRF}(t)$ is the instrumental response function. The factor $G = I_Y/I_X$, where I_Y and I_X are the fluorescence components polarized along the Y and X directions, respectively, with both components being orthogonal to the excitation light, accounts for differences in the polarization sensitivity of the two detection channels. The index k denotes f and b, which correspond to free and enzyme-bound NADPH, respectively. This model allows the determination of isotropic fluorescence decay $I_{iso}(t)$ and anisotropy decay $r(t)$ parameters for free and bound species.

The isotropic fluorescence decay of free and enzyme-bound NADPH exhibited heterogeneous behavior and was described by a sum of exponents:

$$I_{iso}^{f/b}(t) = I_0^{f/b} \sum_{i=1}^n a_i \exp\left(-\frac{t}{\tau_i}\right), \quad (3)$$

where $I_0^{f/b}$ is the fluorescence intensity contribution of the corresponding species at $t = 0$, τ_i are fluorescence decay times, and a_i are their weighting coefficients.

The fluorescence anisotropy decay of free NADPH was described by a single-exponential function:

$$r^f(t) = r_0 \exp\left(-\frac{t}{\tau_r}\right), \quad (4)$$

where r_0 is the anisotropy and τ_r is the rotational diffusion time.

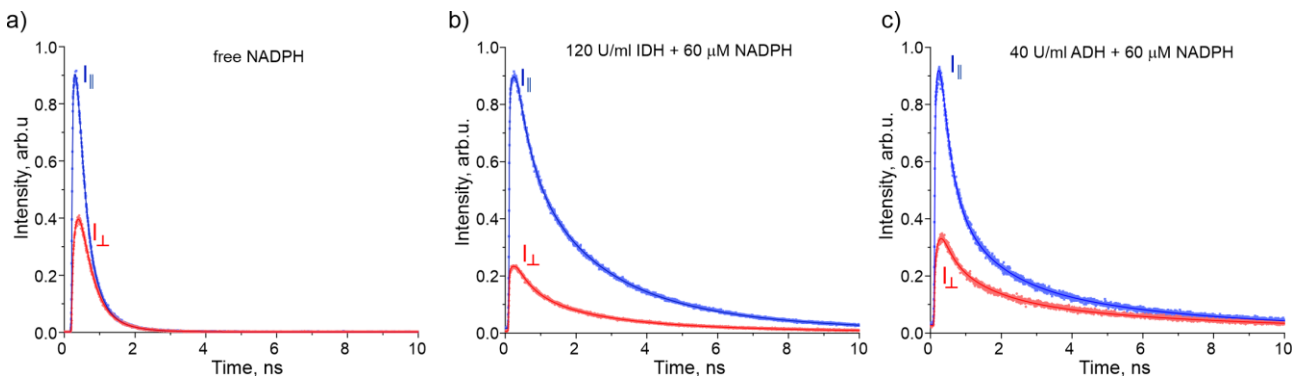


Fig. 1 Orthogonally polarized fluorescence decay signals corresponding to the components parallel ($I_{||}$, blue curve) and perpendicular ($I_{\perp}$, red curve) to the excitation light polarization: (a) free NADPH, (b) solution of 120 U/ml IDH and 60 μM NADPH, and (c) solution of 40 U/ml ADH and 60 μM NADPH. U/mL denotes enzyme activity normalized to the solution volume.

Table 1 Polarized fluorescence decay parameters in NADPH bound with IDH determined by the global fit.

λ_{ex} , nm	τ_{b1} , ns (a_1)	τ_{b2} , ns (a_2)	r_{b1}	r_{b2}	τ_{bv} , ns	τ_r , ns
733	2.27 (0.83)	5.67 (0.17)	0.52	0.02	0.67	50
750	2.28 (0.83)	5.70 (0.17)	0.51	0.02	1.01	50
770	2.26 (0.79)	5.60 (0.21)	0.50	0.06	0.90	50
Uncertainty	± 0.25 (± 0.07)	± 0.30 (± 0.07)	± 0.02	± 0.02	± 0.07	—

In the case of coenzyme-enzyme complexes we took into consideration of different possible depolarization processes in coenzyme excited state and used the expression suggested in our recent publication [19]:

$$r^b(t) = (r_{b1} + r_{b2} e^{-t/\tau_{bv}}) e^{-t/\tau_{br}}, \quad (5)$$

where τ_{bv} characterizes fast depolarization associated with anisotropic vibrational relaxation and τ_{br} characterizes the rotational diffusion time of the entire complex.

3 Results

The fluorescence decay signals of the orthogonal components of polarized fluorescence for IDH and ADH solutions with the addition of 60 μM free NADPH, obtained under two-photon excitation at 733 nm, are shown in Fig. 1(b) and (c), respectively. For comparison, Fig. 1(a) presents the polarized fluorescence decay signals of free NADPH in PBS.

To perform global fitting of the polarized fluorescence decay signals of NADPH in the presence of IDH and ADH, a model accounting for contributions from free NADPH and enzyme-bound NADPH was used. The fluorescence decay parameters of free NADPH were determined independently from the analysis of the signals shown in Fig. 1(a) and were fixed during the fitting procedure. Using this approach, the fluorescence decay parameters of the NADPH-IDH and NADPH-ADH complexes were extracted and are summarized in Tables 1 and 2, respectively.

Table 2 Polarized fluorescence decay parameters in NADPH bound with ADH determined by the global fit.

λ_{ex}, nm	τ_{b1}, ns	r_{b1}	r_{b2}	τ_{bv}, ns	τ_r, ns
733	5.31	0.21	0.22	0.36	10
750	4.90	0.23	0.23	0.24	10
770	5.07	0.23	0.24	0.26	10
Uncertainty	± 0.30	± 0.02	± 0.02	± 0.07	—

It was found that the NADPH-IDH complex exhibits a bi-exponential fluorescence decay with decay times of $\tau_1 \approx 2.3 \text{ ns}$ ($a_1 = 0.83$) and $\tau_2 \approx 5.7 \text{ ns}$ ($a_2 = 0.17$). In contrast, the NADPH-ADH complex is characterized by a single dominant fluorescence decay component with $\tau_1 \approx 5.0 \text{ ns}$. The obtained decay times and their relative contributions were practically independent of the excitation wavelength. Notably, the fluorescence decay times of NADPH bound to IDH and ADH were significantly longer than those of free NADPH in solution. In addition, free NADPH exhibited biexponential fluorescence decay, whereas only a single decay component was observed for the NADPH-ADH complex.

Global fitting of the experimental data in Fig. 1 was further used to analyze the fluorescence anisotropy decay of NADPH-enzyme complexes. The anisotropy decay was described using a model previously developed for coenzyme-enzyme complexes, which allows separation of multiple anisotropic relaxation processes associated with different types of dipole reorientation. These processes include rotational diffusion of the entire complex, local mobility of the coenzyme within the enzyme binding site, reorientation due to nuclear rearrangement during excited-state vibrational relaxation, and resonant energy transfer (FRET) between chromophores located at different binding sites within dimeric enzymes. The fluorescence anisotropy decay parameters obtained from global fitting are summarized in Tables 1 and 2 for the NADPH-IDH and NADPH-ADH complexes, respectively.

As follows from the obtained parameters, the fluorescence anisotropy decay of enzyme-bound NADPH is characterized by two depolarization times. A short depolarization time τ_{bv} of approximately 0.3 ns and 1.0 ns was observed for the NADPH-ADH and NADPH-IDH complexes, respectively, while long rotational diffusion times τ_{br} of approximately 20 ns and 100 ns were determined for the NADPH-ADH and NADPH-IDH complexes, respectively. Significant differences were also observed in the corresponding anisotropy amplitudes r_{b1} and r_{b2} .

For the NADPH-IDH complex, the anisotropy amplitude r_{b1} was 0.52, close to the fundamental anisotropy of free NADPH $r_0 \approx 0.5$, whereas the amplitude r_{b2} associated with the short depolarization component was very small (0.02). In contrast, for the

NADPH-ADH complex, r_{b1} was 0.22, more than two times lower than that of free NADPH and the NADPH-IDH complex. At the same time, the anisotropy amplitude r_{b2} for the NADPH-ADH complex was significantly larger (0.23) than that for the NADPH-IDH complex.

It is noteworthy that the anisotropy amplitude $r_{b2} \approx 0.23$ for the NADPH-ADH complex coincides with the corresponding parameter previously reported for the NADH-ADH complex. However, the short depolarization time $\tau_{bv} \approx 0.3 \text{ ns}$ observed for NADPH-ADH is considerably shorter than that reported earlier for NADH-ADH ($\tau_{bv} \sim 1 \text{ ns}$).

4 Discussion

4.1 Fluorescence Lifetimes of Bound NADH and NADPH

One of the central challenges in NAD(P)H-FLIM is the separation of NADH and NADPH fluorescence signals in cells. Because free NADH and NADPH have identical fluorescence spectra and fluorescence lifetimes, they cannot be distinguished in the free form. For this reason, recent studies [9, 10] have focused on enzyme-bound cofactors and suggested that differences in fluorescence decay times of bound NADH and NADPH may enable their separation. In particular, Blacker et al. reported that the NADPH-NOX complex exhibits a significantly longer fluorescence lifetime than bound NADH and proposed a general rule that enzyme-bound NADPH has a longer fluorescence decay time than NADH, independent of the enzyme. However, studies of NAD(P)H-enzyme complexes in solution demonstrate that fluorescence decay times of NADH and NADPH strongly depend on the specific enzyme binding site. Table 3 summarizes representative fluorescence decay parameters reported for various NADH- and NADPH-dependent dehydrogenases. As can be seen in Table 3, both the absolute decay times and the number of decay components vary significantly between enzymes. In several cases, bound NADH exhibits fluorescence decay times comparable to or even longer than those of bound NADPH, indicating that a simple lifetime-based distinction between NADH and NADPH is not universal. In addition to enzyme specificity, another important factor affecting fluorescence decay analysis is the heterogeneity of free NADH fluorescence.

Table 3 Fluorescence decay times τ_i and weighting coefficients a_i of NAD(P)H-enzyme complexes reported in the literature.

Species	$\lambda_{ex}, \lambda_{fluo}$	τ_1 (ns) (a_1)	τ_2 (ns) (a_2)	Source
NADH-LDH	340 nm, 368–600 nm	1.34 (0.22)	3.20 (0.09)	[10]
	340 nm, 368–600 nm	–	3.34 (0.09)	[21]
	750 nm, 440/40 nm	–	1.60 (1.00)	[14]
NADH-MDH	720 nm, 436/10 nm	–	1.24 (1.00)	[16]
	750 nm, 440/40 nm	–	1.20 (1.00)	[14]
NADH-ADH	720 nm, 436/10 nm	–	4.44 (1.00)	[9]
NADPH-IDH	340 nm, 368–600 nm	1.59 (0.12)	4.44 (0.02)	[10]
	760 nm, 466/30 nm	–	2.20 (1.00)	[16]
NADPH-G6PDH	750 nm, 440/40 nm	–	2.50 (1.00)	[16]

Free NAD(P)H fluorescence decay is heterogeneous and is characterized by two decay components with lifetimes of approximately $\tau_1 \sim 0.25$ ns and $\tau_2 \sim 0.6$ ns, which have been attributed to two distinct cis and trans conformations of the nicotinamide moiety. Approximating the fluorescence decay of free NAD(P)H by a single-exponential function, as was done in several studies listed in Table 3, can lead to quantitative errors in the determination of fluorescence decay times of enzyme-bound NAD(P)H in mixed systems.

In this work, we applied a refined analysis of fluorescence decay in mixtures of free NADPH and NADPH bound to IDH and ADH. In our model, the heterogeneity of free NADPH fluorescence was explicitly taken into account by describing the free coenzyme with a biexponential decay function. This approach allowed us to determine fluorescence decay times of enzyme-bound NADPH with higher accuracy than in previous studies. We found that the NADPH-IDH complex exhibits two nanosecond fluorescence decay components of approximately 2 ns and 5 ns, whereas the NADPH-ADH complex is characterized by a single fluorescence decay time of about 5 ns. The shorter decay component of the NADPH-IDH complex (~ 2 ns) is consistent with the value reported by Leben et al. [16] but differs from that reported by Blacker et al.

Another important result of this study is the determination of the fluorescence decay time of the NADPH-ADH complex, which has not been reported previously. This result enables a direct comparison of fluorescence decay parameters of NADH and NADPH bound to the same enzyme. We show that the fluorescence decay time of NADPH-ADH (~ 5 ns) determined in this work is not significantly longer than that of NADH-ADH (~ 4.5 ns) reported in our previous study [9].

Overall, the strong enzyme specificity of NADPH fluorescence lifetimes and the nearly identical fluorescence decay times of NADH and NADPH bound to ADH demonstrate that lifetime-based separation of NADH and NADPH cannot rely on a universal fluorescence lifetime criterion.

4.2 Fluorescence Anisotropy Decay in Bound NADH and NADPH

Here, we show that fluorescence anisotropy decay of bound NADPH is heterogeneous and characterized by two distinct depolarization times of approximately 1 ns and longer than 20 ns, as summarized in Tables 1 and 2. The longer depolarization time corresponds to rotational diffusion of the entire coenzyme–enzyme complex. Elucidation of the origin of the short depolarization time was the main focus of this study.

Heterogeneous fluorescence anisotropy decay is a general feature of fluorophore-macromolecule complexes and has been discussed extensively for a wide range of systems, including fluorescent proteins, dye-labeled DNA, and membrane-associated probes. Two major theoretical frameworks have been proposed to describe fluorescence anisotropy heterogeneity. The first comprises diffusion-based models that account for multiple types of rotational motion, including fast local motions of the fluorophore, slower dynamics of protein structural elements, and global tumbling of the entire complex [22, 23]. The second framework is based on a more general formalism using spherical tensors and correlation functions, in which multiple relaxation pathways contributing to anisotropy decay are taken into account [9, 19, 24].

Among diffusion-based approaches, the “wobbling-in-a-cone” model has been widely used to interpret anisotropy decay of NAD(P)H-enzyme complexes [10, 25, 26]. Within this framework, fluorescence depolarization of bound NADH and NADPH arises from a combination of global rotational

diffusion of the enzyme and restricted local motion of the coenzyme within a cone defined by the geometry of the binding site. Using this model, Blacker et al. [10] concluded that NADPH bound to IDH exhibits significantly faster local motion than NADH bound to LDH, consistent with differences in steric constraints and electrostatic interactions in the respective binding sites.

In contrast, the model used in the present study to analyze fluorescence anisotropy decay of bound NADPH (Eq. (5)) is based on a more general formalism that explicitly separates different anisotropic relaxation processes occurring in the excited state of NADPH. This model was first proposed in our recent paper [19] and successfully applied to the NADH-ADH complex [9]. Within this framework, the anisotropy parameters r_{b1} and r_{b2} in Eq. (5) characterize the relative orientation of the excitation and emission transition dipole moments in the relaxed state and during anisotropic vibrational relaxation, respectively. The parameter r_{b2} reflects changes in the orientation of the fluorescence transition dipole moment induced by fast nuclear relaxation in the excited state of the coenzyme. Correspondingly, the fast fluorescence depolarization time τ_{bv} in Eq. (5) reflects the rate of anisotropic relaxation associated with nuclear rearrangements in the excited state.

We suggest that the short depolarization time of approximately 1 ns observed for the NADPH-IDH and NADPH-ADH complexes reflects the rate of nuclear rearrangements of NADPH in the excited state and can be used as a measure of local coenzyme mobility within the enzyme binding site. For the NADPH-IDH complex, the anisotropy amplitude r_{b1} is close to the fundamental anisotropy of free NADPH, while the parameter r_{b2} associated with the fast depolarization component is close to zero. This indicates that the orientation of the excitation and emission transition dipole moments changes only weakly during excited-state vibrational relaxation, suggesting limited local mobility of NADPH in the IDH binding site under our experimental conditions. In contrast, for the NADPH-ADH complex, the parameter r_{b2} is significantly larger ($r_{b2} = 0.3$), indicating a pronounced reorientation of the transition dipole moment during excited-state relaxation and enhanced local mobility of NADPH in the ADH binding site.

These findings reflect fundamental differences in binding-site mobility and conformational dynamics of NADPH in IDH and ADH. NMR and X-ray studies have shown that dehydrogenases undergo large-amplitude conformational changes during catalysis and exhibit varying degrees of binding-site mobility [27, 28]. Our results indicate that NADPH bound to ADH exhibits higher local mobility than NADPH bound to IDH and is also more mobile than NADH bound to ADH.

We further compared anisotropy decay parameters of NADPH and NADH bound to ADH and revealed coenzyme-specific dynamics within the ADH binding

site. We found that the fast depolarization time of the NADPH-ADH complex ($\tau_{bv} \sim 0.3$ ns) is considerably shorter than that of the NADH-ADH complex ($\tau_{bv} \sim 1$ ns). At the same time, the anisotropy amplitude r_{b2} associated with the fast component is nearly identical for NADH and NADPH bound to ADH ($r_{b2} \approx 0.23$). This demonstrates that NADPH experiences faster local depolarization dynamics than NADH in the same enzyme environment.

Overall, these results demonstrate that time-resolved fluorescence anisotropy provides a sensitive probe of local coenzyme mobility and binding-site dynamics. When combined with isotropic fluorescence decay analysis, anisotropy measurements enable discrimination between enzyme-bound NAD(P)H states that cannot be resolved by fluorescence lifetime analysis alone, highlighting their potential for improving molecular specificity in FLIM-based metabolic studies.

5 Conclusion

In this work, we performed a comprehensive time-resolved study of fluorescence decay and anisotropy of NADPH bound to isocitrate dehydrogenase (IDH) and alcohol dehydrogenase (ADH). By combining a global fitting procedure of polarized fluorescence decay signals with explicit consideration of the heterogeneous fluorescence decay of free NADPH, we demonstrated enzyme-specific features of polarized fluorescence decay of NADPH. In particular, we found that the fluorescence lifetime of NADPH bound to IDH is approximately two times shorter than that of the NADH-ADH complex, whereas the fluorescence lifetimes of NADH and NADPH bound to ADH are nearly identical. These results confirm that fluorescence lifetime alone cannot be reliably used to separate NADH and NADPH in living cells by FLIM.

The major result of this study is the observation of heterogeneous fluorescence anisotropy decay in NADPH-enzyme complexes. We show that both NADPH-IDH and NADPH-ADH are characterized by a sub-nanosecond fluorescence depolarization component in addition to a long rotational diffusion time of the entire complex exceeding 20 ns. We suggest that the fast depolarization component reflects rapid nuclear rearrangements of NADPH in the excited state and/or local coenzyme mobility within the enzyme binding site. Importantly, the fast depolarization component differs significantly between NADPH-IDH and NADPH-ADH, indicating enzyme-specific local mobility of NADPH. Moreover, comparison of anisotropy decay of NADPH-ADH and NADH-ADH reveals coenzyme-specific anisotropic excited-state dynamics that are not accessible from isotropic fluorescence lifetime analysis.

Overall, our results demonstrate that combined analysis of isotropic fluorescence decay and fluorescence anisotropy provides additional molecular specificity for endogenous NAD(P)H. This approach opens new opportunities for distinguishing enzyme-bound states and lays the groundwork for extending

polarized FLIM toward improved separation of NADH- and NADPH-related metabolic pathways in complex biological systems.

Disclosures

All authors declare that there is no conflict of interest in this paper.

References

1. H. D. Vishwasrao, A. A. Heikal, K. A. Kasischke, and W. W. Webb, “Conformational Dependence of Intracellular NADH on Metabolic State Revealed by Associated Fluorescence Anisotropy,” *Journal of Biological Chemistry* 280 (26), 25119–25126 (2005).
2. R. Niesner, P. Narang, H. Spiecker, V. Andresen, K.-H. Gericke, and M. Gunzer, “Selective Detection of NADPH Oxidase in Polymorphonuclear Cells by Means of NAD(P)H-Based Fluorescence Lifetime Imaging,” *Journal of Biophysics* 2008, 602639 (2008).
3. Q. Yu, A. A. Heikal, “Two-photon autofluorescence dynamics imaging reveals sensitivity of intracellular NADH concentration and conformation to cell physiology at the single-cell level,” *Journal of Photochemistry and Photobiology B: Biology* 95(1), 46–57 (2009).
4. A. A. Mossakowski, J. Pohlen, D. Bremer, R. Lindquist, J. M. Millward, M. Bock, K. Pollok, R. Mothes, L. Viohl, M. Radbruch, J. Gerhard, J. Bellmann-Strobl, J. Behrens, C. Infante-Duarte, A. Mähler, M. Boschmann, J. L. Rinnenthal, M. Fuchtemeier, J. Herz, F. C. Pache, M. Bardua, J. Priller, A. E. Hauser, F. Paul, R. Niesner, and H. Radbruch, “Tracking CNS and systemic sources of oxidative stress during the course of chronic neuroinflammation,” *Acta Neuropathologica* 130(6), 799–814 (2015).
5. O. S. Vasyutinskii, A. G. Smolin, C. Oswald, and K. H. Gericke, “Polarized fluorescence in NADH under two-photon excitation with femtosecond laser pulses,” *Optics and Spectroscopy* 122(4), 602–606 (2017).
6. R. Leben, L. Ostendorf, S. Van Koppen, A. Rakhymzhan, A. Hauser, H. Radbruch, and R. Niesner, “Phasor-Based Endogenous NAD(P)H Fluorescence Lifetime Imaging Unravels Specific Enzymatic Activity of Neutrophil Granulocytes Preceding NETosis,” *International Journal of Molecular Sciences* 19(4), 1018 (2018).
7. M. E. Sasin, I. A. Gorbunova, N. O. Bezverkhni, Y. M. Beltukov, O. S. Vasyutinskii, and J. Rubayo-Soneira, “Polarized Fluorescence in NADH Two-Photon Excited by Femtosecond Laser Pulses in the Wavelength Range of 720–780 nm,” *Technical Physics Letters* 45(7), 672–674 (2019).
8. I. A. Gorbunova, M. E. Sasin, J. Rubayo-Soneira, A. G. Smolin, and O. S. Vasyutinskii, “Two-Photon Excited Fluorescence Dynamics in NADH in Water–Methanol Solutions: The Role of Conformation States,” *The Journal of Physical Chemistry B* 124(47), 10682–10697 (2020).
9. I. A. Gorbunova, M. E. Sasin, D. P. Golyshev, A. A. Semenov, A. G. Smolin, Y. M. Beltukov, and O. S. Vasyutinskii, “Two-Photon Excited Fluorescence Dynamics in Enzyme-Bound NADH: the Heterogeneity of Fluorescence Decay Times and Anisotropic Relaxation,” *The Journal of Physical Chemistry B* 125(34), 9692–9707 (2021).
10. T. S. Blacker, M. R. Duchon, and A. J. Bain, “NAD(P)H binding configurations revealed by time-resolved fluorescence and two-photon absorption,” *Biophysical Journal* 122(7), 1240–1253 (2023).
11. V. I. Shcheslavskiy, M. V. Shirmanova, K. S. Yashin, A. C. Rück, M. C. Skala, and W. Becker, “Fluorescence Lifetime Imaging Techniques—A Review on Principles, Applications and Clinical Relevance,” *Journal of Biophotonics* 18(12), e202400450 (2025).
12. D. L. Pham, A. A. Gillette, J. Riendeau, K. Wiech, E. C. Guzman, R. Datta, and M. C. Skala, “Perspectives on label-free microscopy of heterogeneous and dynamic biological systems,” *Journal of Biomedical Optics* 29(S2), (2024).
13. E. V. Potapova, E. A. Zhrebtssov, V. V. Shupletsov, V. V. Dremin, K. Y. Kandurova, A. V. Mamoshin, A. Y. Abramov, and A. V. Dunaev, “Detection of NADH and NADPH levels in vivo identifies shift of glucose metabolism in cancer to energy production,” *The FEBS Journal* 291(12), 2674–2682 (2024).
14. J. T. Sharick, P. F. Favreau, A. A. Gillette, S. M. Sdao, M. J. Merrins, and M. C. Skala, “Protein-bound NAD(P)H Lifetime is Sensitive to Multiple Fates of Glucose Carbon,” *Scientific Reports* 8(1), 5456 (2018).
15. T. S. Blacker, Z. F. Mann, J. E. Gale, M. Ziegler, A. J. Bain, G. Szabadkai, and M. R. Duchon, “Separating NADH and NADPH fluorescence in live cells and tissues using FLIM,” *Nature Communications* 5(1), 3936 (2014).
16. R. Leben, M. Köhler, H. Radbruch, A. E. Hauser, and R. A. Niesner, “Systematic Enzyme Mapping of Cellular Metabolism by Phasor-Analyzed Label-Free NAD(P)H Fluorescence Lifetime Imaging,” *International Journal of Molecular Sciences* 20(22), 5565 (2019).
17. P. Majumder, T. S. Blacker, L. S. Nolan, M. R. Duchon, and J. E. Gale, “Multiphoton NAD(P)H FLIM reveals metabolic changes in individual cell types of the intact cochlea upon sensorineural hearing loss,” *Scientific Reports* 9(1), 18907 (2019).
18. A. Song, N. Zhao, D. C. Hilpert, C. Perry, J. A. Baur, D. C. Wallace, and P. M. Schaefer, “Visualizing subcellular changes in the NAD(H) pool size versus redox state using fluorescence lifetime imaging microscopy of NADH,” *Communications Biology* 7(1), 428 (2024).

19. I. A. Gorbunova, M. E. Sasin, Y. M. Beltukov, A. A. Semenov, and O. S. Vasyutinskii, “[Anisotropic relaxation in NADH excited states studied by polarization-modulation pump–probe transient spectroscopy](#),” *Physical Chemistry Chemical Physics* 22(32), 18155–18168 (2020).
20. S. Cao, H. Li, Y. Liu, M. Wang, M. Zhang, S. Zhang, J. Chen, J. Xu, J. R. Knutson, and L. Brand, “[Dehydrogenase Binding Sites Abolish the “Dark” Fraction of NADH: Implication for Metabolic Sensing via FLIM](#),” *The Journal of Physical Chemistry B* 124(31), 6721–6727 (2020).
21. S. Ranjit, L. Malacrida, M. Stakic, and E. Gratton, “[Determination of the metabolic index using the fluorescence lifetime of free and bound nicotinamide adenine dinucleotide using the phasor approach](#),” *Journal of Biophotonics* 12(11), e201900156 (2019).
22. C. W. Ko, Z. Wei, R. J. Marsh, D. A. Armoogum, N. Nicolaou, A. J. Bain, A. Zhou, and L. Ying, “[Probing nanosecond motions of plasminogen activator inhibitor-1 by time-resolved fluorescence anisotropy](#),” *Molecular BioSystems* 5(9), 1025–1031 (2009).
23. G. F. Schröder, U. Alexiev, and H. Grubmüller, “[Simulation of Fluorescence Anisotropy Experiments: Probing Protein Dynamics](#),” *Biophysical Journal* 89(6), 3757–3770 (2005).
24. A. Szabo, “[Theory of fluorescence depolarization in macromolecules and membranes](#),” *The Journal of Chemical Physics* 81(1), 150–167 (1984).
25. K. Kinoshita, A. Ikegami, and S. Kawato, “[On the wobbling-in-cone analysis of fluorescence anisotropy decay](#),” *Biophysical Journal* 37(2), 461–464 (1982).
26. G. Lipari, A. Szabo, “[Effect of librational motion on fluorescence depolarization and nuclear magnetic resonance relaxation in macromolecules and membranes](#),” *Biophysical Journal* 30(3), 489–506 (1980).
27. S. B. Raj, S. Ramaswamy, and B. V. Plapp, “[Yeast Alcohol Dehydrogenase Structure and Catalysis](#),” *Biochemistry* 53(36), 5791–5803 (2014).
28. B. V. Plapp, B. R. Savarimuthu, D. J. Ferraro, J. K. Rubach, E. N. Brown, and S. Ramaswamy, “[Horse Liver Alcohol Dehydrogenase: Zinc Coordination and Catalysis](#),” *Biochemistry* 56(28), 3632–3646 (2017).

Doppler Ultrasound Indices of the Major Arteries as Markers of Vascular Tone during Thermal Stress

Anatoly V. Skripal^{1*}, Konstantin V. Mashkov¹, Andrey A. Sagaidachnyi¹, and Elena S. Olenko²

¹Department of Medical Physics, Saratov State University, 83 Astrakhanskaya str., Saratov 410012, Russian Federation

²Department of Fundamentals of Medicine and Medical Technologies, Saratov State University, 83 Astrakhanskaya str., Saratov 410012, Russian Federation

*e-mail: skripalav@info.sgu.ru

Abstract. This study focuses on the effects of moderate thermal heating on microcirculation and peripheral hemodynamics, as well as on the interrelationship between changes detected by Laser Doppler Flowmetry (LDF) and Ultrasound Doppler (USD). Microcirculation and peripheral hemodynamics were evaluated in the context of cardiovascular diagnostics using LDF and USDG. LDF non-invasively assessed tissue perfusion, reflecting the state of arterioles, venules, precapillary sphincters, and capillaries, thereby evaluating blood filling and tone in the microcirculatory bed. USDG was employed to assess blood flow in major arteries and calculate peripheral resistance indices, including resistance and pulsatility indices. General thermal heating from 20°C to 30°C was applied to observe its impact on both microvascular tone and major artery resistance indices. Moderate thermal heating increased perfusion and amplified the amplitude of neurogenic and myogenic oscillation rhythms. Peripheral resistance, as determined by USDG, decreased with rising temperature. The retrograde blood flow component, indicative of the tone of distal muscular-type vessels, disappeared upon warming, suggesting reduced resistance in both major arteries and microvessels. Additionally, the amplitude of antegrade diastolic flow, associated with the properties of elastic-type arteries, increased. Changes in microcirculation and peripheral hemodynamics are interconnected and follow a common physiological pattern: vasodilation and reduced tone in the microcirculatory bed. The decrease in USDG spectral indices in major arteries during heating correlates directly with the reduction in microvascular tone. While USDG can detect vascular resistance changes within a single cardiac cycle, LDF requires significantly more time to register such changes in microvascular tone.

Keywords: vascular elasticity; vascular resistance; microcirculation; major artery; ultrasound Doppler; laser doppler flowmetry; thermal test.

Paper #9587 received 15 Apr 2026; revised manuscript received 7 Jul 2026; accepted for publication 7 Jul 2026; published online 7 Sep 2026. [doi: 10.18287/JBPE26.12.030303](https://doi.org/10.18287/JBPE26.12.030303).

1 Introduction

Almost all major pathologies of the cardiovascular system are accompanied by significant impairments in microcirculation or blood flow in the major arteries. For the real-time assessment of microcirculation [1–3], Laser Doppler Flowmetry (LDF) is a promising method. To

determine the hemodynamics of medium and small-caliber arteries and veins [4–6], Ultrasound Doppler (USDG) has proven to be highly effective. It is generally accepted that these methods provide information on the hemodynamic mechanisms of various functional systems in the body at the micro- and macrovascular levels. In the present work, we aim to demonstrate the relationship

between these methods for assessing hemodynamic changes during a general thermal functional test.

Studies using LDF [7–10] have established that both local and general thermal exposure induces a significant increase in cutaneous blood flow. The perfusion and the amplitude of blood flow oscillations in various frequency ranges are used to assess blood filling and tone in the microcirculatory bed. These oscillations are governed by active mechanisms of endothelial, neurogenic, and myogenic regulation [11–13].

Specifically, a study by Petrofsky et al. [7] demonstrated that exposure to air flows with different temperature parameters (38°C, 40°C, 42°C) for 20 min resulted in a significant increase in both skin blood flow and skin temperature. A similar increase in cutaneous blood flow was also observed in the works of Fedorovich and Skripal [14, 15].

An impaired microvascular response to thermal hyperemia is closely associated with increased cardiovascular mortality, underscoring the prognostic value of this non-invasive test for assessing skin microcirculatory function [16].

The investigation of regulatory mechanisms in microcirculation under various pathological conditions is made possible through the analysis of blood flow oscillation rhythms in tissues [17]. The microcirculatory response to local heating in both healthy individuals and patients with diabetes mellitus has been described in studies by Filina et al. [18] and Mizeva et al. [9]. It was found that a local increase in skin temperature from 25°C to 42°C leads to a significant rise in the microcirculation index, accompanied by a marked tendency for an increase in oscillation amplitude within the neurogenic and myogenic ranges. In patients with diabetes, a reduction in neurogenic vasodilation and an impairment of low-frequency vasomotions were noted, reflecting dysfunction in the autonomic regulation of vascular tone.

According to a study by Mallette et al. [12], passive heat stress significantly increases laser-Doppler flux relative to baseline levels. This is accompanied by an amplification of blood flow oscillations associated with nitric oxide-independent endothelial activity, as well as neurogenic, myogenic, respiratory, and cardiac rhythms. In contrast, oscillations reflecting nitric oxide-dependent endothelial activity showed no significant changes under the influence of passive heat stress.

A number of studies [19, 20] have shown that the indicator of oxidative metabolism of the NADH coenzyme can also be a marker of changes in the tone of the microcirculatory bed when heated. However, the dependence of this parameter, measured by modern LDF devices with integrated fluorescence analysis of blood flow [21], depends on the individual characteristics, the measuring zone, and even the phase of the menstrual cycle in women [20].

In contrast to LDF, the USDG method provides recording of the direction and velocity of blood flow in large arteries and veins, enabling the assessment of macrocirculatory blood flow and the calculation of spectral indices of peripheral resistance. For example, the

Resistance Index (RI) serves as an integral indicator reflecting the total peripheral vascular resistance in the examined vascular bed. Within the classical approach, its calculation is based on the analysis of the direction and magnitude of blood flow during the diastolic phase of the cardiac cycle.

A study by Korkmaz et al. [22] demonstrated that the change in RI following reactive hyperemia of the brachial artery reflects the functional state of the endothelium and can be used in clinical practice as a diagnostic marker of vascular dysfunction. Furthermore, the RI level closely correlates with risk factors for coronary artery disease and serves as an indirect indicator of atherosclerotic changes. This is supported by data obtained from a group of renal transplant recipients, where an association was found between the renal RI and cardiovascular risk factors, as well as the thickening of the carotid intima-media complex – a marker of subclinical atherosclerosis [23].

The Pulsatility Index (PI) is an integral parameter reflecting a combination of peripheral resistance and stiffness (elasticity) of distal arteries, as well as the magnitude of cardiac output. It allows for the quantitative assessment of blood flow velocity over the cardiac cycle period [24].

The effect of local and whole-body heating on vascular function, including flow-mediated dilation and post-occlusive reactive hyperemia, has been comprehensively reviewed [25]. The authors demonstrated that systemic thermal exposure significantly increases forearm blood volume flow within the first minutes, which is accompanied by an increase in shear rate and enhanced antegrade flow, while a retrograde component is also present.

A review presented by Brothers et al. [26] analyzed data from a number of experimental studies investigating the impact of heat stress on dynamic cerebral autoregulation under conditions of significant blood pressure fluctuations. Based on the summarized results, the authors concluded that heat stress is accompanied by a reduction in systemic vascular resistance.

Acute thermal exposure exerts a pronounced effect on vascular regulation by enhancing endothelium-dependent vasodilation in conduit arteries. A study by Coombs et al. [27] analyzed the contribution of both skin temperature response and core body temperature to the flow-mediated dilation of the brachial artery during different modes of passive heating. During systemic thermal exposure without limb occlusion, a marked increase in antegrade blood flow velocity in the brachial artery was observed, accompanied by a progressive reduction of the retrograde component up to its complete disappearance.

Interest in studying limb blood flow obtained via ultrasound Doppler waveform is driven by the ability to visualize both antegrade and retrograde flow, which allows for the assessment of the functional state of the vascular system and the determination of its resistance. Results from a study by Thomas et al. [28] showed that leg immersion in hot water (42°C) for 30 min led to an increase in antegrade blood flow velocity and a decrease

in retrograde velocity. These changes were recorded in both the upper and lower limbs. The observed changes indicate the involvement of the resistive vascular bed in thermoregulatory mechanisms and the influence of the myogenic tone of small arteries and arterioles on peripheral circulation [29].

Literature analysis reveals studies that have investigated the vascular response to thermal exposure using either ultrasound Doppler or LDF. Simultaneous assessment of Doppler ultrasound blood flow indices and microcirculatory parameters could provide insights into the relationship between regulatory processes at the macro- and microcirculatory levels of the vascular system organization.

The aim of this work was to establish the possibility of using ultrasound Doppler ultrasound indices measured in the radial artery as a marker of microvascular tone during thermal stress.

2 Materials and Methods

2.1 Participants

Fifteen healthy volunteers (9 males and 6 females) with a mean age of 22 ± 2 years participated in this study. Prior to the experiment, each participant was thoroughly informed about the study's objectives, procedures, and methods, after which they provided written informed consent in accordance with the ethical standards for biomedical research. The study was approved by the Ethics Committee of the V.I. Razumovsky Saratov State Medical University (Protocol No. 9, April 4, 2023) and was conducted in a clinical setting under the supervision of medical personnel.

2.2 Microcirculation Assessment

Microcirculatory blood flow was assessed using the "LAZMA PF" device (LAZMA, Russia), which operates at a wavelength of 850 nm with a radiation power of 0.7 mW. The penetration depth of the laser beam into skin tissue was approximately 1.0–1.2 mm. The "LAZMA PF" probe was placed on the distal phalanx of the middle finger pad of the right hand. LDF signals were analyzed using Morlet wavelet transform, which decomposed the signals into frequency components of blood flow oscillations. Among the five existing frequency ranges, parameter changes were analyzed in only two: 0.02–0.06 Hz – neurogenic activity (blood flow regulation via sympathetic vascular innervation); 0.06–0.2 Hz – myogenic activity (contractions of vascular smooth muscle cells). This limitation is due to the duration of the LDF signals.

2.3 Ultrasound Doppler

Doppler waveforms were recorded using an "Edan U50" ultrasound system (Edan Instruments, China) with a 7.2 MHz transducer. Hemodynamic parameters were assessed using multiple modes: B-mode for visualization of vascular anatomy; color Doppler (CD) mapping for localization of arterial blood flow; and pulsed-wave

Doppler (PW Doppler) for obtaining quantitative blood flow velocity parameters.

2.4 Study Protocol

The investigation was conducted in two phases. During the first phase, participants were seated in a laboratory room with a controlled temperature of $20 \pm 1^\circ\text{C}$, monitored using a digital thermometer (Xiaomi Mijia Bluetooth Thermometer 2, China). Following a 15-min rest period for acclimatization, baseline measurements of microcirculation parameters were performed on the pad of the left index finger using LDF, along with recording Doppler waveforms of the radial artery using Doppler ultrasound.

In the second phase, participants moved to an adjacent room with an ambient temperature of $30 \pm 1^\circ\text{C}$. After 15 min of whole-body thermal exposure, subsequent measurements of microcirculation were conducted on the pad of the distal phalanx of the left index finger (LDF) and Doppler waveforms of the radial artery (ultrasonography) were recorded while participants remained at rest.

2.5 Laser Doppler Flowmetry Parameters

The perfusion and amplitudes of blood flow oscillations in various frequency ranges were determined from LDF measurements. The perfusion (P) is defined as:

$$P = K \cdot N \cdot V_{avg}, \quad (1)$$

where K is the proportionality coefficient, N is the number of scatterers (erythrocytes) in the measured volume, and V_{avg} is their average velocity [30].

Since the LDF signal consists of slowly and rapidly varying components, the perfusion can be described as:

$$P(t) = Pc + \delta P(t), \quad (2)$$

where Pc represents the constant flow component and $\delta P(t)$ represents the variable flow component.

The constant component Pc reflects the average blood flow level in the microcirculatory bed over the selected observation period or LDF recording analysis interval. The variable component $\delta P(t)$ is determined by factors affecting blood flow stability – specifically, changes in the velocity V_{avg} and concentration N of erythrocytes in the measured volume.

The recorded amplitudes of neurogenic (An) and myogenic (Am) blood flow oscillations are directly related to microvascular lumen size and, consequently, to muscular tone. Neurogenic tone (NT) and myogenic tone (MT) of arterioles and resistive microvessels are related by the expressions [31]:

$$NT = \frac{\sigma \cdot P_{avg}}{An \cdot Pc}, \quad (3)$$

$$MT = \frac{\sigma \cdot P_{avg}}{Am \cdot Pc}, \quad (4)$$

where σ represents the standard deviation of the perfusion, and P_{avg} is the mean pressure in the peripheral vascular system.

A decrease in oscillation amplitude is associated with increased tone and stiffness of the vascular wall, while an increase in oscillation amplitude results from reduced vascular tone [31–33].

2.6 Ultrasound Doppler Parameters

The resistance index (RI) and pulsatility index (PI) were calculated from the ultrasound Doppler data, with averaging performed over three cardiac cycles. Typically, two different equations are used for calculating RI, accounting for peripheral vascular resistance [34]. The first equation is applied in cases of high vascular resistance with presence of retrograde blood flow:

$$RI_{V_d} = \frac{V_{ps} + V_d}{V_{ps}}, \quad (5)$$

where V_{ps} is the peak systolic blood flow velocity and V_d is the peak retrograde blood flow velocity.

The second equation is used for cases of low vascular resistance:

$$RI_{V_{ed}} = \frac{V_{ps} - V_{ed}}{V_{ps}}, \quad (6)$$

where V_{ed} represents the antegrade peak velocity.

It should be noted that in cases of low vascular resistance, the end-diastolic velocity V_{ed} is commonly used for calculating both RI and PI indices. However, in healthy young participants at room temperature, the end-diastolic velocity is typically zero due to high vascular resistance. Therefore, we employed modified formulas for calculating RI and PI indices, utilizing the diastolic velocity V_{ed} corresponding to the maximum antegrade flow. A similar approach was implemented in the monograph by Lelyuk, V.G. and Lelyuk, S.E. [34].

The expression for the PI for cases with high vascular resistance was determined from the following relationship:

$$RI_{V_d} = \frac{V_{ps} + V_d}{TAMX}, \quad (7)$$

where TAMX – the time-averaged maximum velocity as the mean value of all measured maximum blood flow velocities obtained at equal time points throughout one cardiac cycle:

$$TAMX = \frac{\sum_{i=1}^n V_i}{n}, \quad (8)$$

where V_i is the measured value of maximum blood flow velocity in the i -th pixel (cm/s), and n is the number of measured points (pixels on the Doppler waveform) per cardiac cycle.

For cases with low vascular resistance, the PI was determined from the following relationship:

$$PI_{V_{ed}} = \frac{V_{ps} - V_{ed}}{TAMX}, \quad (9)$$

where V_{ed} was also defined as the antegrade peak diastolic velocity.

2.7 Statistical Analysis

Data analysis was performed using mathematical statistics methods. The STATISTICA 10 software package (StatSoft, Inc.) was used for data processing. Given the small sample size and non-normal distribution of data, the Wilcoxon T -test – one of the most powerful non-parametric analysis methods – was applied for comparing quantitative variables in two dependent samples. Parameter differences were considered statistically significant at $p \leq 0.05$, indicating a high probability that the observed changes in blood flow parameters were not due to random factors.

3 Results

3.1 LDF Measurements of Microcirculatory Blood Flow

Figure 1 presents a graph illustrating changes in the perfusion, measured in perfusion units (p.u.), in a healthy 25-year-old subject under two conditions: in a room with air temperature of $20 \pm 1^\circ\text{C}$ and $30 \pm 1^\circ\text{C}$, respectively.

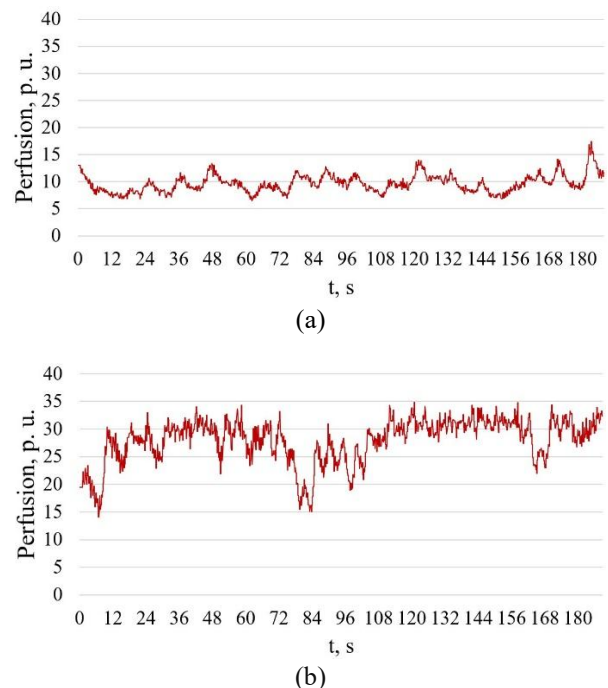


Fig. 1 Perfusion of a 25-year-old study participant: (a) in a room with air temperature of $20 \pm 1^\circ\text{C}$, (b) in a room with air temperature of $30 \pm 1^\circ\text{C}$.

As shown in Fig. 1, the average level of the perfusion in the volunteer before exposure to the warm room was 8 p.u., which corresponds to the baseline state of microcirculation under normal temperature conditions ($20 \pm 1^\circ\text{C}$). After moving to the room with elevated

temperature ($30 \pm 1^\circ\text{C}$), this parameter increased to 27.8 p.u.

Figure 2 presents diagrams showing changes in the perfusion for a group of 15 study participants under different temperature conditions.

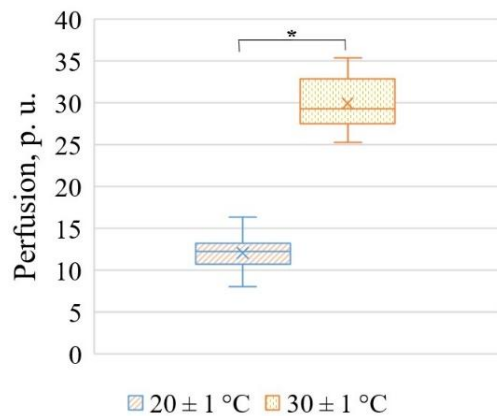


Fig. 2 Diagram of changes in the perfusion for a group of 15 study participants in rooms with air temperatures of $20 \pm 1^\circ\text{C}$ and $30 \pm 1^\circ\text{C}$. *Significance of perfusion difference between the two conditions: $p = 0.0007$.

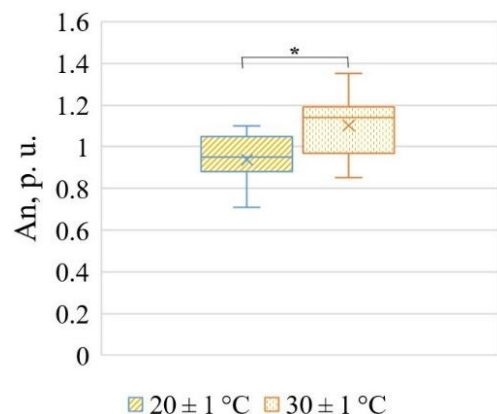


Fig. 3 Diagram of changes in the amplitude of neurogenic oscillations (An) of microcirculation in rooms with air temperatures of $20 \pm 1^\circ\text{C}$ and $30 \pm 1^\circ\text{C}$. *Significance of difference in An values between the two conditions: $p = 0.0026$.

For study participants in the room with a temperature of $20 \pm 1^\circ\text{C}$, the median value of the perfusion was 12.2 p.u.

Increasing the temperature to $30 \pm 1^\circ\text{C}$ led to a significant rise in the microcirculation index, which may indicate a physiological response of the body to external temperature changes. In the cooler environment ($20 \pm 1^\circ\text{C}$), the microcirculation level remained at a relatively low and stable state. Furthermore, under high temperature conditions, greater data variability was observed, potentially reflecting differences in individual responses of participants to the ambient temperature change. Statistical analysis confirmed that the difference in perfusion values between the $20 \pm 1^\circ\text{C}$ room

temperature and the $30 \pm 1^\circ\text{C}$ elevated temperature conditions was statistically significant ($p = 0.0007$).

Of particular interest is the analysis of blood flow oscillations, whose rhythms reflect changes in vascular tone. Figure 3 presents diagrams showing changes in the amplitude of neurogenic blood flow oscillations for a group of 15 study participants under different temperature conditions.

Based on the conducted study, it can be concluded that exposure to elevated temperature induces a statistically significant ($p = 0.0026$) increase in the amplitude of neurogenic oscillations in the blood microcirculation system.

Figure 4 presents diagrams illustrating features of myogenic regulation in a group of 15 study participants under different temperature conditions.

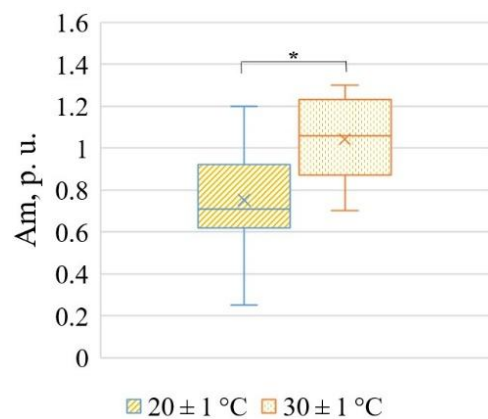


Fig. 4 Diagram of changes in the amplitude of myogenic oscillations (Am) of microcirculation in rooms with air temperatures of $20 \pm 1^\circ\text{C}$ and $30 \pm 1^\circ\text{C}$. *Significance of difference in Am values between the two conditions: $p = 0.001$.

The data presented in Figure 4 show that at a temperature of $20 \pm 1^\circ\text{C}$, the median value of the myogenic oscillation amplitude (Am) was 0.71 r.u. The interquartile range under these conditions varied from 0.62 to 0.92 r.u. The mean value was 0.74 r.u.

The data at $30 \pm 1^\circ\text{C}$ indicates a significant change in the amplitude of myogenic oscillations. The median amplitude value increased to 1.06 r.u., indicating a rise in myogenic activity with increasing temperature. The interquartile range widened, now spanning from 0.87 to 1.23 r.u., suggesting increased variability of the measurements compared to the 20°C condition.

The difference between the Am values under the $20 \pm 1^\circ\text{C}$ room temperature and the $30 \pm 1^\circ\text{C}$ elevated temperature conditions was statistically significant ($p = 0.001$).

3.2 Doppler Ultrasound Measurements of Major Artery Blood Flow

Figures 5a and 5b present an example of the hemodynamic response to thermal exposure typical for the entire subject group. The ultrasound Doppler waveforms of a 25-year-old study participant were

obtained while he was in rooms with air temperatures of $20 \pm 1^\circ\text{C}$ and $30 \pm 1^\circ\text{C}$, respectively.

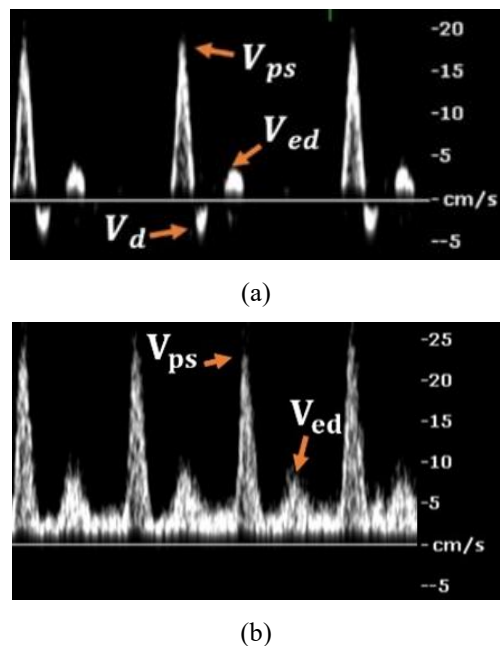


Fig. 5 Ultrasound spectral Doppler waveforms of a 25-year-old study participant showing peak velocities at different time points in the cardiac cycle. (a) in a room with an air temperature of $20 \pm 1^\circ\text{C}$, (b) in a room with an air temperature of $30 \pm 1^\circ\text{C}$.

At an air temperature of $20 \pm 1^\circ\text{C}$, according to Doppler ultrasound data, the peak systolic blood flow velocity (V_{ps}) was 18.5 cm/s, the retrograde blood flow (V_d) had a negative value (-3.5 cm/s), and the maximum antegrade blood flow velocity (V_{ed}) was 3.2 cm/s. Furthermore, the direction of antegrade flow remained positive throughout the entire measurement period.

Under air temperature conditions of $30 \pm 1^\circ\text{C}$, the most pronounced changes concerned the shape of the ultrasound Doppler waveforms. When participants were in the elevated temperature environment, all subjects exhibited alterations in hemodynamic parameters to varying degrees, specifically a significant increase in peak systolic velocity, a decrease in the peak retrograde flow value, and an increase in the peak antegrade flow value. These changes indicate alterations in vascular tone and redistribution of blood flow during thermal exposure.

When the ambient temperature increased to $30 \pm 1^\circ\text{C}$, the study participant (Fig. 5b) showed an increase in peak systolic velocity to 24.3 cm/s. According to Doppler ultrasound data, retrograde blood flow (V_d) was absent, indicating a significant reduction in vascular resistance. The maximum antegrade blood flow velocity (V_{ed}) increased to 7.8 cm/s, while maintaining its positive direction, confirming the primary role of elastic arteries in ensuring blood flow continuity.

Table 1 presents the following parameters calculated for this 25-year-old volunteer: resistance index (RI) and pulsatility index (PI). The parameters at $20 \pm 1^\circ\text{C}$ were

calculated using the equations for high-resistance vessels and the retrograde flow value RI_{V_d} (Eq. (5)), while the parameters at $30 \pm 1^\circ\text{C}$ were calculated using the equations for low-resistance vessels and the antegrade flow value $RI_{V_{ed}}$ (Eq. (6)). All parameters were computed by averaging values across three cardiac cycles.

As shown in Table 1, increasing the ambient temperature to 30°C resulted in a distinct reduction in peripheral resistance indices, which is associated with the development of systemic vasodilation. This effect reflects decreased vascular tone and increased volumetric blood flow.

Figure 6 presents diagrams of the resistance indices RI_{V_d} and $RI_{V_{ed}}$, calculated for a group of 15 study participants under different temperature conditions using Eqs. (5) and (6).

Table 1 Mean parameter values of the 25-year-old study participant under different ambient temperature conditions (based on data from Fig. 5).

Parameter	V_d ($20 \pm 1^\circ\text{C}$)	V_{ed} ($30 \pm 1^\circ\text{C}$)
RI (r.u.)	1.19	0.72
PI (r.u.)	3.64	1.25

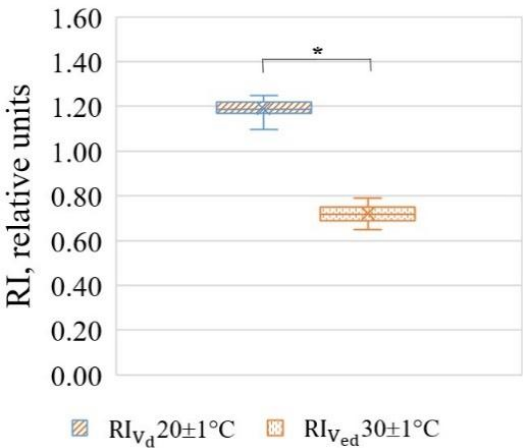


Fig. 6 Diagram of changes in resistance indices RI_{V_d} and $RI_{V_{ed}}$ in rooms with air temperatures of $20 \pm 1^\circ\text{C}$ and $30 \pm 1^\circ\text{C}$. *Significance of difference in RI values between the two conditions: $p = 0.0007$.

The median resistance index decreased by 40%. Statistical analysis performed using the Wilcoxon T -test showed that the differences in resistance index values, calculated using both Eq. (5) and Eq. (6), at ambient temperatures of $20 \pm 1^\circ\text{C}$ and elevated temperature of $30 \pm 1^\circ\text{C}$ were statistically significant ($p = 0.0007$).

Figure 7 presents diagrams showing changes in the pulsatility indices PI_{V_d} and $PI_{V_{ed}}$ for the group of 15 study participants.

The median pulsatility index decreased by 64%. Statistical analysis using the Wilcoxon *T*-test demonstrated that the differences in pulsatility index (PI) values, calculated using both Eq. (7) and Eq. (9), between the ambient temperature of $20 \pm 1^\circ\text{C}$ and the elevated temperature of $30 \pm 1^\circ\text{C}$ were statistically significant ($p = 0.0007$).

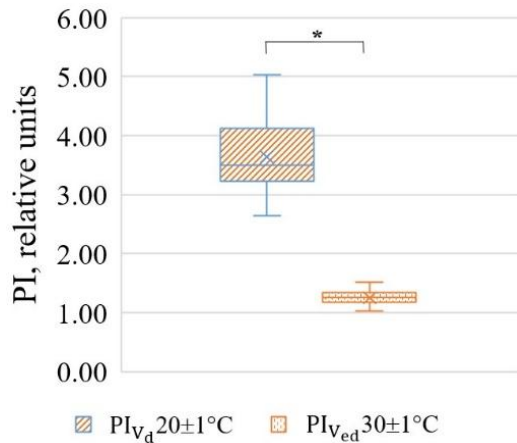


Fig. 7 Diagram of changes in the pulsatility indices PI_{V_d} and $PI_{V_{ed}}$ in rooms with air temperatures of $20 \pm 1^\circ\text{C}$ and $30 \pm 1^\circ\text{C}$. *Significance of difference in PI values between the two conditions: $p = 0.0007$.

4 Discussion

The reduction in arteriolar vascular resistance is attributed to decreased sympathetic tone in the peripheral vascular system, the direct vasodilatory effect of heat on vascular smooth muscles, and the activation of local vasodilators. Under thermoneutral and cold conditions, the sympathetic division of the autonomic nervous system regulates cutaneous blood flow by modulating the frequency of tonic vasoconstrictor impulses [35]. This mechanism is suppressed during heating [22, 36, 27, 25].

Vasodilation can result from both reduced sympathetic activity and the influence of local metabolites (CO_2 , H^+ , K^+ , adenosine, nitric oxide) and humoral factors reflecting current tissue demands for oxygen and nutrients. In some cases, local mechanisms ensuring adequate matching of blood flow to metabolic requirements may play a dominant role in regulating vascular tone [37, 38]. Specifically, activation of $\alpha 1$ -adrenoceptors induces vasoconstriction, whereas stimulation of $\beta 2$ -adrenoceptors leads to vasodilation. Similar conclusions regarding the regulation of vascular tone in muscular-type arteries are reported by Tansey and Johnson [39].

Our analysis of blood flow using laser Doppler flowmetry enabled the determination of the mechanism behind changes in vascular resistance during body heating. The increase in the perfusion reflects enhanced volumetric blood flow in the microcirculatory bed, primarily associated with arteriolar dilation and, consequently, reduced tone in arterial vessels. The

observed increase in the amplitude of neurogenic (An) and myogenic (Am) oscillations during whole-body thermal exposure is also linked to decreased tone in the microcirculatory bed, resulting from smooth muscle relaxation and microvascular vasodilation [40, 41].

Comparison of the results of ultrasound and LDF measurements demonstrates that Doppler ultrasound of the major vessels with the determination of indices RI and PI can serve as a tool for diagnosing the regulation of the general tone of peripheral microvessels. The ultrasound Doppler waveforms clearly reveals the outcome of the interaction between elastic and muscular arteries during heating. Elastic arteries (aorta, pulmonary arteries) are characterized by a predominance of elastic fibers in the tunica media, which enables their ability to stretch during systolic ejection and passively recoil during diastole. This mechanical property ensures continuous antegrade blood flow [42, 43].

Elastic arteries function as a damping mechanism: during systole, their elastic walls stretch due to incoming high-pressure blood, and during diastole, the walls return to their original state, promoting gradual blood movement to the peripheral sections of the vascular system. During body heating, increased cardiac output enhances volumetric blood flow in elastic vessels [24]. These functional characteristics of elastic arteries explain the observed changes in the Doppler spectrum: increased instantaneous and peak amplitudes of blood flow velocity values.

In contrast to elastic-type arteries, muscular-type arteries – including brachial, radial, and femoral vessels – possess a prominent layer of smooth muscle cells in the tunica media. This makes them more capable of active vasoconstriction and vasodilation in response to regulatory mechanisms [44]. Thermal exposure induces vasodilation in the microcirculatory bed, which influences blood flow characteristics in the arteries. This is manifested as reduced peripheral vascular resistance, evidenced by the disappearance of retrograde blood flow due to diminished contractile activity in muscular-type arteries. The decrease in peripheral resistance during heating reflects reduced arteriolar tone. Consequently, a reduction in RI and PI indices is observed during the thermal test.

It should be noted that in the case of certain vascular pathologies, for example, arterial stenosis, the ultrasound Dopplerogram will be strongly distorted, and it will be impossible to assess the tone of the arteries and microcirculatory bed in the corresponding area of the arterial vascular system. If the pathology is associated with atherosclerosis, which leads to increased stiffness of the vascular walls, then the shape of the ultrasound Dopplerogram and the corresponding coefficients will be a marker of changes in the tone of the smooth muscles of the vessels of the microcirculatory bed. Thus, the results obtained may be of practical importance not only in the case of a norm, but also in the case of a malfunction of the mechanisms of regulation of microvascular tone.

5 Conclusions

Laser Doppler flowmetry provides a comprehensive assessment of the vascular response, demonstrating an increase in the perfusion and the amplitudes of neurogenic (An) and myogenic (Am) oscillations during hyperthermia. This indicates the involvement of both sympathetic and local regulatory mechanisms in the process of adaptation to thermal stress.

Ultrasound Doppler proves to be an effective tool for evaluating peripheral resistance in major arteries. Our findings establish that the decrease in ultrasound Doppler ultrasound indices – specifically the resistance index (RI) and pulsatility index (PI), along with the increase in amplitudes of neurogenic (An) and myogenic (Am) oscillations during hyperthermia induced by the general thermal functional test, result from relaxation of smooth muscles and vasodilation of vessels in the microcirculatory bed.

Analysis of ultrasound Doppler data reveals that thermal exposure induces adaptive responses in different types of blood vessels. During whole-body heating, there is a reduction or complete disappearance of retrograde blood flow associated with altered tone in muscular-type vessels. Concurrently, the antegrade blood flow associated with elastic arteries maintains its positive direction.

Thus, the main result of this study is the demonstration that changes in ultrasound Doppler indices in the major arteries during heating can serve as an indicator of vascular resistance in the microcirculatory bed, in particular, in arterioles. Furthermore, ultrasound Doppler enables detection of changes in vascular resistance within a single cardiac cycle, while laser Doppler flowmetry requires substantially more time (minutes) to register comparable changes in vascular resistance compared to ultrasound Doppler.

Funding

The research was supported by a grant from the Russian Science Foundation No. 25-25-01101, <https://rscf.ru/project/25-25-01101>.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

The data are available from the corresponding author on request.

References

1. P. Kvandal, S. A. Landsverk, A. Bernjak, A. Stefanovska, H. D. Kvernmo, and K. A. Kirkebøen, “[Low-frequency oscillations of the laser Doppler perfusion signal in human skin](#),” *Microvascular Research* 72(3), 120–127 (2006).
2. I. Mizeva, P. Frick, and S. Podtaev, “[Relationship of oscillating and average components of laser Doppler flowmetry signal](#),” *Journal of Biomedical Optics* 21(8), 085002 (2016).
3. D. A. Bottino and E. Bouskela, “[Non-invasive techniques to access *in vivo* the skin microcirculation in patients](#),” *Frontiers in Medicine* 9, 1099107 (2023).
4. M. Thoresen, L. Walløe, “[Skin blood flow in humans as a function of environmental temperature measured by ultrasound](#),” *Acta Physiologica Scandinavica* 109(3), 333–341 (1980).
5. T. K. Bergersen, “[A search for arteriovenous anastomoses in human skin using ultrasound Doppler](#),” *Acta Physiologica Scandinavica* 147(2), 195–201 (1993).
6. S. T. Chiesa, S. J. Trangmar, and J. González-Alonso, “[Temperature and blood flow distribution in the human leg during passive heat stress](#),” *Journal of Applied Physiology* 120(9), 1047–1058 (2016).
7. J. S. Petrofsky, L. Berk, F. Alshammari, H. Lee, A. Hamdan, J. E. Yim, Y. Kodawala, D. Patel, B. Nevgi, G. Shetye, H. Moniz, W. T. Chen, M. Alshaharani, K. Pathak, S. Neupane, K. Somanaboina, S. Shenoy, S. Cho, B. Dave, R. Desai, S. Malthane, and H. Al-Nakhli, “[The interrelationship between air temperature and humidity as applied locally to the skin: The resultant response on skin temperature and blood flow with age differences](#),” *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research* 18(4), CR201–CR208 (2012).
8. L. Walløe, “[Arterio-venous anastomoses in the human skin and their role in temperature control](#),” *Temperature* 3(1), 92–103 (2016).
9. I. Mizeva, E. Zharkikh, V. Dremine, E. Zherebtsov, I. Makovik, E. Potapova, and A. Dunaev, “[Spectral analysis of the blood flow in the foot microvascular bed during thermal testing in patients with diabetes mellitus](#),” *Microvascular Research* 120, 13–20 (2018).
10. G. K. Chaseling, C. G. Crandall, and D. Gagnon, “[Skin blood flow measurements during heat stress: technical and analytical considerations](#),” *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 318(1), R57–R69 (2020).
11. Y. Jan, D. M. Brienza, and M. J. Geyer, “[Analysis of week-to-week variability in skin blood flow measurements using wavelet transforms](#),” *Clinical Physiology and Functional Imaging* 25(5), 253–262 (2005).
12. M. M. Mallette, G. J. Hodges, G. W. McGarr, D. A. Gabriel, and S. S. Cheung, “[Spectral analysis of reflex cutaneous vasodilatation during passive heat stress](#),” *Microvascular Research* 111, 42–48 (2017).

13. L. Kralj, H. Lenasi, “Wavelet analysis of laser Doppler microcirculatory signals: Current applications and limitations,” *Frontiers in Physiology* 13, 1076445 (2023).
14. A. A. Fedorovich, “The functional state of regulatory mechanisms of the microcirculatory blood flow in normal conditions and in arterial hypertension according to laser Doppler flowmetry,” *Regional Blood Circulation and Microcirculation* 9(1), 49–60 (2010).
15. A. V. Skripal, Al-Badri Farkad, K. V. Mashkov, A. D. Usanov, and A. P. Averyanov, “Laser flowmetry of microcirculation of the finger depending on the external temperature and the limb position,” *Regional Blood Circulation and Microcirculation* 22(4), 35–41 (2023).
16. C. T. Minson, “Thermal provocation to evaluate microvascular reactivity in human skin,” *Journal of Applied Physiology* 109(4), 1239–1246 (2010).
17. M. Rossi, A. Carpi, F. Galetta, F. Franzoni, and G. Santoro, “The investigation of skin blood flow motion: a new approach to study the microcirculatory impairment in vascular diseases?,” *Biomedicine & Pharmacotherapy* 60(8), 437–442 (2006).
18. M. A. Filina, E. V. Potapova, I. N. Makovik, E. V. Zharkikh, V. V. Dremine, E. A. Zherebtsov, A. V. Dunaev, V. V. Sidorov, A. I. Krupatkin, E. A. Alimicheva, G. I. Masalygina, and V. F. Muradyan, “Functional Changes in Blood Microcirculation in the Skin of the Foot during Heating Tests in Patients with Diabetes Mellitus,” *Human Physiology* 43(6), 693–699 (2017).
19. S. M. Raph, M. M. Dwenger, X. Hu, and M. A. Nystoriak, “Basal NAD(H) redox state permits hydrogen peroxide-induced mesenteric artery dilatation,” *The Journal of Physiology* 601(13), 2621–2634 (2023).
20. E. Ryzhkova, T. Morgunova, E. Potapova, I. Ryzhkov, and V. Fadeyev, “Fluorescence Spectroscopy With Temperature Functional Tests in the Assessment of Markers of Intracellular Energy Metabolism: Spatial Heterogeneity and Reproducibility of Measurements,” *Journal of Biophotonics* 17(10), e202400294 (2024).
21. A. V. Dunaev, V. S. Yanushin, Yu. I. Loktionova, and E. V. Zharkikh, “Capabilities of Human Biotissue Fluorescence Spectroscopy in the Wearable Multimodal Version,” *Modern Technologies in Medicine* 17(3), 29 (2025).
22. H. Korkmaz, M. Akbulut, Y. Özbay, and M. Koç, “A New Noninvasive Method in Evaluating the Endothelial Function: The Measurement of the Resistive Index after Reactive Hyperemia of the Brachial Artery,” *Echocardiography* 27(7), 873–877 (2010).
23. A. Akgul, G. Sasak, C. Basaran, T. Colak, F. N. Ozdemir, and M. Haberal, “Relationship of Renal Resistive Index and Cardiovascular Disease in Renal Transplant Recipients,” *Transplantation Proceedings* 41(7), 2835–2837 (2009).
24. M. Wielicka, J. Neubauer-Geryk, G. Kozera, and L. Bieniaszewski, “Clinical application of pulsatility index,” *Medical Research Journal* 5(3), 201–210 (2020).
25. G. K. Chaseling, A. Debray, H. Gravel, N. Ravanelli, A. Bartlett, and D. Gagnon, “The acute effect of heat exposure on forearm macro- and microvascular function: Impact of measurement timing, heating modality and biological sex,” *Experimental Physiology* 108(2), 221–239 (2023).
26. R. M. Brothers, R. Zhang, J. E. Wingo, K. A. Hubing, and C. G. Crandall, “Effects of heat stress on dynamic cerebral autoregulation during large fluctuations in arterial blood pressure,” *Journal of Applied Physiology* 107(6), 1722–1729 (2009).
27. G. B. Coombs, J. C. Tremblay, D. A. Shkredova, J. M. J. R. Carr, D. J. Wakeham, A. Patrician, and P. N. Ainslie, “Distinct contributions of skin and core temperatures to flow-mediated dilation of the brachial artery following passive heating,” *Journal of Applied Physiology* 130(1), 149–159 (2021).
28. K. N. Thomas, A. M. Van Rij, S. J. E. Lucas, and J. D. Cotter, “Lower-limb hot-water immersion acutely induces beneficial hemodynamic and cardiovascular responses in peripheral arterial disease and healthy, elderly controls,” *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 312(3), R281–R291 (2017).
29. W. F. Jackson, “Myogenic Tone in Peripheral Resistance Arteries and Arterioles: The Pressure Is On!,” *Frontiers in Physiology* 12, 699517 (2021).
30. A. I. Krupatkin, *Laser Doppler flowmetry of blood microcirculation*, Medicine, Moscow (2005). ISBN 5-225-04221-X. [in Russian]
31. A. I. Krupatkin, “The influence of the sympathetic innervation on the skin microvascular tone and blood flow oscillations,” *Human Physiology* 32(5), 584–592 (2006).
32. O. S. Glazachev and E. N. Dudnik, “The functional state of microcirculation and autonomic regulation in young subjects with different levels of subjectively experienced psychological stress,” *Human Physiology* 38(5), 493–499 (2012).
33. I. N. Makovik, A. V. Dunaev, V. V. Dremine, A. I. Krupatkin, V. V. Sidorov, L. S. Khakhicheva, V. F. Muradyan, O. V. Pilipenko, I. E. Rafailov, and K. S. Litvinova, “Detection of angiospastic disorders in the microcirculatory bed using laser diagnostics technologies,” *Journal of Innovative Optical Health Sciences* 11(01), 1750016 (2018).
34. V. G. Lelyuk, and S. E. Lelyuk, *Ultrasound angiology*, 3rd ed., Real Time, Moscow (2007). ISBN 978-5-903025-14-5. [in Russian]
35. T. K. Bergersen, M. Skytjoti, and M. Elstad, “Cold-induced sympathetic tone modifies the impact of endothelium-dependent vasodilation in the finger pulp,” *Autonomic Neuroscience* 203, 97–102 (2017).

36. S. Laurent, P. Boutouyrie, "[The Structural Factor of Hypertension: Large and Small Artery Alterations](#)," *Circulation Research* 116(6), 1007–1021 (2015).
37. A. Sandoo, J. J. C. S. Veldhuijzen Van Zanten, G. S. Metsios, D. Carroll, and G. D. Kitas, "[The Endothelium and Its Role in Regulating Vascular Tone](#)," *The Open Cardiovascular Medicine Journal* 4(1), 302–312 (2010).
38. J. S. Bell, A. O. Adio, A. Pitt, L. Hayman, C. E. Thorn, A. C. Shore, J. L. Whatmore, and C. P. Winlove, "[Microstructure and mechanics of human resistance arteries](#)," *American Journal of Physiology-Heart and Circulatory Physiology* 311(6), H1560–H1568 (2016).
39. E. A. Tansey, C. D. Johnson, "[Recent advances in thermoregulation](#)," *Advances in Physiology Education* 39(3), 139–148 (2015).
40. N. Charkoudian, "[Mechanisms and modifiers of reflex induced cutaneous vasodilation and vasoconstriction in humans](#)," *Journal of Applied Physiology* 109(4), 1221–1228 (2010).
41. J. M. Johnson, D. L. Kellogg, "[Local thermal control of the human cutaneous circulation](#)," *Journal of Applied Physiology* 109(4), 1229–1238 (2010).
42. W. D. Tucker, Y. Arora, and K. Mahajan, "[Anatomy, Blood Vessels](#)," StatPearls Publishing, Treasure Island, FL (2017).
43. M. Ren, X. Li, and M. Xue, "[Aortic Elasticity Evaluated by Pulsed Tissue Doppler Imaging of the Ascending Aorta in Different Diseases: A Systematic Review](#)," *Angiology* 72(5), 403–410 (2021).
44. J. L. Cheng, M. J. MacDonald, "[Effect of heat stress on vascular outcomes in humans](#)," *Journal of Applied Physiology* 126(3), 771–781 (2019).

A Graph Path Interpretability Metric for Explainable Classification of Dermoscopic Skin Lesions

Aleksandra N. Zhdanova

Samara National Research University, 34 Moskovskoe Shosse, Samara, 443086, Russian Federation

e-mail: aleksandra.zhd@yandex.ru

Abstract. Deep neural networks can classify dermoscopic images of skin lesions at a level comparable to that of medical experts. However, their lack of transparency limits the interpretability of model decisions, while existing explainable artificial intelligence techniques often produce pixel-level attributions that are difficult to relate to clinical diagnostic criteria. This study proposes the Graph Path Interpretability (GPI) metric, a quantitative model-independent framework for assessing whether a model decision can be explained through a clinically meaningful path of the form “image – dermoscopic feature – diagnosis”. The study was conducted using the HAM10000 dataset containing 10015 images from seven diagnostic classes. An EfficientNet-B0 classifier achieved an accuracy of 0.81 and an AUC of 0.96 and was combined with five ABCD-rule features and a knowledge graph linking dermoscopic features and diagnoses through statistically significant associations identified using the Benjamini–Hochberg correction. Interpretability was evaluated using the path coverage metric GPI_{cov} , the weighted path metric GPI_w , and a faithfulness test based on sequential concept removal. Three reasoning modules – graph attention, graph convolution, and multilayer perceptron (MLP), were compared under similar classification performance. The value of GPI_{cov} was close to 1.0, indicating that nearly all model decisions could be associated with conceptual paths in the knowledge graph. Under comparable predictive performance, graph-based models generated decisions supported by clinically meaningful paths, whereas the MLP model did not exploit conceptual relationships. The concept-removal test was consistent with the generated explanations. Evaluation on the independent ISIC 2018 dataset yielded an accuracy of 0.78 and an AUC of 0.94, while GPI_{cov} remained close to 1.0. The results indicate that GPI can serve as a complementary quantitative measure of interpretability for dermoscopic image classification models.

Keywords: dermoscopy; skin cancer; explainable artificial intelligence; interpretability; knowledge graph; deep learning.

Paper #9639 received 20 Jun 2026; revised manuscript received 20 Jul 2026; accepted for publication 21 Aug 2026; published online 29 Sep 2026. [doi: 10.18287/JBPE26.12.030304](https://doi.org/10.18287/JBPE26.12.030304).

1 Introduction

Dermoscopy is an optical imaging technique that enhances the visualization of subsurface skin structures and substantially improves the early detection of melanoma and other malignant skin lesions compared with unaided visual examination [1]. Over the past

decade, deep convolutional neural networks (CNNs) trained on large dermoscopic image repositories have achieved performance comparable to, and in some cases exceeding, that of board-certified dermatologists in lesion classification tasks [2, 3]. The release of curated datasets such as HAM10000, which contains 10015 dermoscopic images from seven diagnostic

categories [4], has accelerated progress in this field and enabled reproducible evaluation of machine-learning methods.

Machine learning has also become a central component of optical skin cancer diagnostics in biomedical photonics. Convolutional and transformer-based architectures have been applied to the classification of hyperspectral images of skin lesions [5, 6]; dermoscopic images have been combined with spectral information for lesion identification [7]; and Raman and autofluorescence spectroscopy have been used for optical biopsy of melanoma [8].

Despite the high classification performance of modern deep-learning models, interpreting their decisions remains a challenge for clinical deployment. A model that outputs a prediction such as “melanoma, $p = 0.93$ ” without explaining its rationale limits error analysis, clinical decision-making, and practical adoption [9]. To improve transparency, various explainable artificial intelligence (XAI) techniques have been proposed, including saliency maps, Grad-CAM, and SHAP, which identify influential image regions or estimate feature contributions to a prediction [10, 11]. However, these explanations are often difficult to interpret in clinical terms. Saliency and attribution methods highlight important image regions or features but do not directly relate them to diagnostic criteria routinely used in dermoscopy, such as the ABCD rule and dermoscopic structures [12]. In addition, interpretability is commonly assessed qualitatively, making quantitative comparison between models difficult.

This study introduces the Graph Path Interpretability (GPI) metric, which represents an explanation as a path of the form “image – dermoscopic feature – diagnosis”. Dermoscopic features derived from the ABCD rule are used as the conceptual layer, while relationships between features and diagnostic categories are represented by a knowledge graph [13].

The main contributions of the study are as follows:

- A GPI metric for the quantitative assessment of interpretability in dermoscopic image classification models is proposed.
- A knowledge graph linking ABCD-rule features to the diagnostic categories of the HAM10000 dataset is developed.
- Graph-based and non-graph-based models are compared under similar classification performance, and their interpretability is evaluated quantitatively.

2 Materials and Methods

2.1 Dataset and Data Splitting

The HAM10000 dataset [4] was used in this study. It contains 10015 dermoscopic images from seven diagnostic categories: melanocytic nevi (nv), melanoma (mel), benign keratosis-like lesions (bkl), basal cell carcinoma (bcc), actinic keratoses and intraepithelial carcinoma (akiec), vascular lesions (vasc), and

dermatofibroma (df). The class distribution is highly imbalanced, with melanocytic nevi representing the largest category. Because images may correspond to the same lesion, the dataset was split into training, validation, and test subsets using the lesion identifier (lesion_id). This prevented images of the same lesion from appearing in different subsets, avoiding information leakage between training and evaluation data.

The dataset was divided into training, validation, and test sets in an approximate ratio of 70:15:15, corresponding to 7008, 1515, and 1492 images, respectively. To ensure reproducibility of the experiments, a fixed random seed (seed = 42) was used throughout all data-splitting procedures.

2.2 Baseline Classification Model

EfficientNet-B0 [14], pretrained on the ImageNet dataset, was used as the baseline classification model and subsequently fine-tuned to classify dermoscopic images into seven diagnostic categories. All images were resized to 224×224 pixels.

To address class imbalance, a class-weighted loss was employed, with weights inversely proportional to class frequencies. Model training used the AdamW optimizer (learning rate = 1×10^{-4} , weight decay = 1×10^{-4}) for 30 epochs. The final model was selected according to the highest validation accuracy. Classification performance was evaluated using Accuracy, Balanced Accuracy, macro-F1, and macro-AUC. Ninety-five-percent confidence intervals for all test-set metrics were obtained by non-parametric bootstrap resampling of the test set (2000 resamples).

2.3 Dermoscopic Features

Five interpretable features corresponding to the ABCD rule [12] were extracted for each image. All features were normalized to the range [0, 1]:

- Asymmetry (A) – the degree of mismatch between the lesion mask and its reflections across the principal axes passing through the centroid.
- Border irregularity (B) was defined as $1 - \text{circularity}$, where circularity is calculated as $4\pi \times \text{area} / \text{perimeter}^2$.
- Color variation (C) was measured as the number of color clusters obtained using k-means clustering in the CIE-Lab color space and occupying at least 5% of the lesion area.
- Diameter (D) – the equivalent lesion diameter normalized by image size.
- Structure (S) – structural heterogeneity measured as the normalized Shannon entropy of the lesion brightness histogram.

Lesion masks from the HAM10000 segmentation dataset [4] were used whenever available. For images without provided segmentations, masks were generated using Otsu thresholding applied to the inverted grayscale image. Since all features are computed directly from the image and lesion mask, they are independent of the

classification architecture, ensuring the model-independent nature of the proposed GPI metric.

2.4 Knowledge Graph Linking Features and Diagnoses

A knowledge graph linking dermoscopic features and diagnostic categories was constructed. Graph edges represent the strength and statistical significance of associations between feature c and diagnosis k , estimated exclusively on the training set.

Edge weights were computed using the point-biserial correlation coefficient r_{ck} between the feature value and a one-vs-rest binary class indicator. Multiple hypothesis testing was controlled using the Benjamini-Hochberg procedure. Associations were considered statistically significant when $q < 0.05$ and $r > 0$. Only statistically significant feature–diagnosis associations were included in the graph. Out of 35 possible feature–diagnosis pairs, 11 positive associations satisfied the significance criteria and were retained as graph edges.

2.5 Reasoning Modules

To compare interpretability, three reasoning modules were trained using the same set of five dermoscopic features. Classification was performed on the basis of a shared conceptual feature representation [15].

GAT – feature nodes are aggregated using a graph attention mechanism [16]. The embedding produced by the baseline classifier serves as the attention query. The resulting feature weights w_i , normalized with a softmax function, are interpreted as path weights in the computation of GPI.

GCN – feature nodes are aggregated through graph convolution without an attention mechanism. In the GPI calculation, path weights are determined solely by the presence of the corresponding features.

MLP – the five-dimensional feature vector is provided directly to a multilayer perceptron without using graph structure.

All models were trained under identical optimization settings (AdamW, 50 epochs) and selected according to validation-set performance. The use of a common conceptual representation minimizes the influence of classification accuracy differences when comparing model interpretability.

2.6 GPI Metric

For each image, explanation paths of the form “image – feature – diagnosis” are considered. A path is regarded as valid if feature i is present in the image $c_i > \tau$, $\tau = 0.2$) and the association between this feature and the predicted class $\hat{y}_n$ is statistically significant $q < 0.05$.

The path coverage metric quantifies the proportion of predictions for which at least one valid path exists between the image and the predicted class:

$$GPI_{cov} = \frac{1}{N} \sum_{n=1}^N I_n,$$

where $I_n = 1$ if at least one feature i satisfies $c_i^{(n)} > \tau$ and has a significant edge to the predicted class $\hat{y}_n$, and $I_n = 0$ otherwise.

The weighted interpretability metric measures the proportion of the model contribution associated with valid explanation paths:

$$GPI_w = \frac{1}{N} \sum_1^N \frac{\sum_i w_i^{(n)} \cdot s_i^{(n)}}{\sum_i w_i^{(n)}},$$

where, w_i denotes the importance of feature i for the reasoning module, whereas s_i indicates the clinical validity of the corresponding path for the predicted class. Consequently, GPI_w represents the fraction of the model decision weight attributable to clinically valid explanation paths.

For the GAT model, the values w_i correspond to attention coefficients. For the GCN model, w_i are derived from normalized feature values. The metric GPI_w is not computed for the MLP model because it does not produce explicit feature importance weights.

Both metrics are bounded within the interval [0,1]. High values of GPI_{cov} indicate that explanations can be constructed for most model decisions using paths in the knowledge graph. High values of GPI_w indicate that a substantial proportion of the model contribution is associated with valid feature–diagnosis paths.

2.7 Faithfulness Evaluation

To assess the consistency between generated explanations and model behaviour, a sequential concept-removal test was employed [17]. Features were iteratively removed in descending order of their importance w_i , and the resulting change in the confidence of the predicted class was measured.

If the removal of the most important features leads to a substantial decrease in model confidence, the explanation is considered faithful to the underlying decision-making process. Quantitative evaluation was performed using the area under the deletion curve (Deletion AUC), where lower values indicate a stronger influence of the removed features on the classification outcome.

For the MLP model, which does not generate explicit feature importance weights, the feature-removal order was selected randomly and used as a baseline control condition.

3 Results

3.1 Classification Performance

The EfficientNet-B0 model achieved an Accuracy of 0.807, a Balanced Accuracy of 0.677, a macro-F1 score of 0.668, a macro-AUC of 0.955 on the test set (Fig. 1).

The corresponding 95% bootstrap confidence intervals were [0.787, 0.826] for accuracy, [0.622, 0.728] for balanced accuracy, [0.617, 0.708] for macro-F1 and [0.947, 0.962] for macro-AUC; a per-split comparison is given in Table 1.

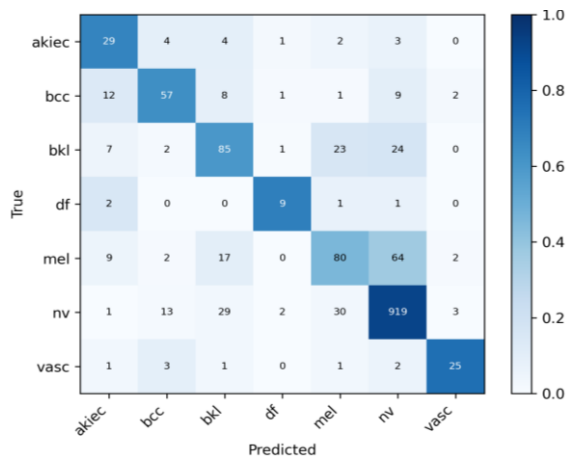


Fig. 1 Normalized confusion matrix for the EfficientNet-B0 classifier on the HAM10000 test set ($n = 1492$). Rows are true diagnoses and columns predicted diagnoses; each row is normalized to sum to one, so the diagonal gives the per-class recall. Class abbreviations: akiec – actinic keratosis, bcc – basal cell carcinoma, bkl – benign keratosis-like lesion, df – dermatofibroma, mel – melanoma, nv – melanocytic nevus, vasc – vascular lesion.

The highest classification performance was obtained for the melanocytic nevus class nv, which represents the largest category in the dataset ($n = 997$ in the test set). The lowest F1 scores were observed for rare classes, particularly dermatofibroma (df, $n=13$), reflecting the pronounced class imbalance of the HAM10000 dataset.

Table 1 Classification performance of the EfficientNet-B0 model on the training, validation and test splits (7008/1515/1492 images). The close agreement between the validation and test metrics, together with the narrow bootstrap confidence intervals, indicates that the reported test performance is a stable estimate; the high training accuracy reflects the expected fit to the training split under early stopping on the validation set.

Split	Accuracy	Balanced Accuracy	macro-F1	macro-AUC
Training	0.996	0.999	0.993	1.000
Validation	0.831	0.692	0.697	0.957
Test	0.807	0.677	0.668	0.955
Test, 95% CI	0.787–0.826	0.622–0.728	0.617–0.708	0.947–0.962

Table 2 Point-biserial correlation coefficients (r) between dermoscopic features (rows) and diagnostic categories (columns), computed on the HAM10000 training set ($n = 7008$). Positive coefficients indicate association with the corresponding diagnosis. Decimal separators are dots.

Feature	akiec	bcc	bkl	df	mel	nv	vasc
Asymmetry	0.02	−0.01	0.00	−0.06	0.03	−0.01	−0.04
Border Irregularity	0.12	−0.01	0.15	−0.06	0.13	−0.19	−0.10
Color Variation	−0.05	−0.19	−0.06	−0.01	0.02	0.14	−0.01
Diameter	0.13	−0.01	0.10	0.03	0.10	−0.17	−0.09
Structural Heterogeneity	−0.09	−0.13	−0.14	−0.10	0.22	0.07	−0.01

Abbreviations: akiec – actinic keratosis, bcc – basal cell carcinoma, bkl – benign keratosis-like lesion, df – dermatofibroma, mel – melanoma, nv – melanocytic nevus, vasc – vascular lesion.

The obtained results are consistent with previously reported performance levels for models of similar complexity on the HAM10000 dataset [4] and were subsequently used as the baseline for interpretability evaluation.

3.2 Knowledge Graph of Feature-Diagnosis Associations

The results of the correlation analysis are summarized in Table 2, and the corresponding knowledge graph is shown in Fig. 2. Of the 35 examined feature-diagnosis pairs, 25 were found to be statistically significant $q < 0.05$, including eleven positive and fourteen negative associations.

Only positive associations $r > 0$, were included in the knowledge graph, resulting in a total of eleven edges.

The largest number of associations were identified for melanoma. This class showed statistically significant positive correlations with structural heterogeneity ($r = 0.22$), border irregularity ($r = 0.13$), lesion diameter ($r = 0.10$), and asymmetry ($r = 0.03$). Positive associations with border irregularity and diameter were also observed for actinic keratosis and benign keratosis-like lesions. No positive associations were identified for basal cell carcinoma or vascular lesions.

The identified relationships are consistent with established clinical knowledge regarding dermoscopic characteristics of skin lesions. In particular, the strongest associations observed for melanoma involved structural heterogeneity and border irregularity.

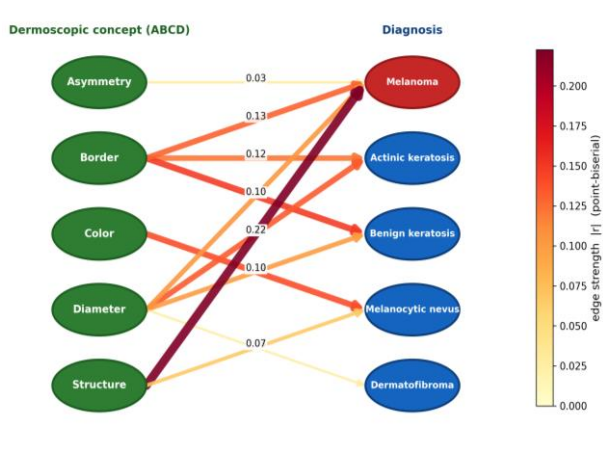


Fig. 2 Knowledge graph of the eleven statistically significant positive feature–diagnosis associations (Benjamini-Hochberg $q < 0.05$, $r > 0$) obtained on the HAM10000 training set. Green nodes are the five ABCD-derived dermoscopic concepts (left) and coloured nodes are diagnoses (right); a directed edge concept $\rightarrow$ diagnosis denotes a significant positive point-biserial correlation, with edge thickness and colour proportional to $|r|$ (colour bar). Basal cell carcinoma (bcc) and vascular lesions (vasc) have no positive edge and are omitted, as their only significant associations are negative (Table 2).

The complete results of the correlation analysis are presented in Table 2. Two observations are noteworthy. First, the absence of positive edges for basal cell carcinoma and vascular lesions does not imply the absence of feature–diagnosis relationships. For these classes, statistically significant negative associations were identified but were excluded from the knowledge graph according to the adopted edge-construction rule. Second, the only positive association involving the color feature was observed for melanocytic nevi $r = 0.14$, whereas no statistically significant association was found

for melanoma. This result may be attributable to the use of the number of color clusters as a measure of color variation and suggests that more detailed color descriptors could improve the representation of chromatic information.

3.3 Comparison of Reasoning Architectures Using GPI

The comparison results for the reasoning modules are presented in Table 3 and Fig. 3. All models used the same set of five conceptual features and differed only in the way these features were aggregated. The Accuracy values for GAT, GCN, and MLP were 0.670, 0.668, and 0.676, respectively, allowing interpretability to be compared under similar classification performance. For all models, the value of GPI_{cov} was close to unity (1.000 for GAT and GCN, and 0.999 for MLP), indicating that at least one knowledge-graph path linking the observed features to the predicted diagnostic class existed for almost all predictions.

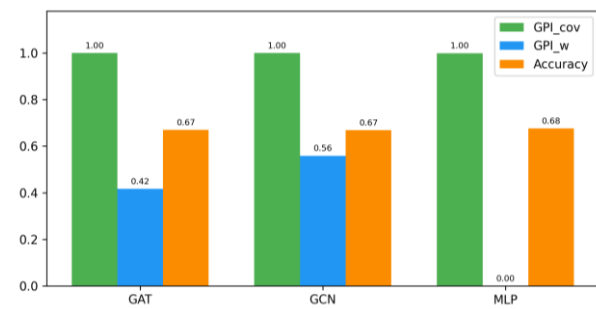


Fig. 3 Interpretability metrics (GPI_{cov} and GPI_w) and classification accuracy of the three reasoning modules (GAT, GCN, MLP). Despite similar accuracy (0.67–0.68), only the graph-based modules achieve non-zero GPI_w , indicating reliance on clinically valid feature–diagnosis paths.

Table 3 Test-set classification accuracy, interpretability metrics (path coverage GPI_{cov} and weighted interpretability GPI_w) and faithfulness (deletion AUC) for the three reasoning modules trained on the same five dermoscopic features of the full HAM10000 dataset (10015 images; test $n = 1492$). A lower deletion AUC denotes a more faithful explanation; for the MLP the value is a random-order control (Fig. 4). The 95% bootstrap confidence intervals, calculated from 2000 resamples of the test set, overlap for all three modules (GAT [0.64, 0.69], GCN [0.65, 0.69], and MLP [0.65, 0.70]), indicating that their accuracies are statistically indistinguishable; for the graph-based modules, $GPI_{cov} = [1.00, 1.00]$. The within-test-set sampling variability of GPI_w is negligible; its dominant source of variation is the training random seed, which is why the GAT-GCN ordering of GPI_w is treated as indicative rather than definitive.

Model	Test Accuracy	GPI_{cov}	GPI_w	Deletion AUC
GAT	0.670	1.000	0.416	0.885
GCN	0.668	1.000	0.559	0.953
MLP	0.676	0.999	0.000	0.691 (Random Baseline)

Abbreviations: GAT – graph attention, GCN – graph convolution, MLP – multilayer-perceptron baseline without graph structure.

The largest differences were observed for GPI_w . Its values were 0.416 for GAT and 0.559 for GCN, whereas GPI_w was equal to 0 for MLP. This result indicates that a substantial fraction of the decision process in graph-based models is associated with clinically meaningful paths represented in the knowledge graph. In contrast, the MLP model does not utilize conceptual relationships and therefore cannot assign decision weight to knowledge-graph paths.

3.4 Faithfulness Evaluation

The results of the sequential feature-removal test are presented in Fig. 4, which shows, for each module, how the predicted-class probability decays as the most important dermoscopic features are progressively removed.

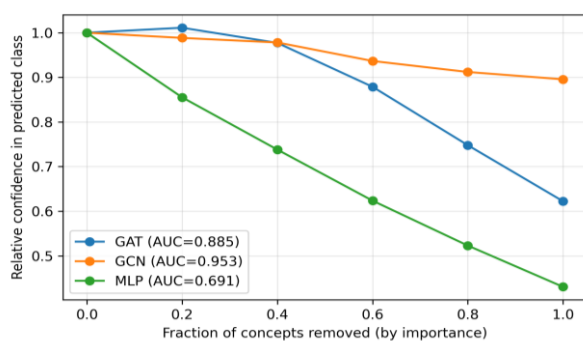


Fig. 4 Faithfulness assessment by sequential concept removal. The five dermoscopic features are deleted one at a time in decreasing order of their model-assigned importance w_i ; the curve shows the retained probability of the predicted class (normalized to its undeleted value, y-axis) versus the fraction of features removed (x-axis). Removing the most decisive features first produces an early, steep decline, so a smaller area under the deletion curve (deletion AUC, given in the legend) indicates a more faithful explanation. Among the graph-based modules GAT (0.885) is more faithful than GCN (0.953). The MLP produces no importance weights; its features are removed in random order and it serves only as a control, so its value (0.691) is not a faithfulness measure and is not comparable with the graph-based modules.

For the GAT model, the area under the deletion curve was 0.885, whereas for GCN it was 0.953. A lower Deletion AUC indicates a larger decrease in model confidence when the most important features are removed and therefore reflects a stronger correspondence between the generated attributions and the underlying decision-making process.

These results suggest that the attention coefficients learned by GAT provide a more faithful representation of feature contributions than the uniform aggregation mechanism employed by GCN. The observed differences between the models remained consistent both on a subset of 5000 images and on the full dataset.

For the MLP model, the feature-removal order was selected randomly because the architecture does not

generate explicit feature importance weights. Consequently, the deletion results for MLP were used only as a baseline control and were not interpreted as a measure of attribution quality.

An example of an explanation path with a GPI_w value of 0.76 for an individual image is shown in Fig. 5.

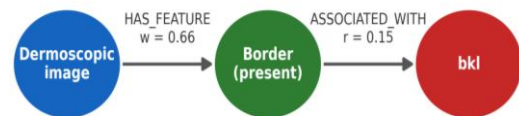


Fig. 5 Example of an explanation path for a single image: the path links the dermoscopic image to a present ABCD feature (edge HAS_FEATURE, weight w equal to the feature value) and then to the predicted diagnosis (edge ASSOCIATED_WITH, weight equal to the point-biserial coefficient r of the corresponding knowledge-graph edge). The path-level interpretability score for this decision is $GPI_w = 0.76$.

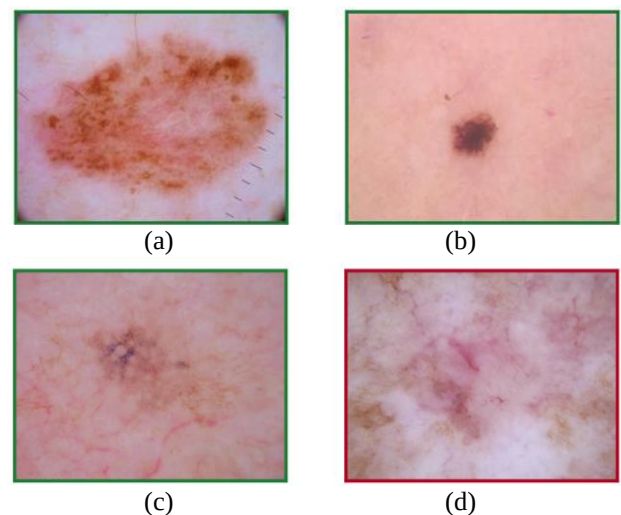


Fig. 6 Explanation paths of the GPI metric on four representative HAM10000 test images. Green frames indicate decisions with a valid explanation path based on a present ABCD feature significantly associated with the true diagnosis; the red frame indicates a case with no valid positive feature-diagnosis path in the ABCD knowledge graph. (a) melanoma explained by structural heterogeneity; (b) melanocytic nevus explained by color variation; (c) benign keratosis-like lesion explained by border irregularity; (d) basal cell carcinoma with no valid positive feature-diagnosis path. Feature abbreviations: A – asymmetry, B – border irregularity, C – color variation, D – diameter, S – structural heterogeneity; r is the point-biserial edge weight (Table 2).

To illustrate how the metric evaluates the ABCD rule on individual images, Fig. 6 shows four representative test cases. In the three lesions with a valid explanation path – melanoma via the Structure→melanoma edge ($r = 0.22$), melanocytic nevus via Color→nevus ($r = 0.14$) and benign keratosis via Border→bkl ($r = 0.15$) – the ABCD features present in the image

connect through a statistically significant edge to the correct diagnosis, so the decision is explainable and contributes to GPI_{cov} . For the basal cell carcinoma none of the present features has a significant positive edge to the bcc class (consistent with bcc showing only negative feature associations, Table 2), so no valid path exists and the metric correctly reports the decision as not explainable by the ABCD graph. These examples confirm that the metric operates directly on the ABCD-rule concepts and distinguishes explainable from non-explainable decisions.

3.5 External Validation on an Independent Dataset

To assess the generalization capability of the proposed approach, an additional evaluation was performed using the independent ISIC 2018 Task 3 dataset, which contains 1511 dermoscopic images from seven diagnostic categories. Model training was conducted on the HAM10000 dataset (8831 images for training and 1184 for validation, split by lesion_id), after which the model was evaluated on ISIC 2018 without further fine-tuning.

On the external dataset, the classifier achieved an Accuracy of 0.784, a Balanced Accuracy of 0.644, a macro-F1 score of 0.660, and a macro-AUC of 0.939. Compared with the results obtained on HAM10000 (0.807, 0.677, 0.668, and 0.955, respectively), only a moderate decrease in performance was observed, indicating that the model retained its effectiveness on previously unseen data.

Correlation analysis performed on the combined training data confirmed the main relationships between dermoscopic features and diagnostic categories. The strongest association remained that between structural heterogeneity and melanoma, with $r = 0.22$ and $q \approx 7 \times 10^{-93}$. In total, 13 statistically significant positive associations were identified.

The primary interpretability metric, GPI_{cov} , remained close to unity, indicating that explanation paths through the knowledge graph were preserved for the vast majority of predictions on the independent dataset as well (Fig. 7).

The conceptual reasoning modules based solely on five dermoscopic features exhibited a more pronounced decrease in performance on the independent dataset than the baseline classifier. This reduction is likely attributable to the limited information content of the compact feature set and differences in lesion segmentation quality. Whereas expert-provided lesion masks were available for HAM10000 images, segmentation masks for part of the external dataset were generated automatically using Otsu thresholding, which may have introduced additional errors into feature extraction.

Despite this reduction in predictive performance, the main findings related to interpretability remained unchanged. The knowledge graph reproduced the principal associations between dermoscopic features and diagnostic categories, and the value of GPI_{cov} remained close to unity. A more detailed investigation of the

robustness of GPI_w on independent datasets will require more reliable methods for extracting conceptual features and is considered a direction for future research.

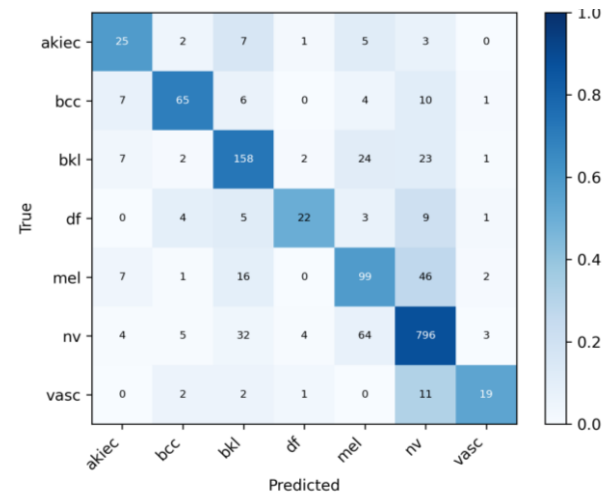


Fig. 7 Normalized confusion matrix for the classifier trained on HAM10000 and evaluated, without retraining, on the independent ISIC 2018 Task 3 test set ($n = 1511$). Rows are true and columns predicted diagnoses (row-normalized); the preserved diagonal structure indicates that the backbone generalizes across datasets. Class abbreviations as in Fig. 1.

4 Discussion

The results demonstrate that the interpretability of dermoscopic image classifiers can be quantified using diagnostic features routinely employed in clinical practice. Unlike saliency maps and other visual explanation techniques, the proposed approach links model decisions to specific ABCD-rule features and statistically validated associations between features and diagnoses. As a result, explanations are expressed in terms familiar to clinicians and can be directly related to established diagnostic criteria.

The problem of making deep dermoscopic classifiers interpretable has motivated a growing body of work. Post-hoc attribution methods such as saliency maps [18], LIME [19], class activation mapping [20], integrated gradients [21] and layer-wise relevance propagation [22] highlight influential pixels, whereas concept-based approaches such as TCAV [23], concept bottleneck models [15], prototype networks [24] and automatic concept extraction [25] relate predictions to higher-level concepts. In dermatology, multimodal and taxonomy-driven frameworks such as ExAID [26] and explainable diagnosis using taxonomies [27] generate concept-level justifications, and recent surveys of dermoscopic feature analysis [28] and of medical explainable AI [29] emphasise that interpretability should be expressed through the diagnostic criteria that clinicians already use. The present knowledge-graph formulation follows this line but adds a quantitative, model-agnostic score rather than a purely qualitative explanation.

The proposed metrics capture different aspects of interpretability. GPI_{cov} reflects the existence of at least one explanation path through the knowledge graph, whereas GPI_w quantifies the extent to which such paths contribute to the final prediction. In the present study, explanation paths were identified for nearly all predictions, both on the original dataset and during external validation. Differences between models were observed primarily in the degree to which these paths contributed to decision making.

The proposed framework distinguishes between models that explicitly exploit conceptual features through a knowledge graph and models that operate without a graph-based representation of knowledge. Under comparable classification performance, graph-based models exhibited a stronger correspondence between model decisions and clinically meaningful explanation paths. The sequential feature-removal analysis further confirmed the consistency of these explanations with model behaviour. Here the graph-based modules used graph attention [16] and graph convolution [30] to aggregate the conceptual features.

The faithfulness of explanations was assessed by sequential concept removal, following the rationale of remove-and-evaluate protocols [17, 31] and their known caveats [32]. Such perturbation-based measures, together with rationale-alignment benchmarks [33] and quantitative evaluations of saliency [34], have become standard tools for verifying that a highlighted feature is actually used by the model. In our setting the deletion analysis separated the graph-based reasoning modules from the black-box baseline, indicating that the GPI paths reflect genuine model behaviour rather than a post-hoc rationalisation.

The use of conceptual features did not lead to a substantial reduction in classification performance compared with models lacking an explicit knowledge representation. This observation is consistent with previous studies on concept-based learning and suggests that interpretability and predictive performance can be combined within a single architecture.

The proposed method complements existing explainable artificial intelligence approaches. Techniques such as Grad-CAM identify image regions that influence model predictions, whereas the graph-based approach relates predictions to diagnostic features and their relationships with diagnostic categories. Combining these complementary forms of explanation represents a promising direction for future research.

From the perspective of clinical deployment, the practical value of the proposed approach lies in expressing each decision through the ABCD criteria [12, 35] that dermatologists apply routinely, which may lower the barrier to trust and adoption relative to opaque predictions [36]. Modern computer-aided diagnosis systems, built on deep convolutional [37, 38] and transformer [39] architectures, already match or exceed dermatologist performance on curated benchmarks [2, 3, 40, 41], and classifiers trained on large clinical image collections have reached specialist-level accuracy

[42, 43]. Importantly, the largest gains in prospective studies arise from human–computer collaboration rather than from autonomous prediction [44]; an interpretability score that exposes whether a model reasons along clinically validated feature–diagnosis paths is directly useful in this collaborative setting. Optical modalities beyond dermoscopy – hyperspectral imaging [5, 6], combined dermoscopic-spectral analysis [7] and Raman or autofluorescence spectroscopy [8] – are increasingly used for skin-cancer diagnosis, and the modality-independent GPI metric could complement them by quantifying how far a model’s reasoning follows established feature–diagnosis relationships. Rather than competing with existing diagnostic pipelines, the metric is intended as an auxiliary, quantitative interpretability criterion reported alongside accuracy when a model is selected or certified for clinical use.

Several limitations should be acknowledged. The study relied on a simplified set of conceptual features derived from the ABCD rule, feature quality depended on lesion-segmentation accuracy, and evaluation was performed on a limited number of datasets. Expanding the feature vocabulary to include more detailed dermoscopic structures and conducting multi-centre validation studies may further improve the practical value of the approach. Future work will also investigate the applicability of the proposed metric to other medical imaging tasks and compare it with existing interpretability assessment methods. Broader validation on additional public collections such as the ISIC 2018 challenge data [45], BCN20000 [46] and seven-point-checklist datasets [47], together with formal notions of interpretability [48], is a natural next step.

5 Conclusions

A GPI metric was proposed for the quantitative assessment of interpretability in dermoscopic image classification models. The metric is based on the analysis of paths linking an image, dermoscopic features, and diagnostic categories through a knowledge graph.

Experiments on the HAM10000 dataset (10015 images) showed that the EfficientNet-B0 classifier reached an accuracy of 0.807, a balanced accuracy of 0.677, a macro-F1 of 0.668 and a macro-AUC of 0.955. A knowledge graph built from ABCD-rule features and comprising 11 statistically significant positive feature–diagnosis associations ($q < 0.05$) provided explanation paths for virtually all classifier decisions ($GPI_{cov} = 1.0$). Under comparable classification accuracy (0.67–0.68 for all reasoning modules), the proposed metric clearly separated models that explicitly use conceptual features ($GPI_w = 0.416$ for GAT and 0.559 for GCN) from the graph-free baseline ($GPI_w = 0$ for MLP). The sequential concept-removal analysis confirmed the faithfulness of the explanations, the deletion AUC being 0.885 for GAT and 0.953 for GCN versus 0.691 for MLP.

External validation on the independent ISIC 2018 dataset reproduced the principal feature–diagnosis associations and preserved interpretability

($GPI_{cov} \approx 1.0$), while classification accuracy decreased to 0.784. These results demonstrate the feasibility of quantitatively assessing interpretability using diagnostic features employed in clinical practice. The proposed GPI metric can serve as an additional criterion for the development and comparison of artificial intelligence models for dermoscopic image analysis.

Disclosures

All authors declare that there is no conflict of interest in this paper.

References

1. G. Argenziano, H. P. Soyer, “[Dermoscopy of pigmented skin lesions – a valuable tool for early](#),” *The Lancet Oncology* 2(7), 443–449 (2001).
2. A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, “[Dermatologist-level classification of skin cancer with deep neural networks](#),” *Nature* 542(7639), 115–118 (2017).
3. T. J. Brinker, A. Hekler, A. H. Enk, et al., “[Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task](#),” *European Journal of Cancer* 113, 47–54 (2019).
4. P. Tschandl, C. Rosendahl, and H. Kittler, “[The HAM10000 dataset, a large collection of multi-source dermoscopic images of common pigmented skin lesions](#),” *Scientific Data* 5(1), 180161 (2018).
5. B. V. Grechkin, V. O. Vinokurov, Yu. A. Khristoforova, and I. A. Matveeva, “[VGG convolutional neural network classification of hyperspectral images of skin neoplasms](#),” *Journal of Biomedical Photonics & Engineering* 9(4), 040304 (2023).
6. S. N., V. A. R., S. R., K. N., P. M. M., and J. K., “[Highly Accurate Skin Cancer Diagnosis Using HMT-NET and Vision Transformer Models](#),” *Journal of Biomedical Photonics & Engineering* 11(4), 040308 (2025).
7. I. A. Matveeva, A. I. Komlev, O. I. Kaganov, A. A. Moryatov, and V. P. Zakharov, “[Multidimensional analysis of dermoscopic images and spectral information for the diagnosis of skin tumors](#),” *Journal of Biomedical Photonics & Engineering* 10(1), 010307 (2024).
8. I. A. Bratchenko, Yu. A. Khristoforova, L. A. Bratchenko, A. A. Moryatov, S. V. Kozlov, E. G. Borisova, and V. P. Zakharov, “[Optical biopsy of amelanotic melanoma with Raman and autofluorescence spectra stimulated by 785 nm laser excitation](#),” *Journal of Biomedical Photonics & Engineering* 7(2), 020308 (2021).
9. A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bénéttot, S. Tabik, A. Barbado, S. García, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, “[Explainable Artificial Intelligence \(XAI\): Concepts, taxonomies, opportunities and challenges toward responsible AI](#),” *Information Fusion* 58, 82–115 (2020).
10. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “[Grad-CAM: visual explanations from deep networks via gradient-based localization](#),” in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, IEEE, 618–626 (2017).
11. S. M. Lundberg, S.-I. Lee, “[A unified approach to interpreting model predictions](#),” in *Advances in Neural Information Processing Systems* 30, 4765–4774 (2017).
12. F. Nachbar, W. Stolz, T. Merkle, A. B. Cognetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun-Falco, and G. Plewig, “[The ABCD rule of dermoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions](#),” *Journal of the American Academy of Dermatology* 30(4), 551–559 (1994).
13. A. N. Zhdanova, “[Ontological approach to predicting psychological characteristics of social media users using knowledge graphs](#),” *Artificial Intelligence and Decision Making* 3, 40–59 (2025). [in Russian]
14. M. Tan, Q. V. Le, “[EfficientNet: rethinking model scaling for convolutional neural networks](#),” in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, PMLR 97, 6105–6114 (2019).
15. P. W. Koh, T. Nguyen, Y. S. Tang, S. Mussmann, E. Pierson, B. Kim, and P. Liang, “[Concept bottleneck models](#),” in *Proceedings of the 37th International Conference on Machine Learning (ICML)*, *Proceedings of Machine Learning Research* 119, 5338–5348 (2020).
16. P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, “[Graph Attention Networks](#),” *arXiv preprint arXiv:1710.10903* (2017).
17. V. Petsiuk, A. Das, and K. Saenko, “[RISE: Randomized Input Sampling for Explanation of Black-box Models](#),” in *Proceedings of the British Machine Vision Conference (BMVC)* (2018).
18. K. Simonyan, A. Vedaldi, and A. Zisserman, “[Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps](#),” *arXiv preprint arXiv:1312.6034* (2013).
19. M. T. Ribeiro, S. Singh, and C. Guestrin, “[“Why Should I Trust You?”: Explaining the Predictions of Any Classifier](#),” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM, 2016)*, pp. 1135–1144.
20. B. Zhou, A. Khosla, A. Lapedrizza, A. Oliva, and A. Torralba, “[Learning deep features for discriminative localization](#),” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2921–2929 (2016).

21. M. Sundararajan, A. Taly, and Q. Yan, “[Axiomatic attribution for deep networks](#),” in Proceedings of the 34th International Conference on Machine Learning (ICML), PMLR 70, 3319–3328 (2017).
22. S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Müller, and W. Samek, “[On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation](#),” PLoS ONE 10(7), e0130140 (2015).
23. B. Kim, M. Wattenberg, J. Gilmer, C. Cai, J. Wexler, F. Viégas, and R. Sayres, “[Interpretability beyond feature attribution: quantitative testing with concept activation vectors \(TCAV\)](#),” in Proceedings of the 35th International Conference on Machine Learning (ICML), PMLR 80, 2668–2677 (2018).
24. C. Chen, O. Li, D. Tao, A. Barnett, C. Rudin, and J. Su, “[This looks like that: deep learning for interpretable image recognition](#),” in Advances in Neural Information Processing Systems 32, 8930–8941 (2019).
25. A. Ghorbani, J. Wexler, J. Y. Zou, and B. Kim, “[Towards automatic concept-based explanations](#),” in Advances in Neural Information Processing Systems 32, 9273–9282 (2019).
26. A. Lucieri, M. N. Bajwa, S. A. Braun, M. I. Malik, A. Dengel, and S. Ahmed, “[ExAID: A multimodal explanation framework for computer-aided diagnosis of skin lesions](#),” Computer Methods and Programs in Biomedicine 215, 106620 (2022).
27. C. Barata, M. E. Celebi, and J. S. Marques, “[Explainable skin lesion diagnosis using taxonomies](#),” Pattern Recognition 110, 107413 (2021).
28. C. Barata, M. E. Celebi, and J. S. Marques, “[A survey of feature extraction in dermoscopy image analysis of skin cancer](#),” IEEE Journal of Biomedical and Health Informatics 23(3), 1096–1109 (2019).
29. E. Tjoa, C. Guan, “[A survey on explainable artificial intelligence \(XAI\): toward medical XAI](#),” IEEE Transactions on Neural Networks and Learning Systems 32(11), 4793–4813 (2021).
30. T. N. Kipf, M. Welling, “[Semi-Supervised Classification with Graph Convolutional Networks](#),” arXiv preprint arXiv:1609.02907 (2016).
31. S. Hooker, D. Erhan, P.-J. Kindermans, and B. Kim, “[A benchmark for interpretability methods in deep neural networks](#),” in Advances in Neural Information Processing Systems 32, 9737–9748 (2019).
32. J. Adebayo, J. Gilmer, M. Muelly, I. Goodfellow, M. Hardt, and B. Kim, “[Sanity checks for saliency maps](#),” in Advances in Neural Information Processing Systems 31, 9505–9515 (2018).
33. J. DeYoung, S. Jain, N. F. Rajani, E. Lehman, C. Xiong, R. Socher, and B. C. Wallace, “[ERASER: a benchmark to evaluate rationalized NLP models](#),” in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (Association for Computational Linguistics, 2020), pp. 4443–4458.
34. W. Samek, A. Binder, G. Montavon, S. Lapuschkin, and K.-R. Müller, “[Evaluating the Visualization of What a Deep Neural Network Has Learned](#),” IEEE transactions on neural networks and learning systems 28(11), 2660–2673 (2017).
35. G. Argenziano, H. P. Soyer, S. Chimenti, et al., “[Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet](#),” Journal of the American Academy of Dermatology 48(5), 679–693 (2003).
36. C. Rudin, “[Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead](#),” Nature Machine Intelligence 1(5), 206–215 (2019).
37. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “[ImageNet classification with deep convolutional neural networks](#),” in Advances in Neural Information Processing Systems 25, 1097–1105 (2012).
38. K. He, X. Zhang, S. Ren, and J. Sun, “[Deep residual learning for image recognition](#),” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778 (2016).
39. A. Dosovitskiy, L. Beyer, A. Kolesnikov, “[An image is worth 16×16 words: transformers for image recognition at scale](#),” arXiv preprint arXiv:2010.11929 (2021).
40. H. A. Haenssle, C. Fink, R. Schneiderbauer, et al., “[Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists](#),” Annals of Oncology 29(8), 1836–1842 (2018).
41. P. Tschandl, N. Codella, B. N. Akay, G. Argenziano, R. P. Braun, H. Cabo, D. Gutman, A. Halpern, B. Helba, R. Hofmann-Wellenhof, A. Lallas, J. Lapins, C. Longo, J. Malvehy, M. A. Marchetti, A. Marghoob, S. Menzies, A. Oakley, J. Paoli, S. Puig, C. Rinner, C. Rosendahl, A. Scope, C. Sinz, H. P. Soyer, L. Thomas, I. Zalaudek, and H. Kittler, “[Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study](#),” The Lancet Oncology 20(7), 938–947 (2019).
42. S. S. Han, M. S. Kim, W. Lim, G. H. Park, I. Park, and S. E. Chang, “[Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm](#),” Journal of Investigative Dermatology 138(7), 1529–1538 (2018).
43. L. Yu, H. Chen, Q. Dou, J. Qin, and P. A. Heng, “[Automated melanoma recognition in dermoscopy images via very deep residual networks](#),” IEEE Transactions on Medical Imaging 36(4), 994–1004 (2017).
44. P. Tschandl, C. Rinner, Z. Apalla, G. Argenziano, N. Codella, A. Halpern, M. Janda, A. Lallas, C. Longo, J. Malvehy, J. Paoli, S. Puig, C. Rosendahl, H. P. Soyer, I. Zalaudek, and H. Kittler, “[Human-computer collaboration for skin cancer recognition](#),” Nature Medicine 26(8), 1229–1234 (2020).

45. N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler, and A. Halpern, “[Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration \(ISIC\)](#),” arXiv preprint arXiv:1902.03368 (2019).
46. M. Combalia, N. C. F. Codella, V. Rotemberg, B. Helba, V. Vilaplana, O. Reiter, C. Carrera, A. Barreiro, A. C. Halpern, S. Puig, and J. Malvehy, “[BCN20000: Dermoscopic Lesions in the Wild](#),” arXiv preprint arXiv:1908.02288 (2019).
47. J. Kawahara, S. Daneshvar, G. Argenziano, and G. Hamarneh, “[Seven-point checklist and skin lesion classification using multitask multimodal neural nets](#),” IEEE Journal of Biomedical and Health Informatics 23(2), 538–546 (2019).
48. F. Doshi-Velez, B. Kim, “[Towards A Rigorous Science of Interpretable Machine Learning](#),” arXiv preprint arXiv:1702.08608 (2017).

Leakage-Free Biomedical Acoustic Signal Classification of Infant Cry for Reliable Neonatal Monitoring

Humam Abdullah Khudair^{1*}, Ziyad K. Farej², and Omar Younis Alani³

¹ Engineering Technical College – Mosul, Northern Technical University, Mosul, Iraq

² Northern Technical University, Mosul, Iraq

³ School of Science, Engineering and Environment, University of Salford, Manchester, United Kingdom

*e-mail: eng.humam@ntu.edu.iq

Abstract. Infant cry analysis is a non-invasive biomedical acoustic signal processing task that can support reliable neonatal monitoring. However, reported high classification accuracy may be affected by dataset imbalance and data leakage rather than true model generalization. This study investigates the effect of leakage-free evaluation and hybrid acoustic feature representation on infant cry classification under imbalanced data conditions. A file-level partitioning protocol was applied before augmentation, feature scaling, and model training to preserve independence between training and testing data. Hybrid acoustic features, including Mel Frequency Cepstral Coefficients (MFCC) statistics, root-mean-square (RMS) energy, zero-crossing rate, pitch-related descriptors, and spectral features, were extracted from eight infant cry categories. Four lightweight machine learning models were evaluated using accuracy, Macro-F1, and balanced accuracy. Under leakage-free evaluation, support vector machines (SVMs) achieved the most balanced performance with 86.41% accuracy, 0.7376 Macro-F1, and 68.43% balanced accuracy. In contrast, leaky evaluation inflated performance above 90% across models. These findings indicate that reliable infant cry classification depends on leakage-free evaluation, imbalance-aware metrics, and biomedical acoustic feature integration rather than model complexity alone.

Keywords: infant cry classification; biomedical acoustic signal processing; neonatal monitoring; leakage-free evaluation; hybrid acoustic features; support vector machine (SVM).

Paper #9636 received 17 Jun 2026; revised manuscript received 21 Jul 2026; accepted for publication 27 Jul 2026; published online 30 Sep 2026. doi: [10.18287/JBPE26.12.030305](https://doi.org/10.18287/JBPE26.12.030305).

1 Introduction

The cry classification of infants is a significant research area in biomedical acoustic signal processing since crying is an early non-invasive behavioral cue of the infant's needs and physiological discomfort. Cry signals can offer complementary data that can be used for neonatal monitoring, such as hunger, pain, fatigue, thermal stress, fear, loneliness, or overall distress. Unlike conventional physiological measures, cry analysis does not necessitate direct contact with the body and thus can become an extra layer of behavior sensing in smart neonatal care systems, and even be used for continuous monitoring [1].

Although there is potential for reliable classification of infant cries, it continues to be a challenge. Cry signals

are highly non-stationary and affected by the infant's physiological state, emotional state and recording environment [2]. This means that there can be acoustic similarities between different cry categories, which makes the classification problem more complicated than just labelling. Discriminative information is spread over multiple acoustic dimensions such as spectral structure, temporal irregularity, vocal intensity, pitch behavior and harmonic content. Hence, one feature type may not be enough to capture the entire structure of infant cry signals [3].

The other major problem is methodological reliability. High classification accuracy has been reported in several studies, but this could be due to data imbalance, inadequate evaluation control, or data leakage. Leakage happens when two samples, which are acoustically

similar or enhanced versions of the same sample, are in both the training and testing sets. In this situation, the model might learn the patterns of the recordings instead of the cry patterns. Furthermore, the accuracy of the overall performance may be misleading in imbalanced data sets because good performance on the majority class may mask poor performance on the minority classes. Therefore, for a reliable performance assessment it is crucial to have leakage-free evaluation and imbalance-sensitive metrics like Macro-F1 and balanced accuracy [4].

To overcome these challenges, this study aims to introduce a leakage-free infant cry classification framework using a combination of hybrid biomedical acoustic features and lightweight machine learning. To ensure independence of the test set, the original recordings are split at the file level prior to augmentation, normalization and model training [5]. A hybrid of feature representation is extracted to include complementary spectral, temporal, energy-based and pitch-related features of the cry signals. Then, four lightweight machine learning models are tested under the same experimental conditions with accuracy, Macro-F1 and balanced accuracy [6].

What is most important in this study is that the classification of infant cries is reliable and that the consistency between acoustic representation, leakage-free evaluation design, and model suitability are crucial. This offers a more accurate foundation for the incorporation of cry analysis into smart neonatal monitoring systems.

2 Literature Review

Cry analysis of infants has become a biomedical acoustic signal processing problem which seeks to extract clinically relevant behavior from raw vocalizations. Previous studies typically used a two-stage pipeline: acoustic signal representation and classification using machine learning. Mel-frequency cepstral coefficients (MFCCs) are still popular handcrafted descriptors due to their ability to give a compact and perceptually meaningful representation of the spectral envelope of infant cry signals. But infant crying is very non-stationary and is not completely described by the spectral envelope information [7]. Other cry characteristics that can vary between distress cries and non-distress cries include temporal instability, intensity of cries, pitch of cries, harmonic structure, and frequency distribution [8]. Therefore, recent studies have turned to hybrid acoustic representations, which include the MFCCs along with root-mean-square (RMS) energy, zero-crossing rate, pitch-related descriptors, spectral centroid, spectral bandwidth, spectral contrast, chroma and Mel-spectrogram statistics. This multi-domain representation is necessary because the discriminative structure of infant cry categories is not always determined by a single acoustic cue, but is more often the result of the interaction of complementary acoustic cues [9].

Infant cry classification has been a focus point of machine learning. Structured acoustic descriptors have been shown to be useful for classification, for example

by classical models like support vector machine (SVM), k-nearest neighbors (KNN), Random Forest or boosting based classifiers, especially when data is limited. The nonlinear decision boundaries of SVM are particularly important, as they can model complex relationships in high dimensional hybrid feature spaces [10]. Meanwhile, deep learning methods such as convolutional neural networks (CNNs), long short-term memory (LSTM), CNN-LSTM, and spectrogram-based models have been used to directly learn time-frequency patterns from the transformed acoustic data. While these models can enhance the representational ability, they strongly rely on the size of the data set, class balance, recording conditions, and evaluation protocol. Thus, complexity of the model does not necessarily imply generalization between small, imbalanced biomedical acoustic datasets [11].

One of the major drawbacks in literature is that high accuracy could be due to methodological artifacts, not necessarily to recognition ability. Leakage may happen if augmented or acoustically correlated samples are used in both the training and testing sets, which results in models learning the patterns of the recording rather than the cry characteristics. Moreover, the overall accuracy may be deceiving in cases of severe class imbalance as good performance on majority class can hide poor performance on minority classes. Hence, to achieve credible evaluation, it is essential to use leakage-free partitioning, training-only augmented data, and imbalance sensitive measures like Macro-F1 and balanced accuracy [12].

In terms of systems, infant cry analysis is best utilized as a non-contact behavioral sensing layer on smart neonatal monitoring platforms. It should be used in conjunction with the physiological and environmental monitoring and not as a substitute for clinical diagnosis [13]. Therefore, models which are interpretable and edge-deployable are of crucial importance for practical applications in the neonatal space, where the computational cost is low, the ability to make local inferences is required, and the models must be robust to real acoustic conditions [14]. This inspires the current research, which focuses on leakage-free evaluation, hybrid acoustic feature fusion and lightweight machine learning for robust infant cry classification.

3 Materials and Methods

3.1 Dataset Description and Preprocessing

The dataset for this study is comprised of 512 original infant cries, divided into eight categories: 1) belly pain, 2) burping, 3) cold/hot, 4) discomfort, 5) hungry, 6) lonely, 7) scared, and 8) tired. The data set is very skewed, with the hungry class being the majority. This imbalance is due to the practical challenge of infant cry classification and thus, evaluation metrics need to take into account class-wise behavior [15]. The distribution of the original recordings across the eight cry categories is presented in Table 1.

Table 1 Distribution of infant cry samples across categories.

Class Label	Number of Files	Percentage (%)
Belly Pain	16	3.13%
Burping	18	3.52%
Cold/Hot	7	1.37%
Discomfort	30	5.86%
Hungry	382	74.61%
Lonely	11	2.15%
Scared	20	3.91%
Tired	28	5.47%
Total	512	100%

3.2 Acoustic Feature Extraction

The feature extraction was designed to be a multi domain representation of the audio, not just a Mel Frequency Cepstral Coefficients (MFCC). The spectral envelope was measured in MFCC statistics, the RMS energy was used for vocal intensity and respiratory effort, the zero-crossing rate (ZCR) was used for irregularity of the vocal

excitation, and pitch-related descriptors for laryngeal tension and vocal stability. Other features, such as chroma, spectral contrast, spectral centroid, spectral bandwidth and Mel-spectrogram statistics, were related to harmonic structure and frequency-energy distribution [16]. This hybrid representation combines complementary spectral, temporal, energy-based and pitch related information to minimize the confusion of acoustically identical cry classes. The audio recordings were resampled to 22,050 Hz and cut to a maximum length of 5 s. For each recording, 13 MFCC coefficients were extracted, and the mean and standard deviation of each coefficient were calculated. Chroma and spectral contrast were computed with the Librosa default parameters and presented as average values. The temporal averages of the spectral centroid, spectral bandwidth, zero crossing rate and RMS energy were used. Global mean and standard deviation were calculated for mel-spectrograms obtained from 64 Mel frequency bands, and added as additional acoustic descriptors. The Librosa Python library was used for all the procedures for feature extraction. The extracted acoustic features and their interpretations are summarized in Table 2. Representative acoustic representations of an infant cry signal, including the waveform, Mel-spectrogram, and MFCCs, are shown in Fig. 1.

Table 2 Extracted acoustic features and their physiological interpretation.

Feature Category	Feature Name	Description	Physiological Interpretation
Spectral	MFCC (Mean, Std)	Compact representation of spectral envelope	Captures vocal tract characteristics and overall acoustic structure of the cry
Temporal	Zero-Crossing Rate (ZCR)	Rate of sign changes in the signal	Reflects excitation irregularity and instability in vocal fold vibration
Energy-based	RMS Energy	Root Mean Square of signal amplitude	Indicates respiratory effort and vocal intensity during crying
Pitch-related	Fundamental Frequency (Pitch Mean)	Average pitch over the segment	Represents laryngeal tension and emotional intensity (e.g., pain vs. discomfort)
Pitch-related	Pitch Variability	Variation of pitch across time	Indicates instability or modulation in vocal production
Spectral	Spectral Centroid	Center of mass of the spectrum	Reflects perceived brightness or sharpness of the cry
Spectral	Spectral Bandwidth	Spread of frequencies around the centroid	Indicates distribution of acoustic energy across frequencies
Spectral	Spectral Contrast	Difference between spectral peaks and valleys	Captures harmonic structure and resonance differences
Harmonic	Chroma Features	Distribution of energy across pitch classes	Represents harmonic content and tonal structure of vocalization
Spectral	Mel-Spectrogram (Statistics)	Time–frequency energy representation	Encodes global acoustic patterns and temporal evolution of cry signals
Dynamic (Temporal)	Δ MFCC (Delta)	First-order derivative of MFCC	Captures temporal transitions in spectral structure
Dynamic (Temporal)	Δ^2 MFCC (Delta-Delta)	Second-order derivative of MFCC	Models acceleration and dynamic variation in cry patterns

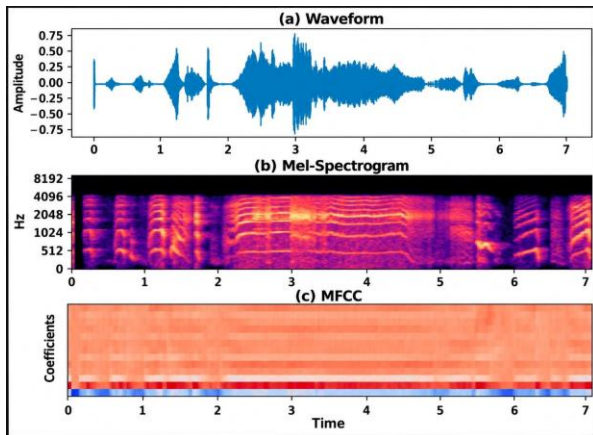


Fig. 1 Acoustic representations of infant cry signals: (a) waveform, (b) Mel-spectrogram, (c) MFCC.

3.3 Evaluation Protocol and Model Comparison

The four lightweight classifiers were tested under the same preprocessing and feature environment: radial basis function (RBF)-SVM, Random Forest, Gradient Boosting and multilayer perceptron (MLP). For severe imbalance, three metrics were used to assess performance: accuracy, Macro-F1, and balanced accuracy. The same models were also tested using a leaky protocol to investigate methodological sensitivity. These results were only used as a comparative measure of performance inflation and were not considered to be valid deployment performance [17]. The leakage-free evaluation workflow is illustrated in Fig. 2. The classifier hyperparameters are summarized in Table 3.

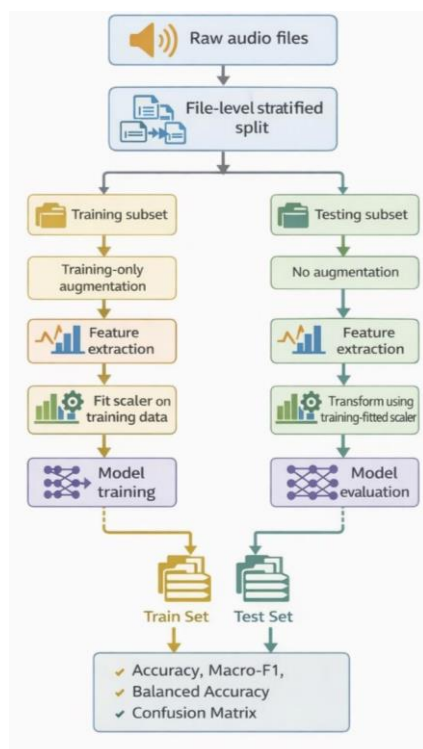


Fig. 2 Leakage-free evaluation pipeline.

Table 3 Hyperparameter Settings of the Evaluated Classifiers.

Model	Hyperparameter Settings
SVM (RBF)	Kernel = RBF, C = 5, Gamma = scale
Random Forest	Number of trees = 250, Max depth = None (unrestricted), Random state = 42
Gradient Boosting	Number of estimators = 200, Learning rate = 0.1, Random state = 42
MLP	Hidden layers = (256, 128), Dropout = (0.35, 0.25), Optimizer = Adam, Loss = Categorical Cross-Entropy, Batch size = 16, Epochs = 100, Early Stopping (patience = 10, restore_best_weights = True)

The hyperparameter settings listed in Table 3 were fixed throughout all experiments to ensure fair comparison among the evaluated classifiers. The experiments were done in Python with the following libraries: Scikit-learn, TensorFlow (Keras), Librosa, NumPy, Pandas, Matplotlib and Seaborn.

4 Results and Discussion

4.1 Data Characteristics and Evaluation Conditions

The dataset also bears a high level of imbalance in the classes, as hungry takes an overbearing size in the sample space. Such an imbalance directly affects the behavior of the classifier since models do not need to be overly accurate in order to obtain relatively high overall accuracy by promoting majority categories. This is why the results that are being reported are mostly interpreted using Macro-F1 and Balanced Accuracy as opposed to accuracy [18]. In order to reduce the class imbalance, the data augmentation was performed only on the training subset following the file-level partitioning. One of the following augmentation methods was applied to each of the songs randomly: additive Gaussian noise, time stretching ($0.75\text{--}0.95\times$ or $1.05\text{--}1.25\times$), pitch shifting (-3 to -1 or $+1$ to $+3$ semitones), or no transformation. The number of training samples per class for minority classes was increased to 180 and the testing set was not augmented. The class-wise distribution of the training and testing sets is presented in Table 4.

4.2 Leakage-Free Performance

The leakage-free protocol yielded the most balanced performance for SVM with accuracy of 86.41%, Macro-F1 of 0.7376 and balanced accuracy of 68.43%.

Table 4 Dataset distribution under the leakage-free protocol.

Cry Category	Original Files	Training Files (Before Augmentation)	Training Samples (After Augmentation)	Testing Files
Belly Pain	16	13	180	3
Burping	18	14	180	4
Cold–Hot	7	6	180	1
Discomfort	30	24	180	6
Hungry	382	305	305	77
Lonely	11	9	180	2
Scared	20	16	180	4
Tired	28	22	180	6
Total	512	409	1565	103

Table 5 Model performance under leakage-free evaluation.

Model	Accuracy (%)	Macro-F1	Balanced Accuracy (%)
SVM (RBF)	86.41	0.7376	68.43
Random Forest	86.41	0.7068	65.46
MLP	81.55	0.6289	65.69
Gradient Boosting	84.47	0.5649	53.68

Random Forest had the same accuracy, but lower Macro-F1 and balanced accuracy scores mean that it is less sensitive to the minority classes. The performance of the evaluated classifiers is summarized in Table 5. The results of MLP and Gradient Boosting indicated a lower class-balanced reliability, thus illustrating that overall accuracy can overestimate the model's performance in the extreme case of class imbalance. Overall, SVM provided the best trade-off between accuracy and class-wise robustness.

4.3 Effect of Data Leakage

All models exhibited significant overperformance when using the leaky protocol. The results showed that MLP obtained the highest accuracy of 94.83%, with the Macro-F1 and balanced accuracy of 0.9483 and 95.02%, respectively, while SVM gave the accuracy of 91.79% and the balanced accuracy of 92.10%. This improvement was observed across models, and was primarily due to the evaluation protocol, and not to the model architecture. The results show that data leakage can greatly overestimate the reliability of classification, and leakage-free evaluation is a prerequisite for meaningful biomedical acoustic evaluation. The model-wise performance comparison is shown in Fig. 3. The effect of data leakage on classification performance is

illustrated in Fig. 4. The detailed results under the leaky evaluation protocol are reported in Table 6.

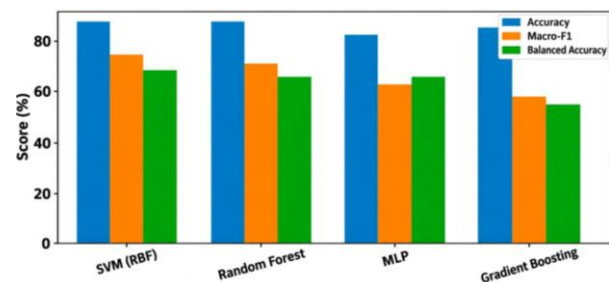


Fig. 3 Model performance comparison.

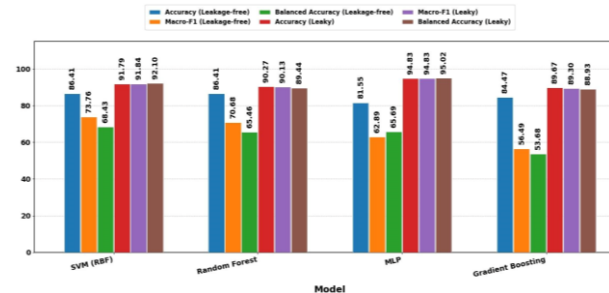


Fig. 4 Effect of data leakage on performance.

Table 6 Model performance under leaky evaluation.

Model	Accuracy (%)	Macro-F1	Balanced Accuracy (%)
SVM (RBF)	91.79	0.9184	92.10
Random Forest	90.27	0.9013	89.44
MLP	94.83	0.9483	95.02
Gradient Boosting	89.67	0.8930	88.93

4.4 Structure and Confusion Patterns of Errors

The analysis of confusion matrix also explains the behavior of models. The leakage-free SVM results depict high diagonal concentration of dominant classes, which are an indication that the model represents the major structure of the dataset. Nonetheless, there are still systematic cases of misclassification between acoustically similar classes, especially between classes which have overlapping energy and pitch properties.

The misclassification errors are not arbitrary and mean something underlying in the feature space. As an illustration, the similarity between the forms of discomfort and tiredness can be explained by their comparable RMS distribution and limited differences in pitch variation, whereas the overlap between the scared and pain-related conditions can be the excessively high pitch in both cases.

Unlike SVM, this model is not as separable with respect to classes, which are represented by the higher the off-diagonal values of the model. This demonstrates the significance of nonlinear separation processes in the determination of the finer acoustic variations. The confusion matrices of the SVM and Gradient Boosting models are presented in Figs. 5 and 6, respectively.

4.5 Multidimensional Feature Space Analysis

To facilitate qualitative interpretation of the multidimensional acoustic feature representation, selected combinations of representative acoustic descriptors (e.g., RMS energy, zero-crossing rate, and pitch-related features) were visualized. These visualizations provide an intuitive representation of the feature distribution rather than a formal dimensionality reduction technique such as principal component analysis (PCA), *t*-distributed stochastic neighbor embedding (*t*-SNE), or uniform manifold approximation and projection (UMAP).

When individual features are viewed separately, it is demonstrated in the visualization that they do not offer clear separation. Rather, the class clusters are just found when features are studied together in a multidimensional

space. Intersection between clusters is a positive indication that the classification problem is inherently nonlinear. This fact is the direct cause of the high performance of SVM with an RBF kernel which can capture smooth surfaces with the underlying geometry of the feature space. Conversely, simple partitioning strategies are not effective in models that use the simplistic strategies. The pairwise distributions of the selected acoustic features are illustrated in Fig. 7.

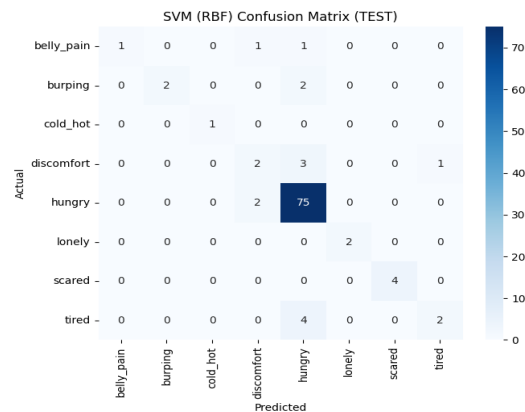


Fig. 5 Confusion matrix of the SVM model.

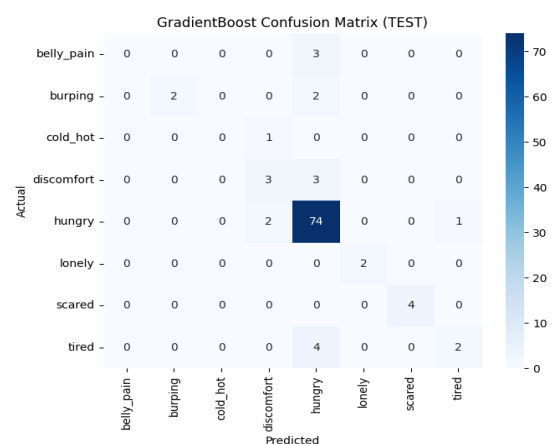


Fig. 6 Confusion matrix of the Gradient Boosting model.

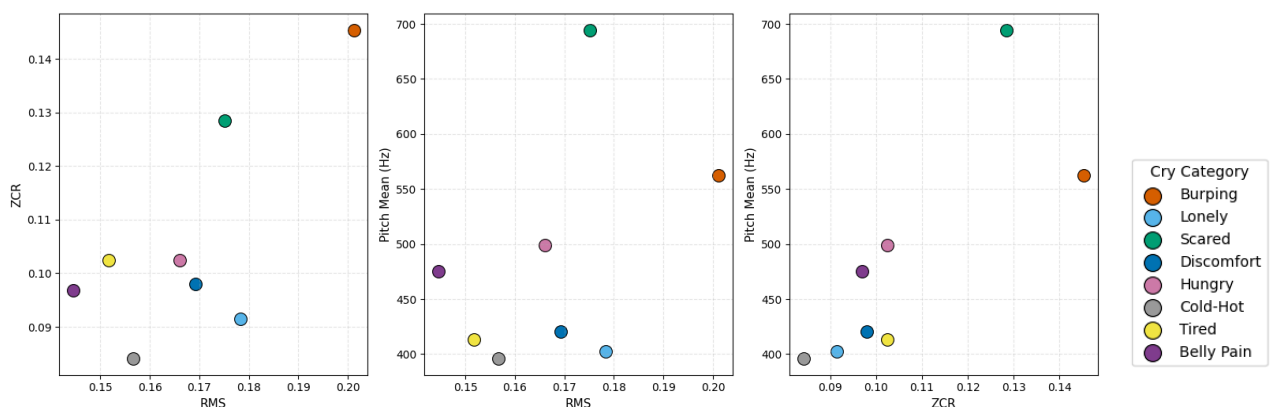


Fig. 7 Acoustic feature space visualization. (a) RMS vs ZCR, (b) RMS vs Pitch Mean, and (c) ZCR vs Pitch Mean.

4.6 Interpretation of Cry Classification As a Feature

The classification behaviour observed indicates that the recognition of infant crying is more related to the interaction of a multitude of acoustic features than to the complexity of the model. The RMS energy represents the intensity of the voice and the effort used to produce it, the ZCR measures the irregularity of the voice's excitation, and the pitch-related descriptors show the tension and stability of the voice. Variations in pitch, temporal irregularity and energy thus combine to produce differences between the cry categories, not just one of these. The results are consistent with the use of hybrid acoustic feature integration to be a more reliable representation of distinguishing overlapping infant cry patterns. Class-wise acoustic feature characteristics are summarized in Table 7.

4.7 Integration into Smart Neonatal Monitoring Systems

For easier implementation, the proposed classification system can be embedded in a smart neonatal monitoring system as illustrated in Fig. 8. The system integrates environmental sensors, physiological sensors and an acoustic sensing unit based on the microphone. The features extracted and processed on the edge device (Raspberry Pi) are then fed into the SVM model for real time inference. The summarized outputs can then be sent remotely via an IEEE 802.11ac wireless network for scalable multi-incubator monitoring [19]. This architecture is not designed to use cry classification as a diagnostic tool, but rather as a behavioral inference layer, working in conjunction with physiological and environmental monitoring, and to help raise the alert of the caregiver [20].

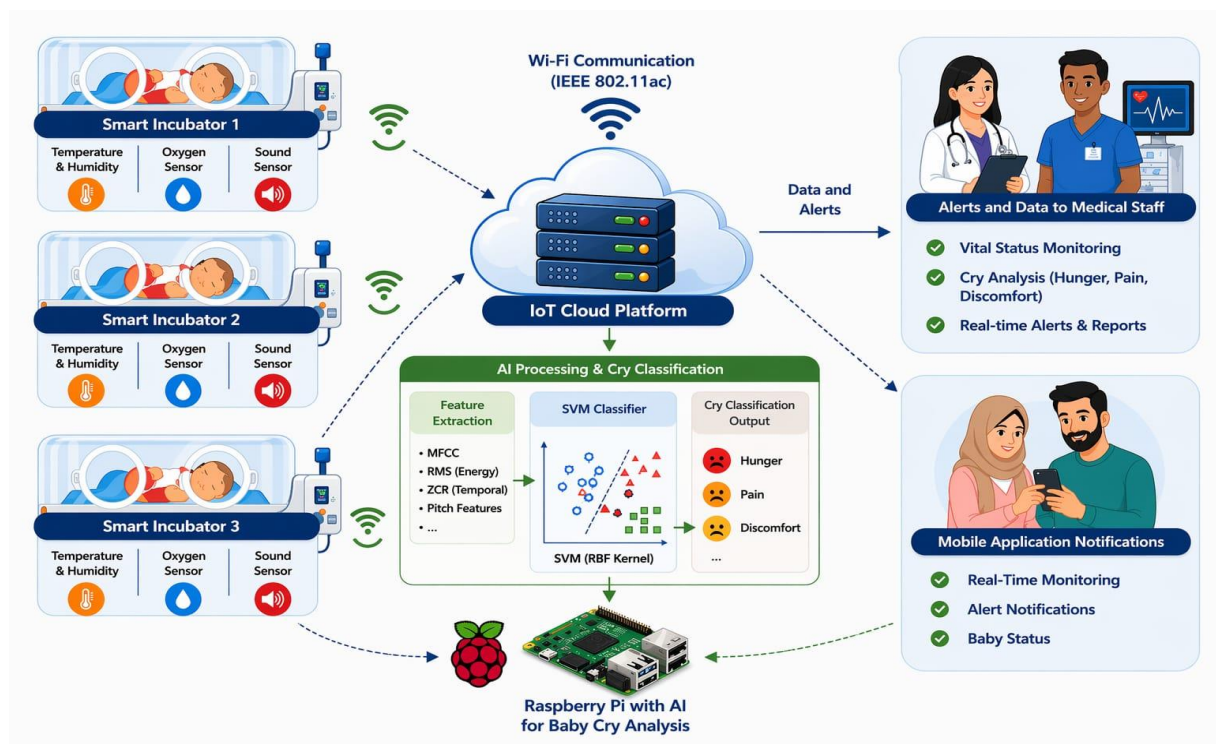


Fig. 8 IoT-based neonatal monitoring architecture.

Table 7 Acoustic feature characteristics across cry categories.

Cry Category	RMS (After)	ZCR (After)	Pitch Mean (Hz)	Pitch Variability (Hz)	Dominant Acoustic Profile	Physiological / Behavioral Interpretation
Burping	0.201178	0.145341	562.609645	164.036446	High energy, high irregularity, moderately high pitch, very high variability	Explosive airflow bursts with unstable excitation; energy–excitation dominated vocalization
Loneliness	0.178320	0.091464	402.334851	54.401323	Relatively low irregularity, low-to-moderate pitch, very low variability	Sustained and stable phonation with limited emotional escalation; low-activation cry

Table 7 (Cont.)

Scared	0.175196	0.128498	694.229246	125.675044	High irregularity, very high pitch, high variability	Emotionally aroused and unstable vocal excitation; fear-related high-arousal cry
Discomfort	0.169246	0.097998	420.154448	116.828038	Relatively high energy, moderate irregularity, moderate pitch, moderate variability	Sustained physical irritation with controlled but effortful phonation; energy-dominant discomfort cry
Hungry	0.166051	0.102475	499.155349	143.861676	Moderate energy, moderate irregularity, moderately high pitch, high variability	Rhythmic and repetitive distress signaling with balanced multidimensional activation
Cold–Hot	0.156627	0.084085	395.957094	82.902765	Low irregularity, low pitch, low variability	Intermittent and relatively weak vocal effort with limited activation; low-energy cry
Tired	0.151715	0.102363	413.109759	89.635140	Low energy, low-to-moderate irregularity, low-to-moderate pitch, low variability	Reduced respiratory support and diminished arousal; fatigue-dominated cry
Belly Pain	0.144520	0.096835	475.233171	84.366574	Low-to-moderate energy, moderate irregularity, moderate-to-high pitch, low variability	Sustained laryngeal tension with limited modulation; pain-related cry dominated by controlled pitch tension

5 Key Insight

The findings indicate that infant cry classification depends on feature interaction, leakage-free evaluation, and model suitability rather than model selection alone for reliable biomedical monitoring.

6 Conclusion

This paper has looked at infant cry classification in terms of methodological reliability as opposed to raw benchmarks performance. The findings reveal that the effect of acoustic representation, imbalance among classes, and testing procedures are leading to highly influential classification results. In a leakage free environment, SVM performed best in terms of balanced performance relative to the tested models, whereas, Random Forest performed equally well in terms of accuracy but not in terms of reliability by class. Limited data were more sensitive to neural performance than to Gradient Boosting, and the balanced behavior of leaky evaluation was the poorest. Increased performance of over 90 percent, in leaky conditions, was not a measure

of greater generalization but less experiment independence. This shows that the effectiveness of evaluation design can overshadow the seemingly successful models. As an applied contribution, the results endorse lightweight, interpretable acoustic model usage as a part of smart neonatal monitoring systems. Nevertheless, they too indicate that cry-only inference is constrained in extreme imbalance and acoustic overlap. It should then be a focus of future development to use a more diverse data, more subject-independent evaluation, and multimodal integration with physiological sensing in order to implement it in intelligent neonatal care settings.

7 Limitations and Future Work

7.1 Limitations

The interpretation of the current findings has a number of limitations. To begin with, the dataset is extremely skewed, and augmentation is not able to replace authentic acoustic diversity in minority classes. Second, the feature representation can be considered statistical

and potentially underrepresents longer temporal dependencies in the sequence of cry. Third, the experiments were done on rather controlled recordings but in the actual incubator settings there is a lot of acoustic interference. Fourth, it could not ensure strict subject-level independence in spite of leakage being managed at the file level. Lastly, the research was narrowed down to the classification of cry only, and is not yet multimodal in terms of physiological context. No formal statistical significance testing (such as McNemar's test or bootstrap confidence intervals) was conducted to compare the classifiers evaluated under the leakage-free evaluation protocol, although the SVM classifier yielded the highest Macro-F1 score and balanced accuracy. So, the results of the observed performance differences should be taken as an empirical difference and not statistically significant. Future work will consider the use of formal statistical hypothesis testing for providing more evidence in favor of model comparison.

7.2 Future Work

In the future all the research must emphasize data-centric enhancement of a system rather than architectural magnification. Balanced and more diverse data sets should be sought especially on underrepresented distresses. Where possible, subject-independent partitioning must be imposed. Statistical frameworks with time-frequency representations can be more effective at capturing global and dynamic structure in a hybrid way. Fusion with physiological measurements, including temperature, SpO₂ and heart rate, is of particular interest to Multimodal fusion as it allows to reduce the similarity of acoustically similar conditions when deploying the smart incubator. Moreover, lighter edge-conscious executions are also to be considered in order to enable real-time inference of embedded neonatal applications.

Disclosures

All authors declare that there is no conflict of interest in this paper.

References

1. S. Sirithemmontree, N. Katchamat, and S. Nuampa, "The application of machine learning for infant cries classification and pathological cries detection: A systematic review," *Science Progress* 109(1), 00368504251410776 (2026).
2. T. Ozcan, H. Gungor, "Baby Cry Classification Using Structure-Tuned Artificial Neural Networks with Data Augmentation and MFCC Features," *Applied Sciences* 15(5), 2648 (2025).
3. G. Aggarwal, K. Jhajharia, J. Izhar, M. Kumar, and L. Abualigah, "A Machine Learning Approach to Classify Biomedical Acoustic Features for Baby Cries," *Journal of Voice* 39(6), 1446–1455 (2025).
4. M. Mekhfioui, W. Fadel, F. E. Hammouch, O. Laayati, M. Bouchouirbat, N. El Bazi, A. Satif, T. Boujiha, and A. Chebak, "Development of a Baby Cry Identification System Using a Raspberry Pi-Based Embedded System and Machine Learning," *Technologies* 13(4), 130 (2025).
5. M. Hammoud, M. N. Getahun, A. Baldycheva, and A. Somov, "Machine learning-based infant crying interpretation," *Frontiers in Artificial Intelligence* 7, 1337356 (2024).
6. K. Rezaee, H. G. Zadeh, L. Qi, H. Rabiee, and M. R. Khosravi, "Can You Understand Why I Am Crying? A Decision-making System for Classifying Infants' Cry Languages Based on DeepSVM Model," *ACM Transactions on Asian and Low-Resource Language Information Processing* 23(1), 1–17 (2024).
7. X. Qiao, S. Jiao, H. Li, G. Liu, X. Gao, and Z. Li, "Infant cry classification using an efficient graph structure and attention-based model," *Kuwait Journal of Science* 51(3), 100221 (2024).
8. F. S. Matikolaie, C. Tadj, "Machine Learning-Based Cry Diagnostic System for Identifying Septic Newborns," *Journal of Voice* 38(4), 963.e1-963.e14 (2024).
9. F. Li, C. Cui, and Y. Hu, "Classification of Infant Crying Sounds Using SE-ResNet-Transformer," *Sensors* 24(20), 6575 (2024).
10. K. Zhang, H.-N. Ting, and Y.-M. Choo, "Baby cry recognition based on WOA-VMD and an improved Dempster-Shafer evidence theory," *Computer Methods and Programs in Biomedicine* 245, 108043 (2024).
11. H. A. Patil, A. Kachhi, and A. T. Patil, "CQT-Based Cepstral Features for Classification of Normal vs. Pathological Infant Cry," *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 32, 4713–4726 (2024).
12. Y. Zayed, A. Hasasneh, and C. Tadj, "Infant Cry Signal Diagnostic System Using Deep Learning and Fused Features," *Diagnostics* 13(12), 2107 (2023).
13. T. Ozseven, "Infant cry classification by using different deep neural network models and hand-crafted features," *Biomedical Signal Processing and Control* 83, 104648 (2023).
14. Z. Wang, Y. Cai, X. Wang, S. Wu, Y. Cao, F. Xu, and M. Huang, "The significance of an infant's cry: a narrative review of physiological, pathological, and analytical perspectives," *Frontiers in Pediatrics* 13, 1558951 (2025).
15. M. Ahmed, "Baby Cry Dataset," Kaggle, 2023 (accessed 11 November 2025). <https://www.kaggle.com/datasets/mennaahmed23/baby-cry-dataset>.
16. A. Abbaskhah, H. Sedighi, and H. Marvi, "Infant cry classification by MFCC feature extraction with MLP and CNN structures," *Biomedical Signal Processing and Control* 86, 105261 (2023).

17. M. Ashraf, F. Abid, I. U. Din, J. Rasheed, M. Yesiltepe, S. F. Yeo, and M. T. Ersoy, “[A Hybrid CNN and RNN Variant Model for Music Classification](#),” *Applied Sciences* 13(3), 1476 (2023).
18. Y.-C. Liang, I. Wijaya, M.-T. Yang, J. R. Cuevas Juarez, and H.-T. Chang, “[Deep Learning for Infant Cry Recognition](#),” *International Journal of Environmental Research and Public Health* 19(10), 6311 (2022).
19. Z. K. Farej and H. Y. Al-hayaly, “[Accuracy Evaluation of Healthcare Monitoring System Based on ESP32 Microcontroller with IoT](#),” in 2023 International Conference on Engineering, Science and Advanced Technology (ICESAT) (IEEE, 2023), pp. 90–94.
20. Z. K. Farej, M. A. Hassan, “[Analysis and Performance Evaluation of the Static and Dynamic Channel Bonding for the IEEE 802.11ac WLAN](#),” *International Journal of Interactive Mobile Technologies* 17(13), 168–185 (2023).

Development of a Method for the Intelligent Analysis of Raman Spectra of Blood Serum

Irina A. Pimenova* and Irina A. Matveeva

Department of Laser and Biotechnical Systems, Samara National Research University, 34 Moskovskoe shosse, Samara 443086, Russian Federation

*e-mail: pimenova.0312@list.ru

Abstract. This paper presents a method for intelligent analysis of Raman spectra of blood serum for chronic heart failure (CHF) diagnosis using machine learning approaches. Raman spectra obtained from blood serum samples of 229 subjects (189 patients with CHF and 40 healthy volunteers) were analyzed using the multivariate curve resolution-alternating least squares (MCR-ALS) method. Eight spectral components were extracted and used as features for classification. Logistic regression, support vector machine, random forest, gradient boosting, and ensemble stacking methods were evaluated. Using only spectral features, the stacking model achieved an area under the receiver operating characteristic curve (ROC AUC) of 0.72 ± 0.04 with a validation accuracy of 0.83 ± 0.03 . The inclusion of clinical parameters (systolic and diastolic blood pressure, final systolic size, total bilirubin, and hematocrit) significantly improved classification performance, increasing the ROC AUC to 1.00 ± 0.00 with a validation accuracy of 0.95 ± 0.01 . Statistical analysis showed no significant differences between the stacking ensemble and individual classifiers ($p > 0.05$), indicating that classification performance was primarily determined by the informativeness of the extracted spectral and clinical features. The obtained results demonstrate the potential of combining Raman spectroscopy with machine learning methods for CHF screening and early diagnosis, particularly in settings where access to advanced imaging techniques is limited.

Keywords: alternating least squares; multidimensional curve resolution method; Raman spectroscopy; blood serum; chronic heart failure; machine learning; stacking.

Paper #9382 received 20 Oct 2025; revised manuscript received 10 Sep 2026; accepted for publication 11 Sep 2026; published online 30 Sep 2026. doi: [10.18287/JBPE26.12.030306](https://doi.org/10.18287/JBPE26.12.030306).

1 Introduction

Cardiovascular diseases occupy a leading place among the causes of death and disability worldwide, which underlines the importance of their diagnosis for healthcare systems [1]. Chronic heart failure (CHF) is one of the most widespread cardiovascular syndromes associated with progressive impairment of cardiac function, reduced quality of life, and increased mortality risk [2]. Early diagnosis of CHF remains a challenging task, especially at the initial stages of disease development, when structural and functional changes may not yet be sufficiently pronounced. Conventional

diagnostic approaches, including echocardiography and biochemical blood analysis, often require expensive equipment, specialized personnel, and considerable examination time, which limits their widespread use in screening programs and preventive medical examinations [3].

Biochemical blood analysis methods are used to quantify the concentration of blood components, enzymes, and hormones, as well as to identify deviations from normal values [4]. However, the prognostic significance of blood analysis can be increased when considering a complex of changes in the biochemical composition of blood rather than isolated biomarkers [5].

In this regard, optical spectroscopic methods capable of providing information about the molecular composition of biological samples attract significant interest [6].

Raman spectroscopy makes it possible to analyze the features of biological objects at the molecular level [7]. In recent years, Raman spectroscopy has opened new perspectives for blood analysis [8]. This method allows obtaining spectral information about the molecular composition of complex biological samples, including blood serum, making it a powerful tool for detecting biochemical changes associated with cardiovascular diseases [9]. Due to its high sensitivity, noninvasiveness, and minimal sample preparation requirements, Raman spectroscopy is becoming an attractive method for clinical diagnostics and biomedical screening applications.

However, Raman spectra of biological samples are characterized by high dimensionality, overlapping spectral bands, noise sensitivity, and substantial inter-patient variability [10]. These factors complicate the interpretation of spectral data and require the use of advanced computational approaches for efficient analysis. In this context, machine learning and intelligent decision-support systems have become increasingly important tools for Raman spectra processing [11].

Various machine learning algorithms have been applied to the analysis of Raman spectra in biomedical research. Classical statistical and machine learning methods, including Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Logistic Regression, Support Vector Machines (SVM) have demonstrated effectiveness in classification tasks involving biological samples [12–14]. The development of information technology has also led to attempts to use artificial neural networks and deep learning methods for spectral data analysis [15]. In particular, convolutional neural networks have shown promising results in automated Raman spectrum recognition and feature extraction [16]. These approaches enable identification of complex nonlinear relationships within spectral data and improve diagnostic accuracy for various pathological conditions [17].

Despite significant progress, the problem of extracting informative and interpretable features from Raman spectra remains unresolved. In many studies, dimensionality reduction is performed using PCA or related statistical methods [18–19]. Although such approaches effectively reduce computational complexity, the resulting components are often difficult to interpret from a biochemical perspective [20]. Therefore, methods capable of extracting physically interpretable spectral components are of particular interest for biomedical applications [21].

One promising approach is multivariate curve resolution using alternating least squares (MCR-ALS), which allows decomposition of spectral data into concentration profiles and spectra of underlying biochemical components [22]. Unlike conventional statistical decomposition techniques, MCR-ALS provides improved interpretability of extracted features

and may facilitate identification of molecular changes associated with disease development [23].

Another important trend is the development of ensemble machine learning methods that combine predictions of several individual classifiers [24]. Ensemble approaches can improve the robustness and generalization ability of classification models, especially when analyzing complex biomedical datasets. In particular, stacking methods combine the advantages of multiple algorithms by constructing a meta-model based on predictions of basic classifiers [25].

The novelty of the proposed approach lies in the combined use of MCR-ALS decomposition, ensemble machine learning methods, and integration of Raman spectral data with clinical patient parameters for chronic heart failure diagnosis. Unlike conventional approaches based solely on spectral preprocessing and standard dimensionality reduction techniques, the proposed framework extracts interpretable biochemical components from Raman spectra and combines them with clinically significant features. In addition, several machine learning algorithms, including logistic regression, support vector machines, random forest, gradient boosting, and stacking ensembles, are comparatively evaluated for CHF classification.

The use of Raman spectroscopy in combination with intelligent data analysis systems is a promising area of research and can increase the effectiveness of population screening and preventive medical examinations. Such approaches can improve the accessibility of diagnostics, contribute to the early detection of cardiovascular diseases, and potentially prevent disease progression [26].

The aim of this study was to develop an intelligent method for the analysis of blood serum Raman spectra using the MCR-ALS method and machine learning algorithms, taking into account additional clinical data of patients with chronic heart failure.

2 Methods and Materials

In this work, the *in vitro* Raman spectra of the blood serum of 229 patients are analyzed, among which 189 are patients with chronic heart failure, 40 are control group (volunteers without heart failure). Blood serum samples were taken from the patient in the morning on an empty stomach and placed in sterile tubes, followed by freezing at a temperature of -16°C . Immediately before spectral analysis, the samples were thawed at room temperature.

The recording of Raman spectra was carried out using surface-enhanced Raman spectroscopy (SERS) technology, described in detail in Ref. [27]. Each 1.5 ml serum sample was applied to an aluminum foil with a layer of silver structures and dried for 30 min. The spectral characteristics of blood serum were analyzed using an experimental stand consisting of a spectrometric system (EnSpectr R785, Spektr-M, Chernogolovka, Russia) and a microscope (ADF U300, ADF, China). The spectra were excited in the near-infrared range using a laser module with a central wavelength of 785 nm. An LMPlan lens with a $50\times$ magnification was used to focus the radiation on the

sample and collect the scattered radiation. The diameter of the laser spot in focus was 5 μm . Human serum was analyzed at a laser power of 10 mW. The spectra were recorded four times for each sample with an exposure time of 4 s. Further, the spectra were subjected to the preprocessing described in Ref. [28].

In this work, the method of resolving multidimensional curves using the alternating least squares (MCR-ALS) method is used to analyze the spectra, which has already shown its effectiveness in isolating the biochemical components of the skin [29]. The main idea of MCR is to decompose the matrix of Raman spectra into two matrices, one of which represents the concentration profiles for each component, and the second represents the matrix of the spectra of these components [30].

In the framework of the MCR-ALS analysis, the term “components” refers to latent spectral profiles extracted from the original Raman spectra during matrix decomposition [31].

In this approach, the concentrations of each individual component isolated in the studied samples serve as features for subsequent classification. The following machine learning algorithms were applied to the profiles of the obtained component concentrations: Logistic Regression, SVM, Random Forest, and Gradient Boosting.

Logistic Regression is a form of multiple regression that aims to analyze the relationship between several independent variables and a single binary dependent variable [32–34]. Random forest is based on the construction of an ensemble of decision trees and is one of the most robust machine learning algorithms [35–37]. The principle of the support vector machine method is to find the optimal hyperplane that maximally separates classes in a multidimensional feature space [38–40]. Gradient boosting is a powerful ensemble learning method in which the sequential construction of decision trees aims to minimize the errors of previous models [41–43].

To improve the quality of the final model and combine the advantages of various methods, the stacking method was applied, in which the basic models formed meta-features, which were then fed to the input of the final model, the logistic regression model. This approach makes it possible to exploit the strengths of individual algorithms, improving classification generalization performance.

In addition to concentration profiles, the clinical data of patients were used as additional features for classification: systolic blood pressure, diastolic blood pressure, final systolic size, total bilirubin, hematocrit.

A 5-fold cross-validation was performed to verify the stability of the classification models. All machine learning algorithms and ROC curve analyses were implemented in Python.

3 Results and Discussions

Using MCR-ALS, eight components have been identified. This amount was determined during

preliminary studies, which examined the range from 3 to 12 components in order to optimally identify substances. Several datasets with different numbers of components were analyzed using various machine learning methods. The ROC curves obtained for logistic regression based on five data sets are shown in Fig. 1. Based on the analysis, it is concluded that the choice of eight components is optimal for achieving the objectives of the study. The primary objective was to retain as much information as possible, which would improve the accuracy of the analysis of the Raman spectra of blood serum without losing key data.

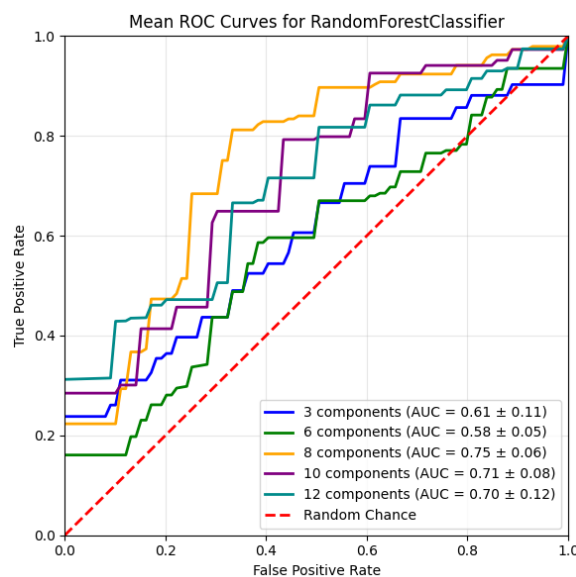


Fig. 1 ROC-curves for logistic regression for different numbers of components.

The main peaks for each component are shown in Table 1. The spectra of the components are shown in Fig. 2. For example, Component 1 was mainly associated with proteins and lipids due to the presence of peaks near 723 cm^{-1} , 1097 cm^{-1} , and 1329 cm^{-1} , corresponding to amide groups, C–O vibrations, and hydrocarbon chain deformations. Component 2 was related predominantly to protein structures and aromatic amino acids because of the characteristic Amide I band near 1652 cm^{-1} and aromatic ring vibrations at 1576 cm^{-1} . Component 3 was associated with metabolites and carbohydrate-related vibrations, while Components 4 and 5 reflected contributions from aromatic compounds, proteins, and lipid-associated molecular structures. Components 6–8 mainly corresponded to carbohydrates, amino acid residues, and low-frequency molecular vibrations observed in biological serum samples [44–45].

Therefore, the obtained concentration profiles reflect variations in the biochemical composition of blood serum associated with chronic heart failure.

As a result, the area under the ROC curve for the case of taking into account only spectral features (component concentrations) and using the basic logistic regression model was 0.72 ± 0.04 (validation accuracy: 0.83 ± 0.03), the support vector machine achieved an

ROC AUC of 0.66 ± 0.08 (validation accuracy: 0.83 ± 0.03), random forest 0.70 ± 0.07 (validation accuracy: 0.85 ± 0.01), gradient boosting 0.62 ± 0.05 (validation accuracy: 0.81 ± 0.02). For the stacking model, the area under the ROC curve was 0.72 ± 0.04 (validation accuracy: 0.83 ± 0.03). The addition of clinical patient data to the feature space has significantly improved the learning performance of the model. The ROC AUC increased to 1.00 ± 0.00 , while the validation accuracy was 0.95 ± 0.01 .

Table 1 Characteristic peaks and assumed molecular bonds of the components.

Component	Peak (cm ⁻¹)	Possible Assignment
Component 1	723.12	Amide group or methylene groups in lipids
	1329.03	C–H deformation in hydrocarbon chains (possibly Amide III)
	1096.99	C–O or S–O stretching in carbohydrates or carboxylic acids
Component 2	1652.38	C=O stretching in amide groups or carbohydrates (Amide I)
	1576.09	Aromatic ring stretching (e.g., phenyl groups)
	639.33	Amino acids associated with aromatic ring vibrations
Component 3	1393.58	Vibrations typical for creatinine or other metabolites
	1021.23	Carbohydrate or amino acid vibrations
	631.35	Amino acid residues (e.g., phenylalanine)
Component 4	1577.62	Aromatic vibrations related to phenyl groups
	1212.39	Vibrations in carbohydrates or amide groups
	1313.32	C–H deformation or aromatic ring vibrations
Component 5	624.35	Amino acids or carbohydrates
	1364.32	C–H or C–C deformations in hydrocarbon chains
	1604.25	Aromatic ring vibrations
Component 6	1053.51	C–O stretching in carbohydrates or metabolites
	677.01	Aromatic ring vibrations or possible methylene groups
	492.15	Low-frequency vibrations in organic molecules
Component 7	645.30	Aromatic rings or carbohydrates
	500.41	Methylene groups or low-frequency vibrations
	1142.56	C–O or methyl group stretching
Component 8	634.34	Amino acids or carbohydrates
	491.12	Vibrations in carbohydrates or amino acids
	1132.10	Presence of amide groups (Amide III region)

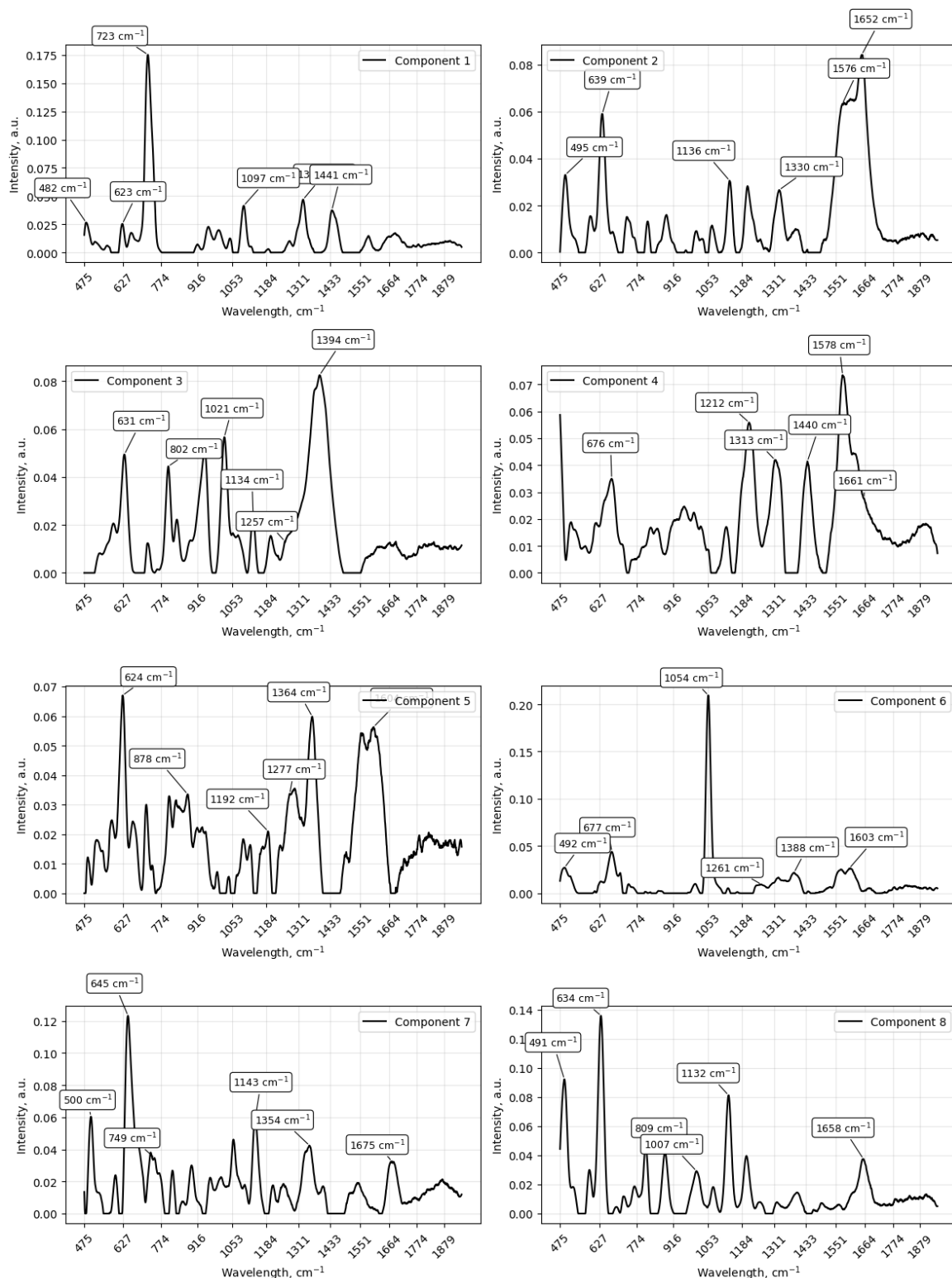


Fig. 2 Raman spectra obtained by *in vitro* MCR analysis of Raman scattering spectra of blood serum.

In the framework of this work, an ensemble classification model was implemented using the stacking method. The aim was to combine several basic models to improve the accuracy of the prediction of CHF based on spectral and clinical features. The ROC curves obtained as a result of the analysis are shown in Fig. 3.

To further validate the obtained results, a statistical significance analysis was conducted comparing the

baseline models with the stacked ensemble. The comparison was performed using the area under the ROC curve (AUC) obtained from 5-fold cross-validation. Since the performance values were calculated on the same data splits, paired statistical tests were applied, including the paired Student's *t*-test [46] and the nonparametric Wilcoxon signed-rank test [47].

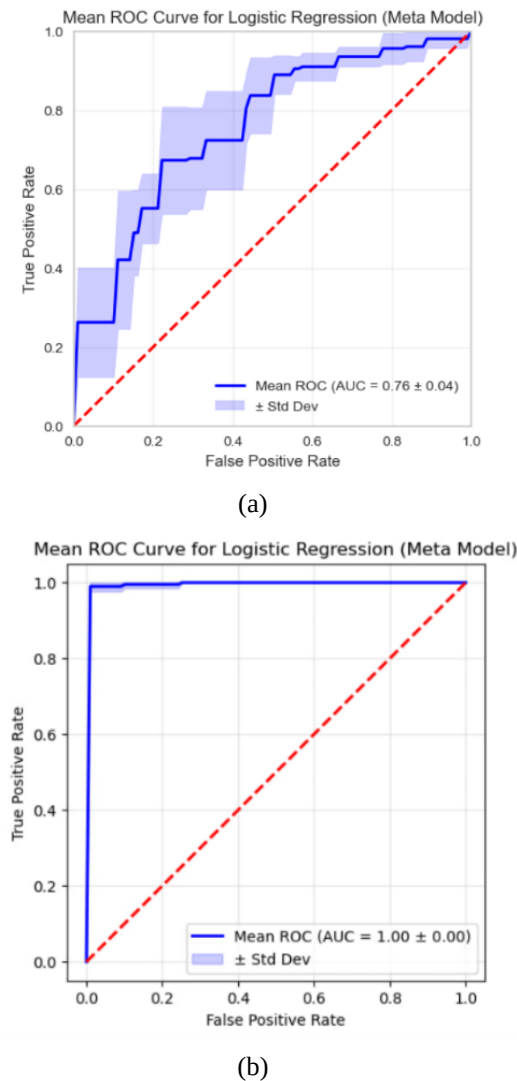


Fig. 3 ROC-curves for classification of healthy people and patients with chronic heart failure: (a) only component concentrations are taken as features; (b) component concentrations and clinical data are taken as features.

The results show that the differences between the stacked model and individual baseline classifiers are statistically insignificant ($p > 0.05$ in all cases). Specifically, the p -values for logistic regression were 0.3408 (t -test) and 0.3125 (Wilcoxon test), 0.2463 and 0.3125 for support vector machines, 0.8445 and 0.8125 for random forest, and 0.0722 and 0.1250 for gradient boosting, respectively. Although gradient boosting approached statistical significance ($p \approx 0.07$), it does not reach the generally accepted significance level of 0.05.

These results are consistent with the similar performance metrics reported above. The lack of statistically significant differences can be explained by several factors. First, the relatively small sample size (229 samples, of which only 40 were healthy controls) leads to increased variability in AUC estimates across different cross-validation folds, reducing the statistical power of the tests. Second, the baseline models demonstrate comparable performance levels and likely

capture similar patterns in the feature space formed by the eight MCR-ALS components. Under these conditions, the advantage of stacking is diminished because the metamodel does not receive sufficiently diverse or additional information.

Thus, the statistical analysis confirms that the stacking approach does not provide a statistically significant improvement over the individual classifiers in this study. This suggests that, within the considered framework, the overall classification performance is primarily determined by the informativeness of the extracted features rather than by the choice of a specific classification algorithm.

We conducted an additional experiment using only five clinical parameters (systolic and diastolic blood pressure, final systolic size, total bilirubin, and hematocrit). The obtained results demonstrated near-perfect classification performance (Fig. 4; Mean Accuracy: 0.96 ± 0.03 ; Mean AUC: 0.98 ± 0.03).

Such high performance can be explained by the strong association between these clinical variables and the diagnosis of chronic heart failure.

Among the clinical parameters used in this study, final systolic size represents a key echocardiographic indicator associated with structural cardiac remodeling and is directly related to the diagnosis of chronic heart failure. Other variables, such as blood pressure, bilirubin, and hematocrit, are not diagnostic criteria per se but reflect systemic and metabolic alterations accompanying the disease.

The combination of these features provides a strong discriminative signal, which may explain the high classification performance observed in this dataset.

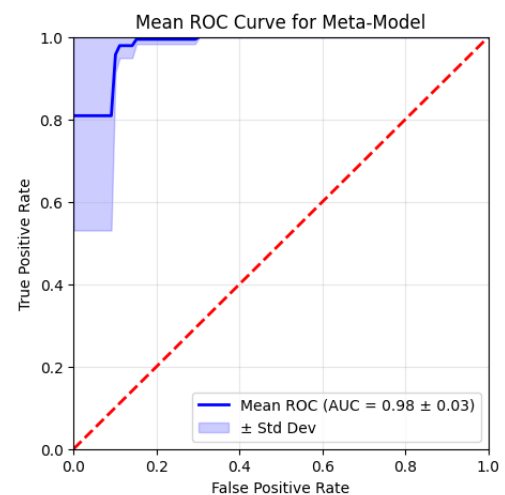


Fig. 4 ROC-curves for classification of healthy people and patients with chronic heart failure; clinical data are taken as features.

However, this result is likely to reflect dataset-specific characteristics and may lead to an overestimation of model performance, especially in the absence of external validation. In real-world clinical settings, variability of patient populations and overlapping clinical features typically result in lower separability. Despite the high

discriminative power of clinical parameters, it should be emphasized that these features are typically obtained using established diagnostic procedures, such as echocardiography, which may not always be available in all clinical settings.

In this context, Raman spectroscopy may provide important practical advantages. In particular, it can be considered as a potential tool for screening in settings where access to advanced imaging techniques is limited, or as a preliminary diagnostic approach prior to comprehensive clinical evaluation.

Furthermore, Raman spectroscopy reflects molecular-level changes in blood serum and therefore may be sensitive to early metabolic alterations that precede pronounced structural cardiac remodeling. This suggests its potential application for early-stage detection, where conventional clinical parameters may not yet exhibit strong diagnostic signals.

Thus, the value of Raman spectroscopy lies not in replacing established clinical diagnostics, but in complementing them, particularly in screening scenarios and early disease detection.

Machine learning methods have been applied to the analysis of biological fluids for the diagnosis and stratification of CHF in a number of recent studies. Thus, Sorokina et al. [48] conducted a systematic comparison of methods for intelligent analysis of Raman spectra of blood serum from patients with CHF and healthy volunteers using combinations of MCR and PLS with different classifiers (Logistic Regression, SVM). They showed that PLS-based approaches provide higher AUC values (up to 0.99), while MCR in combination with single classifiers demonstrated moderate performance (AUC 0.59–0.70), while leaving the possibility of biochemical interpretation of serum components. Other studies, such as De Bakker et al. [49] and Gao et al. [50] used proteomic and biomarker panels, applying ML algorithms to predict outcomes and classify patients with CHF, achieving a c index of 0.80–0.85 or an accuracy of approximately 85%. The approach by Woolley et al. [51] demonstrated the potential of clustering biomarker profiles to distinguish pathophysiological subtypes of HFpEF, but without direct application to automated diagnostics.

In contrast to these studies, this work uses SERS to directly assess the molecular composition of blood serum. The use of the MCR-ALS method allowed us to identify eight informative components with biochemical interpretation (amino acids, carbohydrates, lipids, amide groups). The resulting concentration profiles, combined with patient clinical data, were used to build ensemble machine learning models, including stacking base classifiers (Logistic Regression, SVM, Random Forest, and Gradient Boosting). This approach achieved an AUC of 1.0 ± 0.0 , which exceeds the MCR results reported by Sorokina et al. [48] and is comparable to the best results

from PLS methods, while maintaining the interpretability of features.

4 Conclusion

In this study, a method for intelligent analysis of Raman spectra of blood serum for chronic heart failure diagnosis was developed using MCR-ALS decomposition and machine learning approaches. Raman spectra obtained from 229 serum samples were decomposed into eight latent spectral components corresponding to major biochemical groups, including proteins, lipids, carbohydrates, amino acids, and metabolites. The concentration profiles of these components were used as features for classification.

Several machine learning algorithms, including logistic regression, support vector machine, random forest, gradient boosting, and stacking ensemble methods, were evaluated. Using only spectral features, the best classification performance corresponded to the stacking model and logistic regression model with an ROC AUC of 0.72 ± 0.04 and validation accuracy of 0.83 ± 0.03 . The integration of clinical parameters significantly improved classification results, increasing the ROC AUC to 1.00 ± 0.00 .

Statistical analysis demonstrated that differences between the stacking ensemble and individual classifiers were not statistically significant ($p > 0.05$), indicating that classification performance was primarily determined by the informativeness of the extracted spectral and clinical features rather than by the choice of a particular classification algorithm.

The obtained results demonstrate that Raman spectroscopy combined with machine learning methods can be considered a promising tool for blood serum analysis and chronic heart failure screening. In particular, the proposed approach may be useful in situations where rapid, minimally invasive, and cost-effective preliminary diagnostics are required. Furthermore, Raman spectroscopy may provide additional information about molecular-level biochemical changes associated with disease development.

Future work will focus on expanding the dataset, performing external validation on independent patient cohorts, and investigating additional biochemical and clinical markers to improve the robustness and clinical applicability of the proposed diagnostic approach.

Acknowledgment

This work was supported by the Russian Science Foundation project № 25-75-00146. <https://rscf.ru/project/25-75-00146/>.

Disclosures

All authors declare that there is no conflict of interest in this paper.

References

1. “Cardiovascular diseases (CVDs),” World Health Organization (WHO), 31 July 2025 (accessed 29 December 2025). [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-%28cvds%29?utm_source].

2. V. Valente, R. Laborante, A. Abdin, P. M. Becher, M. Lainscak, M. Polovina, C. Gavina, and G. Savarese, "[The global epidemiology of heart failure: a comprehensive and contemporary review](#)," *European Journal of Heart Failure* xuag121 (2026).
3. P. A. Heidenreich, B. Bozkurt, D. Aguilar, L. A. Allen, J. J. Byun, M. M. Colvin, A. Deswal, M. H. Drazner, S. M. Dunlay, L. R. Evers, J. C. Fang, S. E. Fedson, G. C. Fonarow, S. S. Hayek, A. F. Hernandez, P. Khazanie, M. M. Kittleson, C. S. Lee, M. S. Link, C. A. Milano, L. C. Nnacheta, A. T. Sandhu, L. W. Stevenson, O. Vardeny, A. R. Vest, and C. W. Yancy, "[2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the american college of cardiology/american heart association joint committee on clinical practice guidelines](#)," *Circulation* 145(18), (2022).
4. D. Glick (Ed.), *Methods of Biochemical Analysis*, Hoboken, John Wiley & Sons, NJ, USA (2009).
5. C. G. Atkins, K. Buckley, M. W. Blades, and R. F. Turner, "[Raman spectroscopy of blood and blood components](#)," *Applied Spectroscopy* 71(5), 767–793 (2017).
6. H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N. J. Fullwood, B. Gardner, P. L. Martin-Hirsch, M. J. Walsh, M. R. McAinsh, N. Stone, and F. L. Martin, "[Using Raman spectroscopy to characterize biological materials](#)," *Nature Protocols* 11(4), 664–687 (2016).
7. N. Kuhar, S. Sil, T. Verma, and S. Umapathy, "[Challenges in application of Raman spectroscopy to biology and materials](#)," *RSC Advances* 8(46), 25888–25908 (2018).
8. J. Dybas, F. C. Alcicek, A. Wajda, M. Kaczmarek, A. Zimna, K. Bulat, A. Blat, T. Stepanenko, T. Mohaissen, E. Szczesny-Malysiak, D. Perez-Guaita, B. R. Wood, and K. M. Marzec, "[Trends in biomedical analysis of red blood cells – Raman spectroscopy against other spectroscopic, microscopic and classical techniques](#)," *TrAC Trends in Analytical Chemistry* 146, 116481 (2022).
9. Y. Khristoforova, L. Bratchenko, and I. Bratchenko, "[Raman-based techniques in medical applications for diagnostic tasks: a review](#)," *International Journal of Molecular Sciences* 24(21), 15605 (2023).
10. J. Servet Lerdo De Tejada, J. Vališ, L. Hrubčík, D. Janstová, J. Tomeš, T. Uhlíková, M. Vazdar, J. Mareš, and P. Matějka, "[Mathematical and machine learning-assisted modelling of Raman spectroscopy for biomedical applications](#)," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 362, 128199 (2026).
11. S. Guo, J. Popp, and T. Bocklitz, "[Chemometric analysis in Raman spectroscopy from experimental design to machine learning-based modeling](#)," *Nature Protocols* 16(12), 5426–5459 (2021).
12. E. D. Omar, H. Mat, A. Z. Abd Karim, R. Sanaudi, F. Ibrahim, M. A. Omar, M. Z. H. Ismail, V. Jayaraj, and B. L. Goh, "[Comparative analysis of logistic regression, gradient boosted trees, SVM, and random forest algorithms for prediction of acute kidney injury requiring dialysis after cardiac surgery](#)," *International Journal of Nephrology and Renovascular Disease* 17, 197–204 (2024).
13. J.-M. Yin, Y. Li, J.-T. Xue, G.-W. Zong, Z.-Z. Fang, and L. Zou, "[Explainable machine learning-based prediction model for diabetic nephropathy](#)," *Journal of Diabetes Research* 2024(1), 8857453 (2024).
14. Md. S. Hossain, Md. A. Talukder, and Md. Z. Mahmud, "[Advancements in cardiovascular disease detection: leveraging data mining and machine learning](#)," *bioRxiv* 09, 584222 (2024).
15. J. Liu, M. Osadchy, L. Ashton, M. Foster, C. J. Solomon, and S. J. Gibson, "[Deep convolutional neural networks for Raman spectrum recognition: a unified solution](#)," *Analyst* 142(21), 4067–4074 (2017).
16. M. Kazemzadeh, C. L. Hisey, K. Zargar-Shoshtari, W. Xu, and N. G. R. Broderick, "[Deep convolutional neural networks as a unified solution for Raman spectroscopy-based classification in biomedical applications](#)," *Optics Communications* 510, 127977 (2022).
17. N. Blake, R. Gaifulina, L. D. Griffin, I. M. Bell, and G. M. H. Thomas, "[Machine Learning of Raman Spectroscopy Data for Classifying Cancers: A Review of the Recent Literature](#)," *Diagnostics* 12(6), 1491 (2022).
18. Z. Zhang, A. Castelló, "[Principal components analysis in clinical studies](#)," *Annals of Translational Medicine* 5(17), 351–351 (2017).
19. L. A. Bratchenko, I. A. Bratchenko, "[Avoiding overestimation and the 'black box' problem in biofluids multivariate analysis by Raman spectroscopy: interpretation and transparency with the SP-LIME algorithm](#)," *Journal of Raman Spectroscopy* 56(4), 353–364 (2025).
20. Y. Liu, Y. Wu, J. Wang, J. Qi, C. Zhou, and Y. Xue, "[Recent Advances in Raman Spectral Classification with Machine Learning](#)," *Sensors* 26(1), 341 (2026).
21. R. S. Senger, D. Scherr, "[Resolving complex phenotypes with Raman spectroscopy and chemometrics](#)," *Current Opinion in Biotechnology* 66, 277–282 (2020).
22. J. Felten, H. Hall, J. Jaumot, R. Tauler, A. De Juan, and A. Gorzsás, "[Vibrational spectroscopic image analysis of biological material using multivariate curve resolution–alternating least squares \(MCR-ALS\)](#)," *Nature Protocols* 10(2), 217–240 (2015).
23. M. Ando, H. Hamaguchi, "[Molecular component distribution imaging of living cells by multivariate curve resolution analysis of space-resolved Raman spectra](#)," *Journal of Biomedical Optics* 19(1), 011016 (2013).
24. O. Sagi, L. Rokach, "[Ensemble learning: A survey](#)," *WIREs Data Mining and Knowledge Discovery* 8(4), e1249 (2018).

25. M. A. Pramudito, Y. N. Fuadah, Y. S. Kim, and K. M. Lim, “[Stacking Ensemble Machine Learning for Cardiac Safety Assessment Using hiPSC-CM MEA Data](#),” *Annals of Biomedical Engineering* 54(5), 1334–1344 (2026).
26. A. Chaichi, A. Prasad, and M. R. Gartia, “[Raman Spectroscopy and Microscopy Applications in Cardiovascular Diseases: From Molecules to Organs](#),” *Biosensors* 8(4), 107 (2018).
27. S. Z. Al-Sammarraie, L. A. Bratchenko, E. N. Typikova, V. P. Zakharov, I. A. Bratchenko, and P. A. Lebedev, “[Silver nanoparticles-based substrate for blood serum analysis under 785 nm laser excitation](#),” *Journal of Biomedical Photonics & Engineering* 8(1), 010301 (2022).
28. Y. A. Khristoforova, L. A. Bratchenko, M. A. Skuratova, E. A. Lebedeva, P. A. Lebedev, and I. A. Bratchenko, “[Raman spectroscopy in chronic heart failure diagnosis based on human skin analysis](#),” *Journal of Biophotonics* 16(7), e202300016 (2023).
29. I. A. Matveeva, A. I. Komlev, O. I. Kaganov, A. A. Moryatov, V. P. Zakharov, “[Multidimensional analysis of dermoscopic images and spectral information for the diagnosis of skin tumors](#),” *Journal of Biomedical Photonics & Engineering* 10(1), (2024).
30. J. Jaumot, A. De Juan, and R. Tauler, “[MCR-ALS GUI 2.0: New features and applications](#),” *Chemometrics and Intelligent Laboratory Systems* 140, 1–12 (2015).
31. H. Noothalapati, K. Iwasaki, and T. Yamamoto, “[Biological and Medical Applications of Multivariate Curve Resolution Assisted Raman Spectroscopy](#),” *Analytical Sciences* 33(1), 15–22 (2017).
32. T. G. Nick, K. M. Campbell, “[Logistic Regression](#),” *Topics in Biostatistics* 404, 273–301 (2007).
33. S. Domínguez-Almendros, N. Benítez-Parejo, and A. R. Gonzalez-Ramirez, “[Logistic regression models](#),” *Allergologia et Immunopathologia* 39(5), 295–305 (2011).
34. E. Y. Boateng, D. A. Abaye, “[A review of the logistic regression model with emphasis on medical research](#),” *Journal of Data Analysis and Information Processing* 07(04), 190–207 (2019).
35. L. Breiman, “[Random forests](#),” *Machine Learning* 45(1), 5–32 (2001).
36. B. Talekar, “[A detailed review on decision tree and random forest](#),” *Bioscience Biotechnology Research Communications* 13(14), 245–248 (2020).
37. H. A. Salman, A. Kalakech, and A. Steiti, “[Random forest algorithm overview](#),” *Babylonian Journal of Machine Learning* 2024, 69–79 (2024).
38. C. Cortes, V. Vapnik, “[Support-vector networks](#),” *Machine Learning* 20(3), 273–297 (1995).
39. J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, “[A comprehensive survey on support vector machine classification: Applications, challenges and trends](#),” *Neurocomputing* 408, 189–215 (2020).
40. S. Winters-Hilt, A. Yelundur, C. McChesney, and M. Landry, “[Support vector machine implementations for classification & clustering](#),” *BMC Bioinformatics* 7(S2), S4 (2006).
41. J. H. Friedman, “[Stochastic gradient boosting](#),” *Computational Statistics & Data Analysis* 38(4), 367–378 (2002).
42. A. Natekin, A. Knoll, “[Gradient boosting machines, a tutorial](#),” *Frontiers in Neurorobotics* 7, (2013).
43. H. Belyadi, A. Haghighat, “[Supervised learning](#),” *Machine Learning Guide for Oil and Gas Using Python* (Elsevier) 169–295 (2021).
44. X. Yue, H. Li, J. Tang, J. Liu, and J. Jiao, “[Rapid and label-free screening of echinococcosis serum profiles through surface-enhanced Raman spectroscopy](#),” *Analytical and Bioanalytical Chemistry* 412(2), 279–288 (2020).
45. J. Wang, D. Lin, J. Lin, Y. Yu, Z. Huang, Y. Chen, J. Lin, S. Feng, B. Li, N. Liu, and R. Chen, “[Label-free detection of serum proteins using surface-enhanced Raman spectroscopy for colorectal cancer screening](#),” *Journal of Biomedical Optics* 19(8), 087003 (2014).
46. T. K. Kim, “[T test as a parametric statistic](#),” *Korean Journal of Anesthesiology* 68(6), 540 (2015).
47. F. Wilcoxon, “[Individual comparisons by ranking methods](#),” *Biometrics Bulletin* 1(6), 80 (1945).
48. E. Sorokina, Y. Khristoforova, I. Bratchenko, M. Skuratova, P. Lebedev, and V. Zakharov, “[Comparison of Machine Learning Methods for the Analysis of Serum Raman Spectra in the Detection of Chronic Heart Failure](#),” *Journal of Biomedical Photonics & Engineering* 10(4), 040319 (2024).
49. Y. Gao, X. Bai, J. Lu, L. Zhang, X. Yan, X. Huang, H. Dai, Y. Wang, L. Hou, S. Wang, A. Tian, and J. Li, “[Prognostic value of multiple circulating biomarkers for 2-year death in acute heart failure with preserved ejection fraction](#),” *Frontiers in Cardiovascular Medicine* 8, 779282 (2021).
50. M. De Bakker, T. B. Petersen, A. J. Rueten-Budde, K. M. Akkerhuis, V. A. Umans, J. J. Brugts, T. Germans, M. J. T. Reinders, P. D. Katsikis, P. J. Van Der Spek, R. Ostroff, R. She, D. Lanfear, F. W. Asselbergs, E. Boersma, D. Rizopoulos, and I. Kardys, “[Machine learning-based biomarker profile derived from 4210 serially measured proteins predicts clinical outcome of patients with heart failure](#),” *European Heart Journal - Digital Health* 4(6), 444–454 (2023).
51. R. J. Woolley, D. Ceelen, W. Ouwerkerk, J. Tromp, S. M. Figarska, S. D. Anker, K. Dickstein, G. Filippatos, F. Zannad, M. Metra, L. Ng, N. Samani, D. J. Van Veldhuisen, C. Lang, C. S. Lam, and A. A. Voors, “[Machine learning based on biomarker profiles identifies distinct subgroups of heart failure with preserved ejection fraction](#),” *European Journal of Heart Failure* 23(6), 983–991 (2021).