360º optoacoustic capsule endoscopy at 50 Hz for esophageal imaging

Ali, Zakiullah; Zakian, Christian; Li, Qian; Gloriod, Jerome; Crozat, Sophie; Bouvet, François; Pierre, Guillaume; Sarantos, Vassilis; Di Pietro, Massimiliano; Flisikowski, Krzysztof

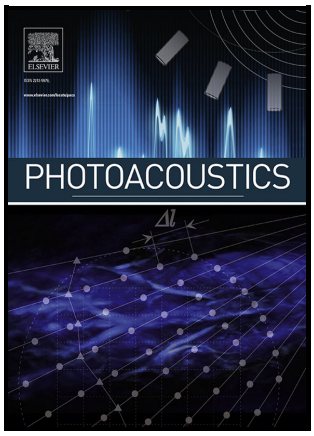
Total number of authors:
13

Published in:
Photoacoustics

Link to article, DOI:
10.1016/j.pacs.2022.100333

Publication date:
2022

Document Version
Version created as part of publication process; publisher's layout; not normally made publicly available



360° optoacoustic capsule endoscopy at 50Hz for esophageal imaging

Zakiullah Ali, Christian Zakian, Qian Li, Jerome Gloriod, Sophie Crozat, François Bouvet, Guillaume Pierre, Vassilis Sarantos, Massimiliano Di Pietro, Krzysztof Flisikowski, Peter Andersen, Wolfgang Drexler, Vasilis Ntziachristos

PII: S2213-5979(22)00005-2

DOI: https://doi.org/10.1016/j.pacs.2022.100333

Reference: PACS100333

To appear in: *Photoacoustics*

Received date: 15 October 2021
Revised date: 10 January 2022
Accepted date: 31 January 2022

Please cite this article as: Zakiullah Ali, Christian Zakian, Qian Li, Jerome Gloriod, Sophie Crozat, François Bouvet, Guillaume Pierre, Vassilis Sarantos, Massimiliano Di Pietro, Krzysztof Flisikowski, Peter Andersen, Wolfgang Drexler and Vasilis Ntziachristos, 360° optoacoustic capsule endoscopy at 50Hz for esophageal imaging, *Photoacoustics*, (2021) doi:https://doi.org/10.1016/j.pacs.2022.100333

360° optoacoustic capsule endoscopy at 50Hz for esophageal imaging

Zakiullah Ali[1,2], Christian Zakian[1,2], Qian Li[3], Jerome Gloriod[4], Sophie Crozat[5], François Bouvet[5], Guillaume Pierre[5], Vassilis Sarantos[6], Massimiliano Di Pietro[7], Krzysztof Flisikowski[8], Peter Andersen[9],Wolfgang Drexler[3], Vasilis Ntziachristos[1,2]

[1]Chair of Biological Imaging, Center for Translational Cancer Research (TranslaTUM), School of Medicine, Technical University of Munich, Munich, Germany
[2]Institute of Biological and Medical Imaging, Helmholtz Zentrum München, Neuherberg, Germany
[3]Center of Medical Physics and Biomedical Engineering, Medical university of Vienna, Vienna, Austria
[4]Statice, Besancon, France
[5]Sonaxis S.A., Besancon, France
[6]RayFos: Scientific Software, Basingstoke, United Kingdom
[7]MRC Cancer Unit, University of Cambridge, Cambridge, United Kingdom
[8]Chair of Livestock Biotechnology, School of Life Science, Technical University of Munich, Freising, Germany
[9]Department of Health Technology, Technical University of Denmark, Lyngby, Denmark

*Corresponding author: bioimaging.translatum@tum.de

Key words

Optoacoustic endoscopy, ultrasound transducer, 360° field of view, fast imaging, distal scanning, slip rings, photoacoustics

Abstract

Gastrointestinal(GI) endoscopy is a common medical diagnostic procedure used for esophageal cancer detection. Current emerging capsule optoacoustic endoscopes, however, suffer from low pulse repetition rates and slow scanning units limit attainable imaging frame rates. Consequently, motion artifacts result in inaccurate spatial mapping and misinterpretation of data. To overcome these limitations, we report a 360°, 50Hz frame rate, distal scanning capsule optoacoustic endoscope. The translational capability of the instrument for human GI tract imaging was characterized with an Archimedean spiral phantom consisting of twelve 100µm sutures, a stainless steel mesh with a pitch of 3mm and an *ex vivo* pig esophagus sample. We estimated an imaging penetration depth of ~0.84mm in vivo by immersing the mesh phantom in intralipid solution to simulate light scattering in human esophageal tissue and validated our findings *ex vivo* using pig esophagus. This proof-of-concept study demonstrates the translational potential of the proposed video-rate endoscope for human GI tract imaging.

Abbreviations

OA	Optoacoustics
RPM	Revolutions per minute
GI	Gastrointestinal
WLSE	White light Surveillance Endoscopy

Introduction

Having caused a total of 544,076 deaths in 2020, esophageal cancer is the sixth deadliest cancer and responsible for 1 in every 19 cancer-related deaths worldwide [1]. The prognosis of patients with advanced esophageal cancer is poor, with survival rates less than 20% after five years. However, early detection of the disease can allow for curative treatment and potential survival rates exceeding 80% [2]. At present, esophageal cancer is detected by white light surveillance endoscopy (WLSE), where the lumen is illuminated with a standard color camera that is designed to create images that replicate human vision , sensitive to only red, green, blue wavelengths (RGB) for accurate colour representation. Unfortunately, changes in contrast arising from low-grade dysplasia or early cancerous lesions are often imperceptible when imaging the esophagus in vivo, even with high definition WLSE. Therefore,

arbitrary quadratic biopsies for histological examination are required to increase the rate of lesion detection [3,4,5]. Recently, a clinically translatable hyperspectral endoscope (HySE) system was developed to improve contrast for in vivo gastrointestinal (GI) tract imaging and features excellent spatial, spectral, and temporal resolution, as well as high color fidelity [6]. HySE applications include quantification of blood oxygenation levels and differentiation of spectral profiles from normal to pathological ex vivo human tissues. However, HySE does not resolve depth information and image feature extraction is still challenged in homogenous or low contrast situations. As an alternative, ultrasound endoscopy is employed to obtain structural depth information within the esophageal wall for cancer staging [7]. Despite ultrasound and WLSE's broad use, the compromises of these technologies include limitations in contrast and depth. WLSE only allows visual inspection of the esophageal surface, ultrasound endoscopy offers no molecular contrast and has low resolution within the upper mucosal layers, and biopsies can often miss early stages of the pathology [7]. As a consequence, depth-resolved information across the esophageal layers is necessary to increase the early detection rate of esophageal cancer and accurately classify the disease's stage. The average healthy esophagus wall is 5.26 mm thick for males and 4.34 mm thick for females [8] and is highly vascularized with vessels ranging from tens to several hundreds of micrometers in size within the mucosa [9].

Optoacoustic (OA) endoscopy is a potentially promising technique for esophageal cancer screening as it can deliver three dimensional, label-free, structural and functional imaging in vivo, offering high resolution molecular imaging through several millimeters of tissue [10,11]. Vascular remodeling and angiogenesis are common landmarks of tumor lesions [12] and the spectral absorption properties of hemoglobin in red blood cells offer an intrinsic optical molecular absorption contrast mechanism to map the vascular network using optoacoustics [13]. The imaging modality yields sub-surface visualization useful for identifying specific alterations of microvascular patterns for the detection of dysplasia and cancer [14]. Visualization of the vascular network through tissue layers can support assessment of tumor invasion and accurate staging of esophageal cancer [15], particularly in early phases of the disease where optimized treatment plans can lead to improved prognoses in the clinic.

OA imaging is based on ultrasound detection induced by light absorption and thermoelastic expansion of molecules within biological tissues, thereby offering high molecular specificity and rich contrast at greater depths than purely optical imaging [16]. These features render OA endoscopy a potentially powerful modality for evaluating and staging malignant tissue within the gastrointestinal (GI) tract. Current OA endoscopes however suffer from low imaging speeds, incomplete fields of view and poor signal-to-noise-ratios (SNR). In order to deliver performance relevant for medical use, imaging the GI tract with optoacoustic endoscopy requires: a) high imaging frame rates to avoid motion artefacts, b) large field of views to expedite the screening procedure, c) careful design of broadband transducers with a miniaturized form factor to be inserted through hollow bodily cavities, and d) appropriate probe geometry to ensure mechanical stability and for maintaining optimal contact between the endoscope and the esophageal circumference for acoustic index matching.

Existing implementations of OA endoscopy include a 4 Hz-distal rotation endoscope which was employed in simultaneous optoacoustic and ultrasonic endoscopy for imaging internal organs of a rabbit [17]. However, the device has a limited field of view due to driving cables crossing over the imaging plane, with the use of reflective ultrasound elements refracting sound and leading to distortions and signal attenuation. Another group showed a proximal scanning catheter with the ability to perform full field of view optoacoustic and ultrasonic dual modality imaging of colorectal tissue of a Sprague Dawley rat [18]. However, this system is limited to a slow frame rate of 5 Hz due to low laser repetition rate and the rotation speed of the rotary joint. A fast endoscope was demonstrated with proximal rotation reaching 50 Hz with a 300 kHz repetition rate laser and high speed rotatory joint to image the rectum wall of a Sprague Dawley rat [19]. This design suffers from poor lateral resolution due to the use of an unfocused acoustic detector and loss of SNR due to lower light transmission efficiency through the fiber optic rotatory joint. The drop in SNR necessitated the use of a 15 mJ/cm^2 fluence to obtain images, significantly exceeding the maximum permissible exposure limit (MPE) at high repetition rates (see discussion). Furthermore, proximal rotation via torque coils are problematic due to uncontrolled coil friction causing heat, limiting the rotation speed, and affecting scanning uniformity during operation. Loss of scanning uniformity results in loss of image formation accuracy.

The above-mentioned configurations [17-19] focused on OA esophageal endoscopes for small animals, aimed at further miniaturization of the probe for insertion into the working channel of white endoscopes. Nevertheless, these solutions exhibit poor OA coupling against the wall due to the dimension of the human esophageal lumen and long propagation distances to the lumen surface, leading to an overall loss of SNR. A side looking optoacoustic capsule endoscope designed to be in direct contact with the esophagus wall could improve the OA coupling and OA sensitivity by placing the sensor closer to the interrogated tissue surface. A prototype capsule-shaped acoustic resolution optoacoustic probe was shown to to characterize the intestinal structures of Crohn's disease in a rabbit [20]. However, the imaging speed was limited due to a slow laser pulse repetition rate of 10 Hz and a low SNR, necessitating the signal to be averaged 30 times, slowing acquisition times considerably. Our group [21] has recently demonstrated a side looking capsule endoscope optoacoustic imaging prototype with a 360° field of view; however, the imaging speed that was limited by a laser repetition rate of 2 kHz and the proximal rotational scanning unit rate of 2.5 Hz.

Clinical translation of current OA capsule endoscopes are limited by their slow scanning speeds. Our approach reported herein instead offers video rate (50 Hz), 360-degrees, side-looking OA signal acquisition by implementing distal rotation coupling of the acoustic sensor and laser illumination, paving the way towards clinical translational of OA capsule endoscopy of the GI track. We characterized the endoscope's lateral resolution by imaging a phantom consisting of twelve 100 µm sutures arranged in an Archimedean spiral with a pitch of 3 mm, and demonstrated a volumetric pullback spiral scan over 1790 mm^3 at 0.24 Hz. To simulate scattering of light in esophageal tissue, the acoustic coupling medium of water was replaced with various concentrations of an intralipid solution to determine an estimated penetration depth of ~0.84 mm in the human esophagus. To validate the estimated penetrated depth, we imaged an *ex vivo* pig esophagus. Our results demonstrate proof-of-principle translational potential of optoacoustic endoscopy for delivering video-rate 3D functional information of the human esophagus.

Material and methods

The benchtop optoacoustic endoscope unit as illustrated in Figure 1 consists of a ns-pulse width, 532nm Q-Switched DPSS laser with an operating repetition rate in the range of 1-100kHz (ONDA, Bright Solutions, Italy). The energy per pulse was regulated by means of a polarizing beam splitter and a beam dump. A 90:10 beam splitter was employed to divert 10% of the light towards a photodetector for triggering. The remaining 90% of light was coupled into a fiber to guide the optical beam to the optoacoustic semi capsule endoscope.

The proposed 16 mm diameter semi-capsule prototype consists of a miniature motor with an ultrafine hollow shaft (Maxon Group, Switzerland) to allow for coaxial low–beam divergence optoacoustic illumination with a custom designed GRIN lens mounted to the distal end of the multimode optical fiber with a core diameter of 105 µm and numerical aperture of 0.2 (GRINTECH, Germany). The detector features a 28 MHz-central frequency, side-looking, planar lithium niobate ($LiNbO_3$) active element transducer with a 6 mm-fused silica acoustic lens, and a central aperture of 2 mm in diameter to allow coaxial illumination (SONAXIS, France). To facilitate side illumination, the transducer is integrates a 90^o silver coated optical reflector with >95% reflectivity at 532 nm, surface flatness of λ/4, and surface quality of 60-40 scratch-dig. Furthermore, a sealing window manufactured with PMMA altuglas SG7 with a thickness of 1 mm was installed at the central aperture of the detector. The optical beam was a low divergence beam with a half divergence angle of 3.9° and focal length of 7.4 mm along the air and water optical path length inside the capsule. Full 360^o unobstructed field of view is achieved by implementing customized low friction miniature gold slip rings (STATICE, France) to enable electrical coupling of high frequency ultrasound signals, thereby avoiding obstruction from driving cables over the field of view. The capsule is encapsulated in water with a cylindrical PMMA capsule lid featuring a thin fiducial marker used for synchronization and interpolation of high frame rate B scans.

In order to drive the rotational motor, an ESCON module 24/2, 4-Q servo driver was operated using ESCON studio via PC. The optoacoustic signals measured during rotational scanning were pre-

amplified externally with an 30dB amplifier (Sonaxis SA, France) and connected to a T-bias (ZFBT-4R2GW+, Mini-Circuits) prior to data acquisition at a sampling rate of 500 MHz. The T-bias is used to provide a DC voltage supply to the pre-amplifier circuit while remaining decoupled from the radiofrequency (RF) line of the transducer, hence voltage fluctuations from the DC voltage source do not affect the RF signal. The semi capsule prototype was mounted to an XYZ linear motorized stage (MTS50-Z8 in 3-axis XYZ Configuration Thorlabs) to allow for controlled repositioning and performed 3D Helical pull-back scans. The optoacoustic data was acquired by a ATS9373 DAQ card (AlazarTech, Canada) synchronized to the board's external trigger port via the photodetector signal. Signal processing was controlled by LabVIEW and data analyzed in MATLAB.

The endoscopic transducer was characterized by linear raster scanning the semi capsule against a 20 µm suture in the YZ plane by shifting the transducer away from the acoustic source generated via the suture (y steps = 10 µm, z steps = 5 µm) to determine the temporal/frequency response at focus, the axial/lateral resolutions, and depth-of-focus (See supplementary for further information).

To demonstrate 50 Hz B-scan frame rate with 100 kHz pulse repetition rate (31.4 µm focal plane lateral sampling interval), a phantom comprising twelve 100 µm sutures arranged in an Archimedean spiral with a pitch of 3 mm as depicted in Figure 2 (a) was imaged. The phantom was designed to ensure that one of the sutures lies within the focal plane of the transducer to determine the influence of laser pulse repetition rate and B-scan frame rate on the lateral resolution at focus.

A 3D helical pull back optoacoustic scan over 10 mm of a stainless steel mesh arranged in an Archimedean spiral with a pitch of 3 mm (Figure 3a) was performed at 50 Hz B-scan rate and 100 kHz repetition rate with a pullback speed of 2.4 mm/s over 4.16 seconds acquiring a volume of 1700 x 2000 x 208 (ρ, φ, z) voxels (or 5mm x 360º x 10mm or 1790mm^3). This experiment was repeated by replenishing the acoustic coupling medium with controlled titration concentrations of intralipid (IL) solution to simulate the scattering of light during the interrogation of human esophageal tissue. A 20% IL solution was diluted to six test IL concentrations including 0.015, 0.031, 0.062, 0.125, 0.25 and 0.5%. The main molecular absorber in mucosa at 532 nm is hemoglobin, which is in circulation in blood vessels (simulated by the mesh) whereas the connective tissue in mucosa acts primarily as a scattering medium at this wavelength (simulated by the IL). The detection depth limit was defined by the deepest position where the signal is greater than the system noise level of 1 mV, which is SNR=1 for the intralipid mesh experiment shown in Figure 4(o).

The same scanning configuration was also performed on an *ex vivo* pig esophageal sample mounted inside a 50 mL-Falcon tube to demonstrate translational potential. In an attempt to preserve the vascular network in the esophagus during sample extracting, a heated blade was used to cut and seal each end of the esophagus. The pulse laser energy was measured at the capsule surface during all imaging and was maintained at ~12 µJ per pulse. Energy measurements were initially made using an Ophir energy meter (model PE10-C) which can operate up to a pulse repetition rate of 25 kHz. The OA signal strength generated by a reference marker was used as a surrogate monitor of the pulse energy for repetition rates beyond the limit of the energy meter to keep the pulse energy ~12 µJ per pulse.

Results

Endoscopic optoacoustic transducer characterization

The structure and setup of the proposed endoscopic transducer is shown in Fig 1a and 1b and described in the Methods section. The detector characterization procedure revealed the spatial features of the transducer. The sensor exhibited a broadband detection performance with a central frequency of 28 MHz and a -6dB bandwidth of 50 MHz as depicted in Figure 1(c). The temporal response indicated a transducer focal length of 6 mm. The axial and lateral resolutions were measured as 35 µm and 200 µm, respectively, as shown in Figure 1(d). A 700 µm depth-of-focus was measured from the Hilbert plot shown in Figure 1(e).

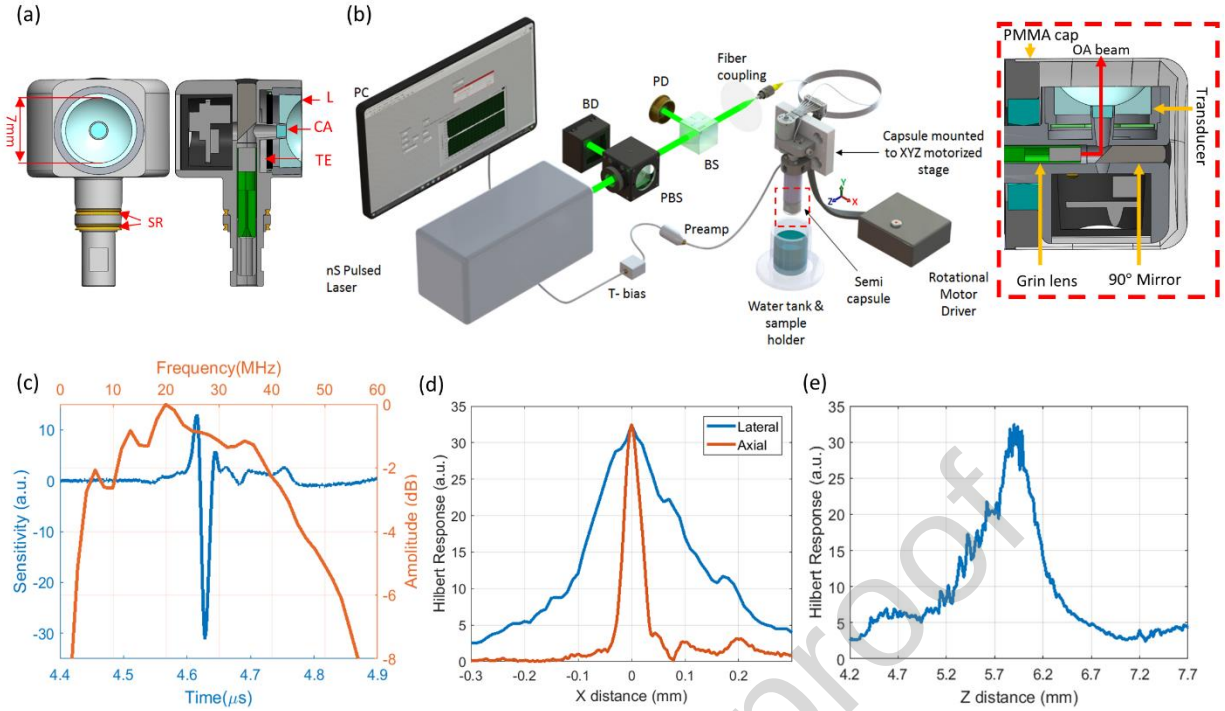


Figure 1: Endoscopic transducer characterization (a) Cross-sectional and en-face view of the endoscopic transducer with a central aperture. TE: Transducer element, EC: Electrical Connector, L: Acoustic Lens, CA: Central Aperture Window, SR: Slip rings. (b) Benchtop optoacoustic endoscopy experimental setup. PBS: polarizing beam splitter, BS: beam splitter, BD: beam dump, PD: photodiode. (c) Optoacoustic transducer temporal and frequency response. (d) Axial and lateral Hilbert transducer resolutions (e) Hilbert depth-of-focus response.

OA endoscopic imaging of a suture phantom at 50 Hz B-scan frame rate

To characterize endoscope's lateral resolution performance as a function of laser pulse repetition rate and scanning speed, we first imaged a phantom comprised of twelve 100 µm sutures arranged in an Archimedean spiral with a pitch of 3 mm (Figure 2a). Figure 2(b) depicts a B scan acquired at 50 Hz with a pulse repetition rate of 100 kHz. The B scan highlights the near-field, focus, and far-field response of the transducer, with the sensor sensitivity reaching its maximum value at focus which is located 2mm away from the capsule surface.

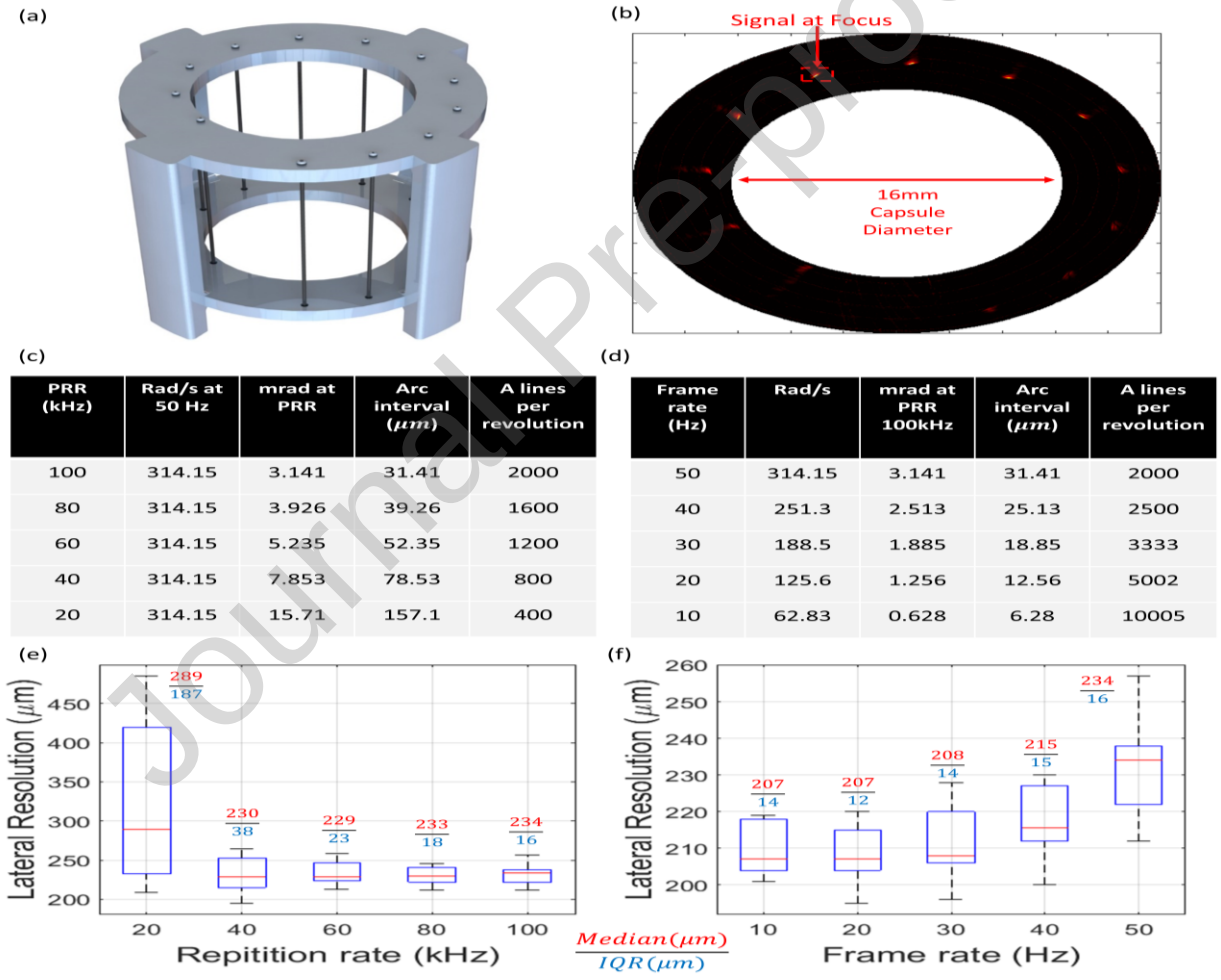
To assess the influence of the arc sampling interval and scanning speed on the lateral resolution at focus, we acquired continuous scans at the maximum frame rate (50 Hz) with different laser pulse repetition rates (20 kHz to 100 kHz), followed by scans at the maximum repetition rate (100 kHz) with different frame rates (10 Hz to 50 Hz). For each case we acquired a set of 10 continuous revolutions and calculated the lateral resolution for each rotation scan. The median and interquartile ranges (IQR) of the measurements were used as indicators of resolution and variability for each condition.

Table (c) in Figure 2 shows the arc sampling intervals and number of A lines obtained per revolution at a constant B frame rate of 50 Hz for pulse repetition rates ranging between 20 kHz and 100 kHz. The median lateral resolution between repetition rates 40 to 100 kHz is approximately the same, varying between 229-234 µm. However, at 20kHz the median lateral resolution degrades to a value of 289 µm. The reduced median lateral resolution of 289 µm at 20 kHz laser pulse rate is attributed to a low lateral arc sampling interval of 157 µm at the focal plane. This case yielded values ranging between 209 µm and 485 µm, which is expected due to the violation of a Nyquist sampling criteria for sampling at a minimum of double the resolvable lateral resolution. For the proposed transducer this would correspond to an approximately 100 µm arc sampling interval. Additionally, as the repetition rate is

increased from 20 to 100 kHz, the variability between revolutions decreases due to an increased angular sample resolution per B frame; and we therefore observed a decrease in the IQR from 187 µm to 16 µm.

Table (d) in Figure 2 shows the arc sampling intervals and number of A lines obtained per revolution at a constant laser pulse repetition rate of 100 kHz for B frame scanning rates ranging between 10 Hz and 50 Hz. The resulting lateral resolution distribution measurements are plotted in Figure 2(f). Here, the median lateral resolution improved from 234 µm (50 Hz) to 207 µm (10 Hz), almost matching the lateral resolution of the transducer of 200 µm as characterized in Figure 1. Although improvement in arc sampling interval is always obtained when reducing frame rates, in this instance no significant improvement in the IQR, fluctuating between 12-16 µm, was observed over the 10-50 Hz test B frame rates

Overall, by reducing the B scan frame rate and using a higher laser pulse repetition rate, arc sampling intervals are reduced during measurement which, however, results in an increased number of A scans per revolution, impacting file size and memory management. Nevertheless, our results show arc sampling intervals below 18.8 µm obtained using a 30 Hz scan rate and 100 kHz repetition rate, indicating no significant improvement in the median value of the lateral resolution corresponding to one tenth of the transducer lateral resolution.



PRR (kHz)	Rad/s at 50 Hz	mrad at PRR	Arc interval (µm)	A lines per revolution
100	314.15	3.141	31.41	2000
80	314.15	3.926	39.26	1600
60	314.15	5.235	52.35	1200
40	314.15	7.853	78.53	800
20	314.15	15.71	157.1	400

Frame rate (Hz)	Rad/s	mrad at PRR 100kHz	Arc interval (µm)	A lines per revolution
50	314.15	3.141	31.41	2000
40	251.3	2.513	25.13	2500
30	188.5	1.885	18.85	3333
20	125.6	1.256	12.56	5002
10	62.83	0.628	6.28	10005

Figure 2: Endoscope performance at 50 Hz B-scan frame rate (a) Twelve sutures with 100µm thickness arranged in an Archimedean spiral phantom with a pitch of 3mm. (b) 360º B scan image of phantom acquired at 50Hz and 100kHz pulse repetition rate. (c) Table listing the number of A lines per revolution and arc sampling interval at the transducer focal plane as a function of pulse repetition rate at a fixed frame rate of 50Hz. (d) Table listing the number of A lines per revolution and arc sampling interval at the transducer focal plane as a function of frame rate at a fixed pulse repetition rate of 100kHz (e) Lateral resolution measured at transducer focus at 50Hz frame rate as a function of laser pulse repetition rate under the same experimental configurations listed in table (c). (f) Lateral resolution measured at transducer focus at 100kHz laser pulse repetition rate as a function of B scan frame rate under the same experimental configurations listed in table (d).

Volumetric OA endoscopy imaging of a mesh phantom at 50 Hz B-scan frame rate

We next set out to demonstrate the volumetric imaging capability of the endoscope. We performed a Cscan of a stainless steel mesh at 50 Hz B-scan rate arranged in an Archimedean spiral set at a pitch of 3 mm (Figure 3a) which served as a phantom sample, with the minimum diameter of 16 mm and the maximum of 22 mm falling within the range of a human esophagus diameter [22]. Figure 3(b) depicts the 3D projection of the helical scan at 50 Hz and 100 kHz repetition rate pulled back over 10 mm at 2.4 mm/s. The full scan was acquired in 4.16 seconds with an arc sampling resolution of 31.4 µm at focus (r = 6 mm), an axial sampling resolution of 3 µm (Fs = 500 MHz), and a helical pitch of 48 µm. Figure 3(c) and (d) illustrate the 2D maximum top and axial intensity projections generated from the helical scan, respectively, with the fiducial marker and focal plane highlighted. The maximum intensity projections indicate that the intensity values are highest at focus, as expected for a focus transducer; however, the mesh appears to be distinguishable over the entire field of view across the total depth of 3 mm.

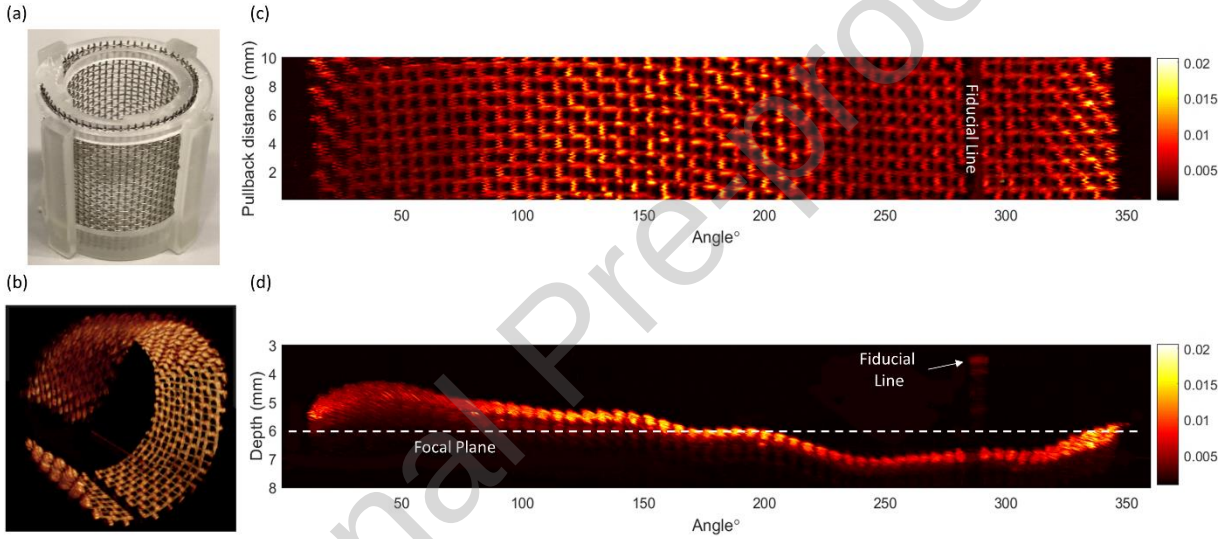


Figure 3: Helical volumetric Cscan over 10 mm longitudinal pullback distance at 50 Hz B-scan rate and 100 kHz repetition rate (a) Aluminum phantom mesh mounted in an Archimedean spiral with a pitch of 3 mm. (b) 3D polar projection of Helical optoacoustic scan generated on Amira (c) 2D maximum intensity projection of Helical scan depicted in Cartesian coordinates. (d) 2D Axial maximum intensity projection of Helical scan depicted in Cartesian coordinates.

Volumetric OA endoscopic imaging of a mesh phantom immersed in intralipid solution to mimic tissue scattering at 50 Hz B-scan frame rate

To assess the imaging penetration depth of the endoscope under light scattering conditions mimicking those of the human esophagus, we embedded the mesh in an intralipid (IL) solution at different concentrations. The maximum intensity projections from the top and side views of the volumetric scans were obtained and the fiducial marker masked out to avoid signal saturation and to restrict the analysis to intensities within the dynamic range from the mesh in the scattering media.

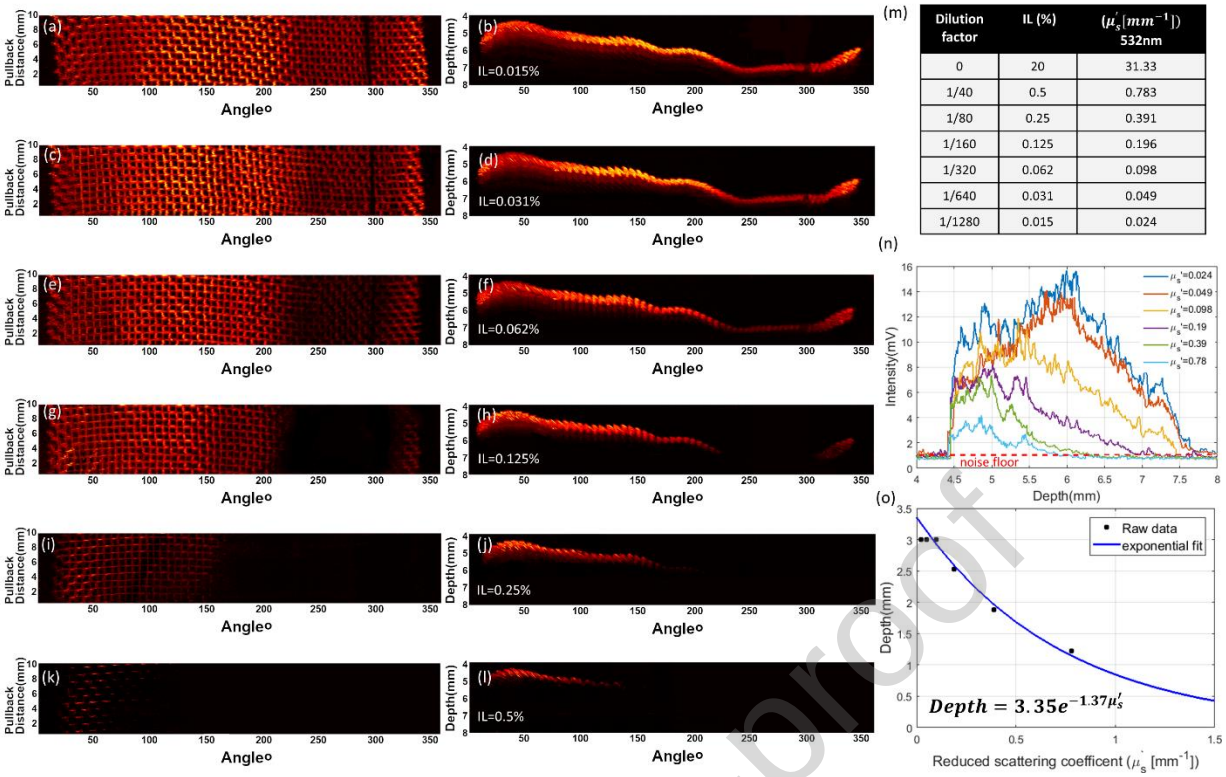
At 0.015% IL (Figures 4a and 4b), the maximum signal strength was 15.9 mV and the entire mesh was visible over the full 3 mm depth of the mesh. At 0.031% IL (Figures 4c and 4d), the maximum signal strength was reduced to 14 mV, but the entire mesh remained observable over the full depth. At 0.062% IL (Figures 4e and 4f), the maximum signal strength dropped to 11.9 mV, with the mesh still perceivable over the full 3 mm depth, but losing definition of the grid pattern in the far field. At 0.125% IL (Figures 4g and 4h), the maximum signal strength was 8.1 mV and the mesh in the far field was no longer observable, hence limiting the imaging depth to 2.53 mm. At 0.250% IL (Figure 4i and 4j), the maximum signal strength and depth were further reduced to 7.5 mV and the mesh was only

observable in the near field, limiting the imaging depth to 1.88 mm. Finally, at 0.50% IL (Figures 4k and 4l), the maximum signal strength reduced to 4.05 mV, the grid pattern of the mesh was no longer defined and the imaging depth was limited to 1.22 mm.

Light scattering and diffusion affects the amount of light available for OA generation from tissue. To estimate the expected imaging depth in the human esophagus, we calculated the reduced scattering coefficient, μ_s', at 532 nm for the IL solutions in our experiments based on previously reported values at 20% IL concentration [23]. Table (m) in Figure 4 shows the μ_s' values for each IL concentration (0.015% to 0.5 %).

Figure 4(n) shows the maximum intensity values for the volumetric mesh images as a function of depth. We determined the position at which the signal intensity drops below the noise level, which was identified from the average intensity value at depths between 4 mm and 4.4 mm (Figure 4n), where no part of the mesh is present in our sample, yielding a background noise level of ~1 mV. We therefore used 1 mV-noise level to estimate the imaging depth achievable for each scattering conditions using our OA endoscopic configuration. Interestingly, we observed that the intensity depth profiles agree with the response of our focused detector (see Figure 1e) for the low scattering cases with μ_s' = 0.024 and μ_s' = 0.049. Here, the peak intensity is seen at 6 mm, corresponding to the focal point of the transducer. For μ_s' at and above 0.098, the peak intensity shifts towards the near field of the transducer as expected from the reduced light penetration with increased scattering. This shift suggests that the focal distance of the transducer could be tuned close to the surface of the sample, matching the highest sensitivity of the transducer to the attainable light penetration available for OA signal generation, therefore increasing the SNR and thereby penetration depth under high scattering media conditions.

Figure 4(o) illustrates the relationship between μ_s' and penetration depth. The data is fit to an exponential equation of the form 3.35exp(-1.37μ_s'). Healthy human esophagus has been previously reported with μ_s'=$1.01 \pm 0.14 mm^{-1}$ at 526 nm [24]. We chose a scattering coefficient reported for superficial mucosal layers (highlighted in Table S1) rather than a fully perfused esophagus, as our model splits the contribution of blood absorption using the mesh inserted in the phantom. Assuming an insignificant change in μ_s' from 526 nm to 532 nm, and following our derived expression for the penetration depth and μ_s', we predict an imaging penetration depth in the human esophagus of ~0.84 mm exceeding the MPE by a factor of 4 at a 100 kHz-pulse repetition rate and 50 Hz-B-scan frame rate (see Supplementary).



Dilution factor	IL (%)	$(\mu_s'[mm^{-1}])$ 532nm
0	20	31.33
1/40	0.5	0.783
1/80	0.25	0.391
1/160	0.125	0.196
1/320	0.062	0.098
1/640	0.031	0.049
1/1280	0.015	0.024

Figure 4: 360º helical volumetric scanning over a 10 mm longitudinal pullback distance with intralipid solution. (a,b) 2D top and side maximum intensity projections, respectively, with 0.015% intralipid solution. The corresponding figure pairs (c,d), (e,f) , (g,h) ,(i,j) , (k,l) represent the 2D projections with 0.031% , 0.062%,0.125% ,0.25%,0.5% intralipid concentration, respectively . (m) Table listing the corresponding reduced scattering coefficient μ_s' at each IL concentration at 532nm (n) The maximum intensity profile as a function of depth for all scattering coefficients (o) The relationship between depth and reduced scattering coefficient.

Volumetric OA endoscopic imaging of an ex vivo pig esophagus at 50 Hz B-scan frame rate

To validate our phantom model and demonstrate translation potential of our endoscope, we imaged an *ex vivo* female pig esophagus (Figure 5a) mounted inside a 50 mL falcon tube. Figure 5(b) illustrates the 3D volume of the acquired pullback scan at 50 Hz B scan rate and 100 kHz over a pullback distance of 12 mm at 2.4 mm/s. The full scan was acquired in 5 seconds with an arc sampling resolution of 31.4 µm at focus, an axial sampling resolution of 3 µm, and a helical pitch of 48 µm. Figure 5 (c) and (d) depict B scans taken from two arbitrary positions from the helical pull back scan demonstrating the ability to acquire data over a 360° field of view. Figure 5(e) and (f) illustrate the 2D enface and axial maximum intensity projections generated from the helical scan, respectively, with the fiducial marker and capsule surface highlighted. Although in the *ex vivo* setting the vascular network has collapsed and the blood has drained, we were still able to measure remnants of blood in vessels visible over the 360° field of view and up to a penetration depth of ~1.2 mm.

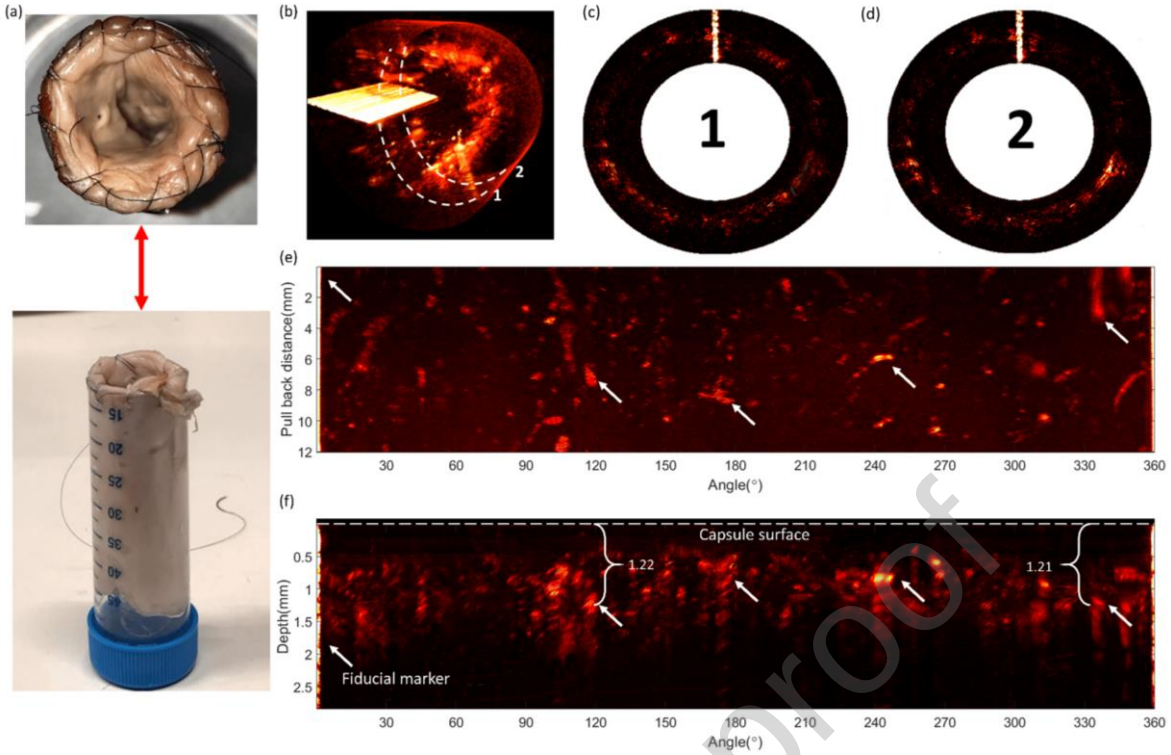


Figure 5: 360º helical volumetric scanning of an *ex vivo* female pig esophagus over a 12 mm longitudinal pullback distance. (a) *Ex vivo* pig esophagus sample mounted on a 50 mL-Falcon tube for support (b) 3D volumetric render of helical scan (c,d) B scans taken at unique positions within the helical scan. Maximum intensity projection for (e) enface and (f) side views.

Discussion and conclusion

While holding tremendous promise for GI tract diagnostics, optoacoustic endoscopes have incomplete fields of view, low B scan frame rates, inappropriate probe construction and insufficient SNR for human esophageal imaging. We have enabled for the first time a 50 Hz, full 360° field of view optoacoustic capsule endoscope by integrating a side-looking broadband focused transducer electrically coupled with slip rings and a central aperture for coaxial illumination. We demonstrate imaging of 1700 x 2000 x 208 (ρ, φ, z) voxels, corresponding to a 0.24 Hz volume rate. We show imaging depths of 3 mm in low scattering media and an estimated imaging depth of ~0.84 mm for esophagus-mimicking phantoms at 100 kHz-laser repetition rates. The overall esophageal wall thickness in patients with Barrett's and early cancer is 3-4 mm (including muscle layer) [25]. The thickness of Barrett esophagus has been measured by volumetric laser endomicroscopy and found to vary between 200 µm and 700 µm and it is generally always less than 900 µm [26]. Therefore, we believe that our OA endoscopy configuration should be able to image tissue within the depths expected from metastatic transformations occurring in Barrett's esophagus.

Optoacoustic imaging was previously shown to reveal the vascular structure in skin melanoma and provide blood oxygenation maps to assess tumour angiogenesis and depth [14]. This performance could be useful for tracking esophageal lesions where the wall is rich in vasculature. Our group has recently proposed optoacoustic capsule endoscopy for radial imaging of the GI track [21]. This approach enabled uninterrupted 360° imaging for the first time, yet the B frame rate was limited to 2.5 Hz, preventing fast volumetric scanning. In the current work, we have achieved a 20-fold higher B scan frame rate by integrating the rotation coupling directly onto the transducer housing inside the capsule with customised miniature slip rings. To our knowledge, this enabled for the first time video-rate optoacoustic capsule endoscopy at 50Hz which could lead to real time optoacoustic endoscopy surveillance of the GI track and can mitigate motion artefacts. We assessed the influence of pulse repetition rate (20-100 kHz) and B scan frame rate (10-50 Hz) on lateral resolution. Increasing the laser pulse repetition rate or the B scan frame leads to a reduced angular sampling resolution, which is

expected from the increased interval between pulses available on each rotation cycle (number of A lines per revolution). We found that the maximum lateral resolution can be reached by sampling at 1/10th of the characteristic transducer lateral resolution at focus. Transducer characterisation yielded a 200 µm lateral resolution, leading to arc sampling of 20 µm required at focus. At 100 kHz-pulse repetition rate and 50 Hz-B frame rate, the arc sampling interval was 33 µm ($\varphi = 0.18°$), which showed a degradation of the maximum lateral resolution to 234 µm, with an interquartile variation range of 16 µm. In contrast, at the same repetition rate, reducing the B scan frame rate to 30 Hz results in an arc sampling interval of 18 µm ($\varphi = 0.1°$), improving the achievable resolution to 208 µm, closer to the maximum attainable value of the transducer. As expected, reducing the B scan frame rate further allowed a decrease down to 6.2 µm ($\varphi = 0.035°$) at 10 Hz; however, this does not lead to a resolution improvement as it lies below the transducer limit.

We showcase the ability to acquire large volumetric imaging at 0.24 Hz over a total volume of $5\ mm \times 360° \times 10\ mm\ (\rho, \varphi, z)$ or 1790 mm^3, at a high sampling resolution of $3\ \mu m \times 0.18° \times 48\ \mu m\ (\rho, \varphi, z)$. This high-resolution volumetric imaging speed can enable a complete scan of the entire human esophagus up to a length of 250 mm [27], in under 2 minutes. The ability to rapidly capture large volumes enables analysis of multiple high risk regions within the human GI tract including large Barrett's esophagus segments. In media without scattering, the OA imaging penetration depth exceeded 3 mm. However, we show that the penetration depth exponentially depends on the scattering coefficient of the medium. The maximum transducer sensitivity at focus is preserved in low scattering cases ($\mu_s' \leq 0.049$); however for higher scattering media, light diffusion becomes dominant over the transducer sensitivity response, shifting the maximum sensitivity towards the near-field of the transducer. Following the exponential relationship with increasing μ_s', we estimate the imaging penetration depth in the human esophagus of our system to be 0.84 mm (assuming a $\mu_s' \sim 1.01$ @ 532 nm). We validated our phantom model by performing a helical pullback scan of an *ex vivo* pig esophagus at the same volumetric image rate of 0.24 Hz. Although the vascular network in the sample had collapsed, we were still able to measure remnants of vessels up to a depth of ~1.2 mm. We observed deeper structures in the pig esophagus than could be predicted from our phantom imaging (Figure 4). The discrepancy can be attributed to the greater proportion of lower acoustic frequency components from the thick vessels in the pig esophagus, which experience lower propagation losses. For instance, the mesh in the phantom has a dimeter of 0.25 mm, whereas the blood vessels in the pig esophagus measured in Figure 5(f) were ~0.7 mm in diameter. A second pig esophagus sample was scanned under the same conditions and revealed preserved vascular structures at imaging depth up to 0.96 mm (see supplementary Figure S1).

This depth is on par with previous OA studies, demonstrating images of small animal esophagus or colon tissue with penetration depths of ~ 1 mm in vivo at < 5 Hz [17, 18, 20]. As a result of the low frame rate coupled with low laser pulse repetition rates, the MPE rule 1 (20mJ/cm^2) is not violated (see supplementary). Under these conditions He et al [21] were able to demonstrate an imaging penetration depth of 2 mm in pig esophagus ex vivo. Increasing the pulse repetition rates imposes tighter limitations on the MPE calculation where 20 mJ/cm^2 is no longer an applicable limit. In our case, the MPE limit is 1.53 mJ/cm^2 (see supplementary). Previously, a 300 kHz pulse repetition rate with 30 µJ-pulse energies with a 1 mm grin lens producing a collimated beam were used in a catheter-based optoacoustic endoscope system [19]. The authors reported an energy density of 15 mJ/cm^2, and following the limited aperture criterion of 3.5 mm of the ANSI standard, the MPE is exceeded by 174-fold in this case. These findings demonstrate that the translational potential of optoacoustic endoscopy is driven by the ability to rapidly scan the human esophagus, albeit currently exceeding the MPE by a factor of 4, while offering high resolution over several mm's of depth.

The ex vivo performance of the optoacoustic capsule design presented herein represents a significant leap toward video-rate in vivo optoacoustic endoscopy of the human GI tract. For *in vivo* measurements, we anticipate that the capsule will stretch the esophageal track and thereby achieve optimal contact between the tissue surface and capsule by peristaltic force. More sophisticated balloon configurations could alternatively be explored. However, in order to translate the design for in vivo imaging, future efforts are required to address capsule size, sensitivity, resolution and motor stability compensation.

While our current capsule design has a diameter of 16 mm, similar to standard WLSE, further miniaturization is required to ease swallowing and match the size of commercially available white light capsules, with sizes typically less than 32 mm in length and 13 mm in diameter [28,29]. For instance, the overall length and diameter can be reduced by reducing the size of the transducer and by integrating a miniature motor inside the capsule. To further increase the sensitivity and imaging depth, a transducer pre-amplification stage can be integrated inside the capsule [21] while focusing the transducer sensitivity towards the surface of the capsule, thereby increasing the SNR for imaging in high scattering media. To increase the resolution of the optoacoustic images, we anticipate that a transducer with higher central frequency and reduced focal length, as currently employed in raster scanning optoacoustic mesoscopy, could be incorporated [10,30]. Furthermore, image post-processing techniques can be explored to compensate for motor instability during scans. In this study only a single light source at 532 nm was utilized due to the lack of high repetition rate pulse lasers available at alternative wavelengths. To increase penetration depth by employing light sources in NIR region, where light attenuation in tissue is reduced, our group is investigating the possibility of over driving laser diodes of several wavelengths at hundreds of kilohertz and realizing their potential for video rate optoacoustic endoscopy [31]. Furthermore, oxygenation saturation measurements require illumination with at least two wavelengths, which would require wavelength multiplexing and reducing imaging speed by a factor equal to the number of wavelengths used. Alternatively, to retain high imaging speed, the spatial angular sampling rate can be reduced by a factor corresponding to the number of wavelengths used at the cost of image resolution. Our group recently investigated the use of nanosecond-pulsed laser diodes for OA imaging using frequency comb modulation to enable simultaneous laser wavelength illumination at high pulse repetition rates and enable oxygen saturation measurements without compromising imaging speed and quality [32].

In conclusion, a 16 mm optoacoustic capsule endoscope prototype with a side viewing central aperture endoscopic transducer with integrated optics and slips rings allowed for full field of view optoacoustic imaging at 50 Hz frame rate with angular sampling resolution of 0.18º attributed to the 100kHz pulse repetition rate. Its functional potential was assessed in phantoms and an *ex vivo* pig esophagus where lateral resolution accuracy and imaging depth, respectively, were demonstrated. Overall, these experimental results suggest that the novel video-rate optoacoustic capsule endoscope design can mitigate motion artefacts during imaging, offering promising translational characteristics for human GI tract molecular imaging in the future.

References

1. World Health Organization. International Agency for Research on Cancer. The Global Cancer Observatory. (2020). https://gco.iarc.fr/today/fact-sheets-cancers Accessed Jan 21, 2021.
2. P. Visaggi, B. Barberio, M. Ghisa, M. Ribolsi, V. Savarino, M. Fassan, M. Valmasoni, S. Marchi, N. de Bortoli, E. Savarino, Modern Diagnosis of Early Esophageal Cancer: From Blood Biomarkers to Advanced Endoscopy and Artificial Intelligence, Cancers 13(13) (2021).
3. B. Weusten, R. Bisschops, E. Coron, M. Dinis-Ribeiro, J.M. Dumonceau, J.M. Esteban, C. Hassan, O. Pech, A. Repici, J. Bergman, M. di Pietro, Endoscopic management of Barrett's esophagus: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement, Endoscopy 49(2) (2017) 191-198.
4. E.L. Bird-Lieberman, A.A. Neves, P. Lao-Sirieix, M. O'Donovan, M. Novelli, L.B. Lovat, W.S. Eng, L.K. Mahal, K.M. Brindle, R.C. Fitzgerald, Molecular imaging using fluorescent lectins permits rapid endoscopic identification of dysplasia in Barrett's esophagus, Nat Med 18(2) (2012) 315-321.
5. S. Menon, N. Trudgill, How commonly is upper gastrointestinal cancer missed at endoscopy? A meta-Analysis, Endosc Int Open 2 (2014) E46-50.
6. J. Yoon, J. Joseph, D.J. Waterhouse, A.S. Luthman, G.S.D. Gordon, M. di Pietro, W. Januszewicz, R.C. Fitzgerald, S.E. Bohndiek, A clinically translatable hyperspectral endoscopy (HySE) system for imaging the gastrointestinal tract, Nature Communications 10(1) (2019) 1902.

7. T. Attila, D.O. Faigel, Role of endoscopic ultrasound in superficial esophageal cancer, Dis Esophagus 22(2) (2009) 104-12.
8. F. Xia, J. Mao, J. Ding, H. Yang, Observation of normal appearance and wall thickness of esophagus on CT images, European Journal of Radiology 72(3) (2009) 406-411.
9. Kierszenbaum, A. L. & Tres, L. Histology and Cell Biology: An Introduction to Pathology with Student Consult Online Access 3rd edn. (Mosby, St. Louis, 2011).
10. J. Aguirre, M. Schwarz, N. Garzorz, M. Omar, A. Buehler, K. Eyerich, V. Ntziachristos, Precision assessment of label-free psoriasis biomarkers with ultra-broadband optoacoustic mesoscopy, Nature Biomedical Engineering 1 (2017) 0068.
11. M. Schwarz, D. Soliman, M. Omar, A. Buehler, S.V. Ovsepian, J. Aguirre, V. Ntziachristos, Optoacoustic Dermoscopy of the Human Skin: Tuning Excitation Energy for Optimal Detection Bandwidth With Fast and Deep Imaging in vivo , IEEE Transactions on Medical Imaging 36(6) (2017) 1287-1296.
12. R.H. Farnsworth, M. Lackmann, M.G. Achen, S.A. Stacker, Vascular remodeling in cancer, Oncogene 33(27) (2014) 3496-3505.
13. K. Haedicke, L. Agemy, M. Omar, A. Berezhnoi, S. Roberts, C. Longo-Machado, M. Skubal, K. Nagar, H.-T. Hsu, K. Kim, T. Reiner, J. Coleman, V. Ntziachristos, A. Scherz, J. Grimm, High-resolution optoacoustic imaging of tissue responses to vascular-targeted therapies, Nature biomedical engineering 4(3) (2020) 286-297.
14. M. Omar, M. Schwarz, D. Soliman, P. Symvoulidis, V. Ntziachristos, Pushing the Optical Imaging Limits of Cancer with Multi-Frequency-Band Raster-Scan Optoacoustic Mesoscopy (RSOM), Neoplasia 17(2) (2015) 208-214.
15. S.J. Hong, T.J. Kim, K.B. Nam, I.S. Lee, H.C. Yang, S. Cho, K. Kim, S. Jheon, K.W. Lee, New TNM Staging System for Esophageal Cancer: What Chest Radiologists Need to Know, RadioGraphics 34(6) (2014) 1722-1740
16. V. Ntziachristos, Going deeper than microscopy: the optical imaging frontier in biology, Nat Methods 7(8) (2010) 603-614.
17. J.-M. Yang, C. Favazza, R. Chen, J. Yao, X. Cai, K. Maslov, Q. Zhou, K.K. Shung, L.V. Wang, Simultaneous functional photoacoustic and ultrasonic endoscopy of internal organs in vivo, Nature Medicine 18 (2012) 1297.
18. Y. Li, R. Lin, C. Liu, J. Chen, H. Liu, R. Zheng, X. Gong, L. Song, In vivo photoacoustic/ultrasonic dual-modality endoscopy with a miniaturized full field-of-view catheter, Journal of Biophotonics 11(10) (2018) e201800034.
19. Y. Li, Z. Zhu, J. Jing, J. Chen, A. Heidari, Y. He, J. Zhu, T. Ma, M. Yu, Q. Zhou, Z. Chen, High-Speed Integrated Endoscopic Photoacoustic and Ultrasound Imaging System, IEEE Journal of Selected Topics in Quantum Electronics PP (2018) 1-1.
20. H. Lei, L.A. Johnson, K.A. Eaton, S. Liu, J. Ni, X. Wang, P.D.R. Higgins, G. Xu, Characterizing intestinal strictures of Crohn's disease in vivo by endoscopic photoacoustic imaging, Biomedical Optics Express 10(5) (2019) 2542-2555.
21. H. He, A. Stylogiannis, P. Afshari, T. Wiedemann, K. Steiger, A. Buehler, C. Zakian, V. Ntziachristos, Capsule optoacoustic endoscopy for esophageal imaging, Journal of Biophotonics 12(10) (2019) e201800439.
22. M. Ferhatoglu, T. Kıvılcım, Anatomy of Esophagus, 2017
23. R. Michels, F. Foschum, A. Kienle, Optical properties of fat emulsions, Opt. Express 16(8) (2008) 5907-5925.
24. J.A. Sweer, M.T. Chen, K.J. Salimian, R.J. Battafarano, N.J. Durr, Wide-field optical property mapping and structured light imaging of the esophagus with spatial frequency domain imaging, Journal of Biophotonics 12(9) (2019) e201900005.
25. K.R.S. Gill, M.S. Ghabril, L.H. Jamil, M. Al-Haddad, S.A. Gross, S.R. Achem, T.A. Woodward, M.B. Wallace, M. Raimondo, L.L. Hemminger, H.C. Wolfsen, Variation in Barrett's Esophageal Wall Thickness: Is It Associated With Histology or Segment Length?, Journal of Clinical Gastroenterology 44(6) (2010).
26. I.J.M. Levink, H.C. Wolfsen, P.D. Siersema, M.B. Wallace, G.J. Tearney, Measuring Barrett's Epithelial Thickness with Volumetric Laser Endomicroscopy as a Biomarker to Guide Treatment, Digestive Diseases and Sciences 64(6) (2019) 1579-1587.
27. B. Kuo, D. Urma, Esophagus - anatomy and development, GI Motility online (2006).
28. N. Shamsudhin, V.I. Zverev, H. Keller, S. Pane, P.W. Egolf, B.J. Nelson, A.M. Tishin, Magnetically guided capsule endoscopy, Medical Physics 44(8) (2017) e91-e111

29. A. Koulaouzidis, D.K. Iakovidis, A. Karargyris, E. Rondonotti, Wireless endoscopy in 2020: Will it still be a capsule?, World J Gastroenterol 21(17) (2015) 5119-5130.
30. Z. Ali, C. Zakian, V. Ntziachristos, Ultra-broadband axicon transducer for optoacoustic endoscopy, Scientific Reports 11(1) (2021) 1654.
31. A. Stylogiannis, L. Prade, A. Buehler, J. Aguirre, G. Sergiadis, V. Ntziachristos, Continuous wave laser diodes enable fast optoacoustic imaging, Photoacoustics 9 (2018) 31-38.
32. A. Stylogiannis, L. Prade, S. Glasl, Q. Mustafa, C. Zakian, V. Ntziachristos, Frequency Comb Optoacoustic Tomography, bioRxiv (2021) 2021.05.12.443808.
33. J.-W. Su, Y.-H. Lin, C.-P. Chiang, J.-M. Lee, C.-M. Hsieh, M.-S. Hsieh, P.-W. Yang, C.-P. Wang, P.-H. Tseng, Y.-C. Lee, K.-B. Sung, Precancerous esophageal epithelia are associated with significantly increased scattering coefficients, Biomedical optics express 6(10) (2015) 3795-3805.
34. C. Holmer, K.S. Lehmann, J. Wanken, C. Reissfelder, A. Roggan, G. Mueller, H.J. Buhr, J.P. Ritz, Optical properties of adenocarcinoma and squamous cell carcinoma of the gastroesophageal junction, J Biomed Opt 12(1) (2007) 014025.
35. B. Roland, A.W. Georges, D. Robert, C.M. Jerome, B. Daniel, B. Hubert van den, S. Jean-Francois, M. Philippe, Clinical measurements of tissue optical properties in the esophagus, Proc.SPIE, 1994.
36. M. Hohmann, B. Lengenfelder, R. Kanawade, F. Klämpfl, A. Douplik, H. Albrecht, Measurement of optical properties of pig esophagus by using a modified spectrometer set-up, J Biophotonics 11(1) (2018).

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 732720 (ESOTRAC), as well as from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 721766 (FBI). We thank Dr. Sergey Sulima for helpful discussions and for editing the manuscript.

Supplementary

Maximum permissible exposures for optoacoustic endoscopy

The maximum permissible exposures (MPE) are a set of maximum laser energy values to which the eye or skin may be exposed without adverse effects or injury occurring in normal circumstances. The levels have been established in standard ANSI Z136.1 various wavelengths and exposure durations in the eye or skin. Due to a lack of scientific data on the laser safety level on esophageal tissue, we based our safety requirements on laser radiation of the human skin, with the important note that the MPE of esophageal tissue might be lower than that.

For repetitively pulsed lasers the MPE is always the smallest MPE value determined by the application of the ANSI Rules 1 and 2 [ANSI Std. Z136.1-2000 (8.2.3.2)] applied over the area of the limiting aperture of 3.5 mm for skin imaging

$$MPE = \min[MPE_{rule1}, MPE_{rule2}]$$

Each of the MPE rules are defined as follows:

MPE Rule 1 (Single Pulse): No single pulse in a train of pulses shall exceed the MPE for a single pulse of the same pulse duration. For laser pulse duration of 1ns to 100 ns the single pulse radiant exposure over the wavelength of 400-700nm is 20 mJ/cm^2. Our laser has a pulse duration of 1 ns with an operating wavelength of 532nm.

$$MPE_{rule1} = 20\frac{mJ}{cm^2}$$

MPE Rule 2 (Average Power): The average MPE per pulse distributed over a number of laser pulses (N) over the exposure time(t) is given by:

$$MPE_{rule2} = \frac{MPE_T}{N}$$

Where N is the number of pulses during the exposure period given by the product of the pulse repetition rate(PRR) and exposure time (t) and the MPE_T radiant exposure for an exposure time between 100 ns to 10 s over the wavelength range of 400-700nm from the ANSI skin safety table is given by:

$$MPE_T = 1.1 * 10^4 t^{0.25} J/m^2 = 1.1 * t^{0.25} J/cm^2$$

Therefore the average MPE per pulse distribution as a function of PRR and exposure time t is given as follows:

$$MPE_{rule2} = \frac{1.1 * t^{0.25}}{Int(PRR * t)} J/cm^2$$

Note: the number of pulses in an exposure is always considered as an integer for laser hazard evaluations.

The exposure time interrogating the same sample location is a function of the laser spot size at the capsule window interface, in our case we apply the limited aperture for skin imaging with a diameter of 3.5 mm. Rotating at the maximum speed of 50 Hz (314.15 rad/s) at 100kHz pulse repetition rate the exposure time t is given by

$$t = \frac{Interrogation\ spot\ central\ angle\ (\ rad)}{frame\ rate\ (\frac{rad}{s})} = \frac{0.437}{314.15} = 1.39ms$$

Therefore:

$$MPE_{rule2} = \frac{1.1 * 1.39ms^{0.25}}{Int(100kHz * 0.139ms)} = 1.53 \frac{mJ}{cm^2}$$

Summary:

Therefore, the maximum permissible limits for a pulsed laser beam interrogating tissue with a spot size of approximately 500 µm, central wavelength of 532 nm, pulse duration of 1 ns and repetition rate of 100 kHz is given by the following table:

MPE rule 1 (mJ/cm^2)	MPE rule 2(mJ/cm^2)
20	1.53

The minimum of the ANSI rules is MPE rule 2, yielding the MPE of our endoscope to be 1.53 mJ/cm^2

During the course of all experimentation performed in this report an energy per pulse of approximately 12 µJ was employed. Hence energy density E_D employed was:

$$E_D = \frac{0.012}{\pi * (0.025^2)} = 6.11 \frac{mJ}{cm^2}$$

Our density is $6.11 mJ/cm^2$ and we thus exceed the maximum permissible energy of skin exposure by a factor of 4. To ensure this limit is not compromised, the repetition rate could be reduced from 100 kHz to 25 kHz to give a $MPE_{rule2} = 6.1\ mJ/cm^2$, although reducing repetition rate and keeping the frame rate fixed at 50 Hz corresponds to a lower lateral resolution as illustrated in Figure 2(e) at 50 Hz frame rate the lateral resolution is reduced from 234 μm at 100 kHz to 289 μm at 20 kHz repetition rate. To ensure the sampling interval is not compromised during high speed imaging the SNR of our transducer requires further enhancement by at least a factor of 4.

Helical volumetric scan of an ex vivo female pig esophagus sample

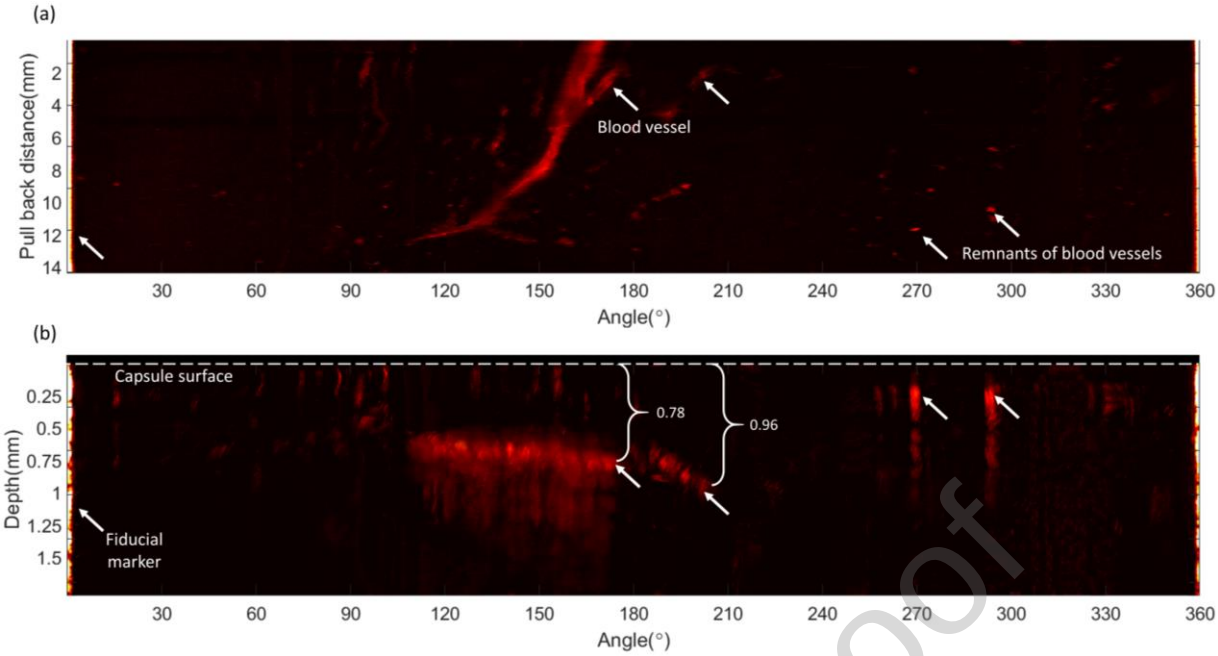


Figure S1: 360º helical volumetric scanning of an ex vivo female pig esophagus sample over a 14 mm longitudinal pullback distance. Maximum intensity projection for (a) enface and (b) side views.

Comparison of reported human and pig oesophageal optical properties

In our phantom model, we split the contribution of blood absorption using the mesh and the scattering contribution using intralipid. Authors in Ref [33] argue that the large variability reported is due to difference in the sample processing leading to high interpatient variability. Moreover, the standard deviation in Ref [34] was not given at 532 nm. Authors in Ref [35] argue that the large measurement variability for the reduced scattering coefficient reported is due to instrumental errors and sample inhomogeneity such as those introduced by narrow blood vessels, altering therefore the measurement. Moreover, it is possible that measurement variability can also results from unknown sampling depth in vivo. In Ref [36], the value for the reduced scattering coefficient is given as the mean of four measured layers of the esophagus, including the mucosa, submucosa, muscularis, and adventia, and are weighted by the layer thickness. The amalgamation of layers could explain the larger value reported, rather than that expected for the mucosal layer alone. We therefore chose a scattering coefficient reported for with the smallest standard deviation for superficial mucosal layers in Ref [24].

Table S1: Reported esophagus optical properties.

Reference	Year	Wavelength (nm)	μ_s' (mm^{-1})	g	μ_s (mm^{-1})	μ_a (mm^{-1})	Sample type
21	1994	514	1.9 ± 1.3	-	-	-	In vivo, human esophagus
22	2007	532	0.82 §	0.94 §	13.6 *	0.61*	Ex vivo, human esophagus covering the entire mucosal depth
23	2015	532	0.82 ± 0.48 §	0.946 ± 0.028	15.2± 4.2	-	Ex vivo, healthy human esophagus epithelium
24	2018	514	2.84 ± 0.48	0.77 ± 0.04	11.83 ± 1.39	0.23 ± 0.03	Ex vivo, pig esophagus in full, including muscle layers
13	**2019**	**526**	**1.01 ± 0.14**	**-**	**-**	**0.176 ± 0.042**	**Ex vivo, superficial human esopahgus**

* Extracted from graph

§ Calculated

Optoacoustic endoscopy precision golden slip rings design

The endoscopic slip ring coupling configuration is depicted in figure S2. The wires maintain electrical connection while the rings spin. To minimize friction induced electrical noise during rotation, the rings are covered with a 5µm- thick gold layer. Two wires are in contact with each ring, making four V contact points per pole, thereby reducing the risk of losing the electrical contact during rotation. To further support electrical contact between the rings and wires, a brass casing around the joint ensures the wires are unable to stretch outwards. The brass enclosure is also grounded to the coaxial cable transmitting optoacoustic signals in an attempt to reduce electromagnetic interference.

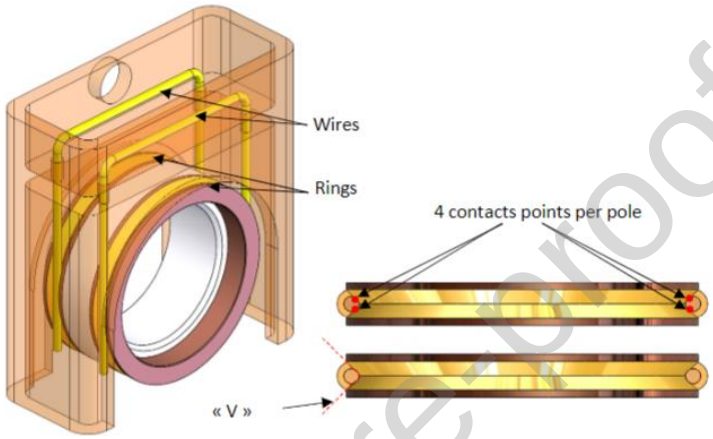


Figure S2:Transducer slip ring coupling design

Optoacoustic endoscopy probe characterization procedure

The optoacoustic endoscope transducer characterization configuration is illustrated Figure S3. The endoscope is mounted onto a 3-axis motorized linear stage XYZ (MTS50-Z8 in 3-axis XYZ Configuration Thorlabs) and immersed inside a 3D printed water tank with a 20 µm suture generating broadband optoacoustic signals. To determine the transducer's characteristics, the linear motorized stages are used to linearly raster scan in the YZ-plane in step intervals of 5 µm in the z-axis and 10 µm in the y-axis, which encompasses the maximum envelope of each A-scan when obtaining the transducer's sensitivity map. The global maxima from the sensitivity map determines the A-scan, which was obtained at the transducers focal length. The FFT and envelope of this A-scan represent the bandwidth and axial resolution at focus, respectively. The lateral resolution is determined from a line scan across the suture in the z-axis and the depth-of-focus is determined from a line scan moving along the y-axis with respect to the suture.

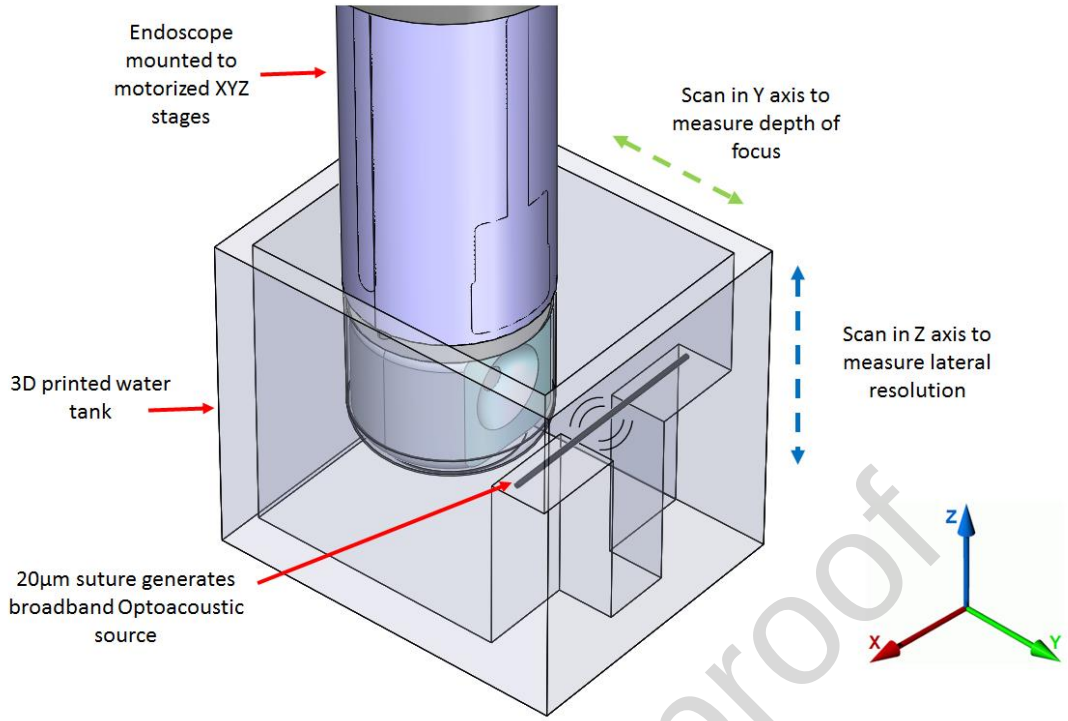
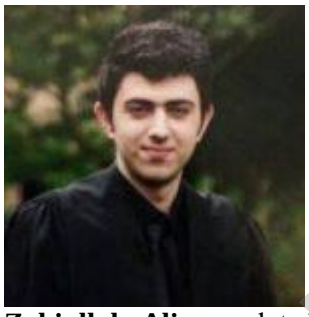


Figure S3:optoacoustic endoscope transducer characterization setup

Vitae



Zakiullah Ali completed his Bachelor's degree specializing in the field of Optoelectronics at the University of Auckland, NZ. Following undergraduate study, he worked as a medical devices engineer at Fisher and Paykel Healthcare Ltd, NZ focusing in the respiratory humidification sector. Having gained 3 years of work experience, he decided to return to academia by commencing the Integrated Photonic and Electronic System Masters program at the University of Cambridge, UK. Subsequently, he finished a mini-project characterizing bio-optical detectors at University College London, UK. He is currently working towards the development of a multi modal biophotonic imaging endoscope for early disease detection of esophageal cancer at the Technical University of Munich and Helmholtz Zentrum Munich.

Dr. Christian Zakian is the Group Leader for Optoacoustic Endoscopy at the Helmholtz Zentrum München. His main research interest is in the translation and adoption of optical and optoacoustic imaging methods to assist in medical diagnosis, and in particular for advanced endoscopy. He received

his PhD in physics from the University of Manchester in 2007 and has complemented his background with a life sciences Masters degree in the field of Inflammation at the William Harvey Research Institute, Queen Mary University London, in 2013. He has over 15 years' experience developing diverse imaging methods implemented for clinical studies.

Dr. Qian Li received his Ph.D. in biomedical engineering from Shanghai Jiao Tong University, China in 2018. Since then he has been a postdoc researcher at the Center for Medical Physics and Biomedical Engineering, Medical University of Vienna. His main research interests are developing optical imaging technologies, such as photoacoustic imaging and optical coherence tomography, for preclinical and clinical applications.

Jérôme Gloriod is a biomedical engineer. He obtained his DUT (Technical Degree) in Production and Mechanic Engineering in 2008, and became a graduated biomedical engineer in 2011 after 3 years training at ISIFC (Besançon). He then worked at Cisteo Medical in Besançon from 2011 to 2015 and joined Statice in 2016 as a project manager in the R&D department. From 2019 he has been leading the R&D implantable medical devices department. He will lead the technical development of the project.

Mrs Sophie CROZAT graduated from the Engineering School of Ceramics (ENSCI) in Limoges in 2001. She joined Sonaxis at the dawn of the company where she is now leading the operational activities. Her field of interests are include ultrasound probe technologies. She is actively participating in Sonaxis's research programs and is a co-inventor of Sonaxis's patents.

Mr François Bouvet, graduated in Electronics from the ESIGELEC School of Engineering, Rouen in 1987. He worked for 20 years with SEPEMA and further with SEPEMA technologies which were companies specialized in ultrasound generation and UT inspection scanning systems. Since 2010, he joined SONAXIS to run the electronics department. His fields of interests are with low and very high frequency electronic ultrasonic devices. He is actively participating in Sonaxis's research programs and he is a co-inventor of Sonaxis's patents.

Mr Guillaume PIERRE graduated from the Engineering School of Centrale Marseille in 1988. Since then, He has been active in the field of Optics and Acoustics over a 30-year period. He founded Sonaxis in 2002, a company based in Besançon, France, specialized in ultrasound probe making. He is currently serving as CTO and CEO of Sonaxis.

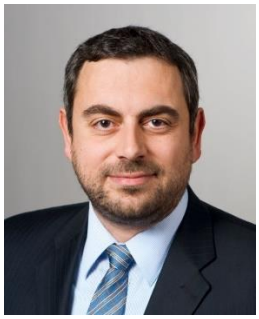
Vassilis Sarantos studied Electrical Engineering at Aristotle University in Thessaloniki, Greece and he is the founder of RayFos, a UK-based company, founded in 2012, in the field of scientific software engineering and real-time processing systems. He currently holds the position of the Chief Executive Officer and Chief Technical officer and his research interests include real time image/video processing, embedded system development, software development, medical imaging, and metrology system development.

Dr. di Pietro qualified from Federico II University of Naples, Italy and obtained an MD at the Institute of Molecular Cancer Research (University of Zurich, Switzerland). He is Senior Clinical Investigator Scientist at the MRC Cancer Unit, University of Cambridge and Honorary Consultant Gastroenterologist at Addenbrooke's Hospital, where he provides regional referral for endoscopic treatment of pre-cancerous and early malignant conditions of the upper GI tract, including endoscopic mucosal resection, endoscopic submucosal dissection and Radiofrequency Ablation. Dr di Pietro has a research interest in oesophageal and gastric pathologies, including Barrett's oesophagus, oesophageal cancer and hereditary gastric cancer. He has co-authored the current British and European guidelines for management of Barrett's oesophagus and the International Guidelines for the management of Hereditary Diffuse Gastric Cancer.

Professor Peter E. Andersen received his MSc.E.E. in 1991 and his Ph.D. in 1995 from the Technical University of Denmark, and since 2015 Dr. Andersen is Fellow of OSA and Fellow of SPIE. Dr. Andersen is with DTU Health Tech at the Technical University of Denmark, where he is group leader of the Bio Photonic Imaging Group, and Head of DTU Health PhD School. From 2006-2013, he was Research Professor at Department of Atomic Physics, Lund University, Sweden. He has more than 20 years of research experience within bio photonics, biomedical optics, light source technology, multimodal imaging concepts comprising optical coherence tomography, multiphoton tomography and nonlinear microscopy, and translation of these modalities into clinical applications, including endoscopic applications.

Professor Wolfgang Drexler was Full Professor of Biomedical Imaging at Cardiff University, UK from 2006-09. Since 2009 he is a Full Professor and the Head of the centre for Medical Physics and Biomedical Engineering at the Medical University of Vienna, Austria. He spent 2 years at the MIT,

Cambridge, USA, received the Austrian START Award (2001) and the COGAN Award (2007), published more than 200 publications, more than 20000 citations, has an h-index of 69 (Scopus), gave more than 250 invited/keynote talks and accomplished € 16 million research grant income since 2000.



Professor Vasilis Ntziachristos studied Electrical Engineering at Aristotle University in Thessaloniki, Greee and received his M.Sc. and Ph.D. from the Bioengineering Department of the University of Pennsylvania. He served as assistant professor and director of the Laboratory for BioOptics and Molecular Imaging at Harvard University and Massachusetts General Hospital. Currently, he is the Director of the Institute for Biological and Medical Imaging at the Helmholtz Zentrum in Munich, Germany, as well as a Professor of Electrical Engineering, Professor of Medicine and Chair for Biological Imaging at the Technical University Munich, Germany. His work focuses on novel innovative optical and optoacoustic imaging modalities for studying biological processes and diseases as well as the translation of these findings into the clinic.

Conflict of interest statement

The authors declare no conflict of interest