WEBINAR SUMMARY:

AI in Medical Practice

Talk by Dr Maki Sugimoto and Dr Stamatia (Matina) Giannarou

Tuesday 21st April 2026

Presentation by Dr Maki Sugimoto:

It is a great honor to be invited to give this lecture. Today, I would like to present my research work on image-guided surgery using virtual reality (VR), artificial intelligence (AI), robotics, and the metaverse.

In pre-operative conferences, spatial understanding in surgical planning is crucial. Have you ever considered how many hours surgeons actually spend on this? Imagine spending more than three or four hours per patient before surgery.

Surgical techniques often remain unsystematised, relying heavily on tacit knowledge, which makes their transmission and standardisation challenging. Surgeons frequently rely on 2D images and subjective expertise, which can lead to disparities in experience and skills. This is the so-called tacit knowledge in surgical technique that is not systematised.

Just as you would use a car GPS, surgical navigation helps surgeons navigate intricate anatomy, reducing risks and enhancing accuracy. You cannot use an old map to explore a new world. Therefore, we developed a surgical navigation system tailored to individual patients, operating much like a car GPS.

However, conventional 2D displays cannot fully convey the spatial positioning of organs, lesions, and instruments reconstructed from CT and MRI scans. Current single-modality imaging for pre-operative surgical planning and intraoperative guidance has limitations in clearly identifying the actual anatomy, which is crucial for surgeons to make correct decisions during intraoperative procedures. We traditionally use paper printouts from a patient's CT scan. Even though we have 3D data from CT scans, why do we only view it on 2D paper or 2D displays? The problem lies within the display itself. This limitation is why surgical techniques remain unsystematised and rely on tacit knowledge.

To bridge the gap between a surgeon's skill level and their spatial awareness, we developed a surgical guidance system using extended reality (XR). Extended reality encompasses virtual reality (VR), augmented reality (AR), and mixed reality (MR), combined with the metaverse.

We developed a web-based XR application named ‘Hollow Eyes’, of which I am the founder. From CT or MRI images, organs and lesions are extracted into colored 3D polygons. We represent these in a spatial and layered space using our original web-based XR application. By integrating a commercially available cloud service, we can generate these XR models in only five minutes. Furthermore, we have already received Class II medical device certification in Japan.

As a certified medical device, all a surgeon needs to do is upload the clinical images, and after five minutes, they can download the VR application to their head-mounted display. This cloud service automatically converts medical images captured by CT or MRI into VR, spatially reconstructing the anatomical structures.

We refer to this technology as extended reality (XR), which is an umbrella term for immersive technologies like virtual reality (VR), augmented reality (AR), and mixed reality (MR). Virtual reality places you completely inside an artificial, digital world. Augmented reality overlays a digital layer of contextual information onto the built environment, while mixed reality is an interactive blend of physical and digital environments. The difference is very simple: is the background a digital environment or the actual, physical environment?

While older generations might ask, "What is it?", younger generations ask, "What can I do with this?" The core of this technology is the stereoscopic view. We leverage side-by-side display technology on VR headsets and smartphones to enable stereoscopic vision of medical imaging for the left and right eyes. While conventional screens provide 2D vision, using two separate displays allows us to offer 3D vision through parallax technology. This approach ensures a highly immersive 3D experience by delivering distinct visual inputs to each eye.

With this technology, you can upload the patient's data, and within 10 to 15 minutes, you receive 3D polygons of each organ, including tumors, cancerous cells, and blood vessels. Today, we can use AI to automatically segment each organ. Using AI technology, you can select specific organs or lesions from the patient's data to rebuild their 3D polygons, which can then be exported for virtual reality applications.

After selecting the organs and areas of interest, you can review the 3D models suspended in the air. Currently, we use video pass-through head-mounted displays. Several cameras are located on the surface of the VR goggles, and these cameras project the real-world surroundings onto the virtual reality screens for both the left and right eyes. This video pass-through function allows you to see the actual environment through the virtual monitor. You can view the patient's CT data floating in front of you, and we can even overlay it onto a 3D-printed organ for training purposes.

This pass-through ability enables mixed reality experiences where virtual elements are overlaid directly on your physical environment. While augmented reality is typically used with smartphones or tablets, a head-mounted display allows surgeons to view the actual environment during live surgery, even while wearing sterile gloves. This is mixed reality, the merging of real and virtual worlds to produce new environments and visualisations where physical and digital objects coexist and interact in real time.

We applied holography as a 3D guide to organs, lesions, and procedures based on individual CT and MRI scans, creating a stereoscopic view in the space right above the surgical field. This allows viewing from all directions. Holograms are given presence in our world by remaining anchored in place as we move around the physical space. When we share these coordinate systems between devices, we can create a shared holographic experience where multiple users can participate in the same virtual environment. We have already published these findings in the Annals of Surgery, a journal with an impact factor of over nine.

Allowing users to move around and share these models enhances surgical planning by optimising layer selection, preventing anatomical misidentification, and minimising both unnecessary bleeding and miscommunication. Furthermore, by using intraoperative fluoroscopy, we can reconstruct the patient's actual state during surgery, bridging the knowledge gap between young and experienced doctors.

For example, in laparoscopic pancreatic surgery and robotic surgery for pancreatic cancer, we can display the patient data right in front of the surgeon. Using sterile gloves, surgeons can manipulate the visual orientation as desired. Today, we are also utilising the Apple Vision Pro to display patient data in the air. The Apple Vision Pro has highly detailed, high-resolution displays (4K for each eye), providing rich visual context to navigate our procedures.

During surgery, we can accurately reproduce the organs, arteries, veins, peripheral tissues, and tumors from CT scans as holograms, displaying the dissection line to guide surgical procedures in a sterile space above the patient's abdomen. We have applied this in several procedures: a partial cancer resection (to preserve organ function), a prostate cancer surgery using the da Vinci robotic system, a hysterectomy for uterine cancer using a laparoscope, and breast cancer surgery to detect lymph node metastasis. We can overlay this data directly onto the patient's breast to see the location of each metastatic lymph node highlighted in green.

This is also applicable for orthopedic surgeries, such as artificial joint replacements for hip joints, helping us select the correct orientation. In spine surgery, where we need to place several screws along the spine, it is traditionally very difficult to detect and select the correct location, position, and insertion angle. However, the holographic guide can project a green holographic bar directly in front of the patient to navigate the insertion line. All the surgeon has to do is insert the needle along this green holographic line, which simplifies the process significantly.

This tool is highly valuable for educational purposes and for training young doctors. It benefits not only surgeons but also internal medicine physicians who use ultrasound to guide needles into liver cancer to target and heat cancer cells. We can show the correct needle insertion path using the holographic guide to demonstrate the stereoscopic procedure.

This system is also available for endoscopic procedure guidance. By utilising fluoroscopic data of the biliary tree, we can project complex biliary reconstructions in the air during endoscopic surgery. Here, you can see an endoscopic procedure for a biliary cancer patient where a stent is being placed into an obstructed biliary tree.

The endoscopist can select the correct line, direction, and route for the insertion. We have published several articles on this.

Furthermore, in endovascular interventions for brain aneurysms, it is very easy to pre-shape the catheter to match the individual patient's anatomy. Using the holographic patient data, we can customise the catheter's shape for each patient. Here is an example of a transcatheter aortic valve replacement (TAVR) using a catheter. We can run a simulation before the surgery, and even during the procedure, we can align the correct location, position, and size of the stent for the patient.

As the saying goes, "Learning is experience; everything else is just information." We want to share this experience across distant locations. Using this spatial, immersive telesurgery training system with robotic endoscopy, we can share the movements and the voices of multiple doctors within a metaverse environment.

Through spatial immersive surgery training, live surgical skills and techniques can be shared within the metaverse and XR environments among several doctors. We can merge endoscopic 3D data and 3D reconstructed CT data in the metaverse, allowing young surgeons to learn by aligning their arm movements with the robotic system.

We can share this content on smartphones using low-cost tools like Google Cardboard (costing only $1 USD). Students can load this data onto their own devices, which makes educational sharing highly accessible, even for large classes of over 100 students. This BYOD (Bring Your Own Device) approach keeps the cost of virtual reality classroom education under $100.

Our system has already been adopted by more than 70 clinical institutions in Japan for various procedures, including open, laparoscopic, and robotic surgeries. It is used not only by surgeons but also by physicians for diagnostics and interventional therapies. We have published more than 200 medical papers, received over 30 national and international awards, and secured numerous research grants.

In summary, I believe this is the future of medical imaging and display. In the past, we relied on 2D imaging like X-rays. Currently, we view 3D reconstructed data on 2D displays. In the near future, viewing 3D holograms in 3D space will be the standard. This is the era of extended reality and the metaverse. These technologies are no longer just "nice-to-have"; they are "must-have" tools, acting as a crucial aid for patient treatment. Thank you very much for your attention.

Presentation by Dr Matina Giannarou:

Hello, I am Matina. The use of AI in medical practice has evolved significantly over the last decade. Artificial intelligence has moved beyond experimental pilots and is now embedded in several areas of routine clinical practice, particularly where pattern recognition, prediction, and automation add clear value.

One of the most mature areas is medical imaging. Numerous AI models are currently used in radiology and pathology to detect abnormalities, grade tumors, and quantify biomarkers. The advantage of AI here is that it reduces inter-observer variability and lightens the workload pressure.

Another major area is clinical decision support. We now have AI models that can predict and estimate the risk of conditions like sepsis or general clinical deterioration, enabling earlier intervention and proactive care. The aim of using AI in this context is to augment rather than replace the clinical judgment of medical teams.

AI is also increasingly used to streamline administrative and operational workflows. We have models that can automatically generate clinical documents or summarise clinical notes, reducing administrative burdens. Recently, we have also seen virtual assistants and chatbots supporting patient communication by answering common questions, providing medication reminders, and guiding patients through pre- or post-procedure instructions.

Furthermore, AI has reshaped surgery, shifting it from a discipline dominated by manual skill and experience to one augmented by data analytics and real-time decision support. Currently, image-guided and robotic systems have already improved the precision of minimally invasive surgery. However, AI adds a new layer of contextual intelligence. For example, some AI models can analyse intraoperative imaging and video data to assist surgeons in identifying anatomical landmarks, delineating critical boundaries, and thereby reducing intraoperative errors. AI also enables the objective assessment of surgical skills, which improves training and treatment quality. Recently, there has been increasing interest in developing AI models to predict surgical complications and errors, allowing for timely intervention and improved patient outcomes.

Focusing even further, AI has been used extensively in surgical oncology. The goal here is to analyse multi-source information to enable surgeons to make informed, patient-specific treatment decisions. We know that the ultimate aim of surgical oncology is to resect the entire tumor while sparing healthy tissue. In practice, however, this is challenging. Particularly in neurosurgery, it can be extremely difficult to distinguish where healthy tissue ends and the tumor begins.

Current intraoperative methods in surgical oncology have significant limitations and often cannot provide complete boundary detection. For example, intraoperative MRI disrupts the surgical workflow, while fluorescent imaging does not always highlight the entire cancerous region.

Recently, we have seen advances in biophotonics techniques, such as probe-based confocal laser endomicroscopy (PCLE), which allows for the visualisation of tissue at a microscopic scale during surgery. For instance, PCLE uses small, flexible probes applied directly to the tissue surface to reveal cellular morphology, helping surgeons make decisions at a microscopic scale.

At a macroscopic scale, we can reveal tissue characteristics using hyperspectral imaging. This is a rich imaging modality that captures information in both the spatial and spectral domains. Preliminary studies show that hyperspectral imaging can play a significant role in detecting tumour boundaries during oncological surgery.

By combining multiple imaging and sensing techniques, we can enable surgeons to make informed decisions and guide the surgery to dissect the entire cancerous area without causing iatrogenic injury. This is where my research focuses: I aim to develop a robotic platform that combines multiple imaging modalities to assist in surgical navigation and enable real-time tissue characterisation to improve the efficacy and safety of oncological surgery.

I conduct this research at the Hamlyn Centre for Robotic Surgery. The Hamlyn Centre was established to develop safe, effective, and accessible technologies in three main areas: imaging, sensing, and robotics. While we focus on technological innovation, we place a strong emphasis on clinical translation, aiming for direct patient benefit and global impact.

At the Hamlyn Centre, I lead the 'Cognitive Vision in Robotic Surgery' group. We develop computer vision and AI methods to assist surgical navigation, enable the autonomous execution of surgical tasks using robotic platforms, support online diagnosis, enhance intraoperative visualisation, and provide objective methods for surgical skill assessment.

In our project to develop the "Neuro-platform," we utilise cameras already present in the operating theater to reconstruct a 3D representation of the surgical scene. We use these anatomical structures to guide the scanning of tissue using robotic tools holding our imaging probes. By scanning the tissue, we capture characteristics that are analysed using AI models to characterise tissue types and assist in decisions regarding tumor resection.

Regarding intraoperative surgical navigation, we have developed numerous techniques to reconstruct depth and retrieve the 3D structure of surgical scenes. We have employed traditional computer vision methods as well as more modern deep learning techniques. We combine 3D reconstruction with temporal tissue tracking to estimate tissue deformation and scene dynamics.

In our work, this is integrated with a biomechanical model to estimate the forces applied to the tissue surface without relying on external force sensors. This is crucial in microsurgery and neurosurgery, where the safety threshold for applied force can be as low as 0.3 Newtons. Even for an experienced surgeon, exceeding this threshold is very easy. By providing real-time force alerts, our system can prevent iatrogenic injuries. In the video shown here, an aneurysm ruptured because the surgeon applied force exceeding the safe limit.

We have also proposed a concept called "see-through vision." In constrained surgical environments, tools often occlude significant portions of the surgical scene. To address this, we detect the occluding tool and use an AI model that learns to reconstruct the tissue contact and structures underneath the occlusion, presenting this view to the surgeon in real time. The goal is to prevent iatrogenic injury by allowing surgeons to see what lies beneath active surgical tools. In the middle video of a da Vinci robotic surgery, the far left panel shows the segmentation of the tool shafts, while the far right panel shows the reconstructed view with the tools completely removed.

Once we have reconstructed the 3D structure of the environment, the next step is to apply robotic tools for tissue scanning. The probes used in PCLE are very small and highly flexible. If we attempt to perform this scan manually, it is nearly impossible to capture stable, high-quality endomicroscopic information suitable for diagnosis. In our work, we attach these probes to a robot, achieving highly repeatable and systematic tissue scanning.

We developed AI models that regress the distance and orientation between the probe and the tissue. Our goal is to keep the probe at an optimal working distance and perpendicular to the tissue surface to capture clear endomicroscopic details. In the left video, you can see that when scanning begins, there is initially no endomicroscopic information, only a black screen or blurry images. The view becomes progressively clearer as our AI model guides the robotic system to position the probe optimally. The right video demonstrates the optimisation of the probe's orientation, where the best quality is achieved when the probe is perfectly perpendicular.

We have integrated these AI systems with a robotic platform, specifically using the SmartPod robot. During scanning, the AI model detects whether the probe is within the optimal working distance, and if it is not, the robotic system automatically rectifies its position. This auto-rectification ensures the probe remains at the optimal distance and orientation.

Once we capture these characteristics, the next step is to analyse the data to assist the surgeon's decision-making. Since PCLE is a relatively new technique, surgeons, who are not histopathologists, may find it difficult to interpret the endomicroscopic videos in real time. Therefore, we aim to develop AI techniques that automatically analyse this content.

We have developed models to classify the two most common types of brain tumors: meningioma (benign) and glioblastoma (malignant). Our models achieve an accuracy of approximately 95% when analysing single frames, and about 98% when analysing short video segments.

To demonstrate how challenging this classification is: the first row displays meningioma samples, and the third row displays glioblastoma samples. The second and fourth rows show misclassifications. For example, some glioblastoma samples in the second row look remarkably similar to genuine meningiomas and were incorrectly predicted as such, and vice versa.

However, surgeons do not want to use AI as a "black box," nor will they trust a system they do not understand. We must make the model's predictions explainable. In our work, we aim to ensure that the model focuses on clinically meaningful features to gain the trust of surgeons. For example, we know that meningioma is characterised by structures called psammoma bodies, while glioblastoma is characterised by hypercellularity. If the model highlights these specific features, surgeons will have more confidence in accepting its predictions.

Recently, we developed a method that can classify meningioma and glioblastoma in both ex vivo and in vivo data. At the same time, it visualises the attention maps to show the surgeon which areas the AI focused on to arrive at its decision. These highlighted attention areas coincide perfectly with psammoma bodies in meningioma cases and hypercellular areas in glioblastoma cases.

Regarding tissue characterisation at a macroscopic scale, we segment volumetric hyperspectral data. We have developed a method combining classical spectral classification with modern interactive segmentation. The advantage of interactive segmentation is that we can leverage the researcher's real-time expertise to improve fully automated methods and enable the classification of novel, unseen tissue types—which is highly valuable during active surgery.

To illustrate, the top-left panel shows raw data captured from the operating room. Inside the black ring at the bottom is the tumorous area. On the top-right image, using spectral information alone successfully identifies the tumor but also misclassifies many unrelated areas. The bottom-left and middle-bottom images display classifications using only spatial information; they focus on the boundaries inside the ring, behaving similarly to edge-detection algorithms. However, our proposed method (shown in the third bottom image) clearly outperforms both approaches by accurately detecting the entire tumorous region within the ring. Furthermore, we believe the highlighted area outside the ring corresponds to active tumor infiltration.

We have combined these two modalities: PCLE and hyperspectral/multispectral imaging, into a single multimodal imaging platform. By attaching a multispectral camera to the surgical microscope, we can observe the surgical field, detect and track the PCLE probe, and map the data from both modalities onto a single coordinate frame.

We have applied this method to analyse hyperspectral data to detect tumorous versus healthy tissues. In the top-right image, green indicates healthy tissue, and red indicates tumorous tissue. We also use the PCLE probe to locally scan the tumorous area and identify microscopic boundaries. In the middle-right panel, you can see the co-registered overlay from both modalities, highlighting the precise tumorous margins that need resection in real time.

Overall, our vision is to combine multiple intraoperative imaging and sensing technologies. This allows surgeons to view real-time overlays of exact tumor boundaries, leading to safer, more effective resections and reduced surgical workloads.

However, we still have a long way to go before achieving widespread clinical adoption of AI in surgery. We must ensure we have datasets that satisfy strict requirements in terms of size and quality, including data from underrepresented clinical classes to achieve robust generalisability across different institutions, geographies, and patient groups. Additionally, AI models must satisfy explainability and trust requirements from surgeons. We also need to overcome regulatory and validation challenges, and ensure we address ethical considerations regarding data privacy, consent, and transparency. Finally, we need to design AI models that integrate seamlessly into the standard surgical workflow.

So far, most developed systems remain task-specific tools operating under direct human supervision. The goal of this technology is not to diminish the surgeon's role, but to reframe it. Surgeons will act as supervisors of intelligent systems, making high-level clinical judgments while delegating specific execution or monitoring functions to the technology.

Over time, I believe AI will help ensure more consistent outcomes, reduce unwanted surgical variation, and enable complex procedures to be performed more accurately, safely, and reproducibly. This will be achieved by fostering a closer partnership between surgeons, AI, and robotic systems, updating surgical training to emphasise the understanding of AI outputs, and building closer collaborations between clinical teams, engineers, and data scientists.

I would like to extend a big thank you to my technical and clinical collaborators, my entire team, and our funders. Thank you very much.