SEMMELWEIS EGYETEM

DOKTORI ISKOLA

Ph.D. értekezések

3405.

MÓGA KRISTÓF JÁNOS

Vaszkuláris patofiziológia / atherosclerosis

című program

Programvezető: Dr. Prohászka Zoltán, egyetemi tanár

Témavezető: Dr. Ferencz Andrea, egyetemi docens

Dr. Haidegger Tamás, egyetemi tanár

AUGMENTED AND MIXED REALITY IN SURGERY

Ph.D. Thesis

Kristóf János Móga MD

Theoretical and Translational Medicine Division
Semmelweis University

Supervisor:	Andrea Ferencz, MD, DSc Prof. Tamás Haidegger, PhD
Official reviewers:	Prof. Péter Banczerowski, MD, DSc Gábor Kertész, PhD habil.
Head of the Complex Examination Committee:	Prof. László Harsányi, MD
Members of the Complex Examination Committee:	Prof. László Kóbori, MD, DSc Attila Zaránd, MD, PhD

Budapest
2025

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AI Artificial Intelligence

AR Augmented Reality

ATE Angular Trajectory Error

CAS/CIS Computer-Assisted/Integrated Surgery

CBCT Cone-Beam Computer Tomography

CC BY 4.0 Creative Commons 4.0. license

CT Computer Tomography

FDA U.S. Food and Drug Administration

FH Free Handed (vertebral procedure)

GRS Gertzbein–Robbins Scale

HMD Head-Mounted Display

ICG Indocyanine Green Fluorescence

IEC International Electrotechnical Commission

IGS Image-Guided Surgery

ISO International Organization for Standardization

LTE Linear Tip Error

MIS Minimal Invasive Surgery

MR Mixed Reality

MRI Magnetic Resonance Imaging

N/A Not Added

OT Operation Time

PS Pedicle Screw insertion

PSP Pedicle Screw Placement

RAS Robot-Assisted Surgery

SLR Systematic Literature Review

UEQ User Experience Questionnaire

VR Virtual Reality

XR Extended Reality

1. INTRODUCTION

1.1. Surgical navigation and guidance

In the past decades, computer technology-driven medical and surgical innovations have transformed modern healthcare. The advancements in medical diagnostic imaging, minimally invasive surgical methods and tools, bio-compatible implants, modern anesthesiology, and the new aspects of digital medical technology have benefited overall patient outcomes (1). The advantages include faster recovery and improved rates of patients' reintegration into society and the economy (2).

Surgical navigation systems have transformed the landscape of surgical procedures by significantly enhancing precision and safety across various specialities (3). These systems are designed to assist surgeons in planning, guiding, and monitoring operations in real-time, thus improving patient outcomes and minimising complications. Traditionally, the surgical field relied heavily on methods, such as anatomical landmarks and X-ray fluoroscopy, which have been foundational in practices like neurosurgery and orthopaedics. Earlier clinical applications included image-guided stereotaxy utilised for brain tumour resections in neurosurgery. Yet, these methods often lacked the immediate step-by-step feedback for procedural subtasks and adaptability to inclement conditions that modern navigation technologies provide (4).

Recent advancements in surgical navigation incorporate sophisticated imaging technologies (CT and MRI scans) and computational algorithms that facilitate real-time feedback during surgery. These technologies could enable surgeons to visualise critical structures dynamically, which is pivotal in minimising the risk of injury to healthy tissues (5). In orthopaedics and neurosurgery, Robot-Assisted Surgery (RAS) plays a vital role by ensuring the precise placement of prostheses, thus optimising alignment and improving long-term functional outcomes for patients. RAS is a cornerstone for minimally invasive surgical techniques (MIS) across various surgical specialities, enhancing both dexterity and visual feedback for surgeons, leading to improved patient outcomes and reduced recovery times (6, 7).

In general surgery, there is a growing trend towards integrating advanced imaging modalities, like Indocyanine Green (ICG) fluorescence, to enhance tumour visualisation and assess tissue viability in real time. These tools empower surgeons to perform more accurate tumour excisions while preserving healthy tissues. Additionally, the field of dentistry has been revolutionised by tools such as Cone Beam Computer Tomography (CBCT), which provides detailed three-dimensional imaging crucial for planning implant placements effectively. Alongside this, Computer-Aided or Integrated Surgery (CAS/CIS, where computer-based technology is implemented to plan, train and simulate or execute surgical procedures for higher precision and patient safety) and Image-Guided Therapy (IGT, where real-time medical imaging modalities are used to guide therapeutic interventions) are essential in navigating complex surgical cases with improved precision (5, 8).

1.2. Extended Reality: VR, AR, MR

With the advancement of information technology, especially regarding computing and processing capacities, graphical interfaces, and the decreased size of personal computers, immersive Virtual Reality (VR) has stepped ahead. From the late 20th century, this new form of Extended Reality, the computer-generated artificial reality, also permeated research. More complex aspects and different layers have been appearing, resulting in the need for further classification. As a pioneer, *Milgram* published his theory about the Reality-Virtuality continuum in 1995 (Figure 1) (9).

Based on *Milgram*, VR, AR, and MR concepts can be distinguished. The concepts of Virtual Reality (VR) and head-mounted displays (HMDs) are not necessarily strongly coupled, as the immersive display device is generally placed on the head of the viewer to exclude the real environment. The user perceives, senses, and interacts with a computer-generated virtual 3D environment. Based on the International Electrotechnical Commission (IEC), VR is a ”*simulation of the physical presence of the user in an environment produced with the help of a computer, enabling the user to interact with this environment*” (10).

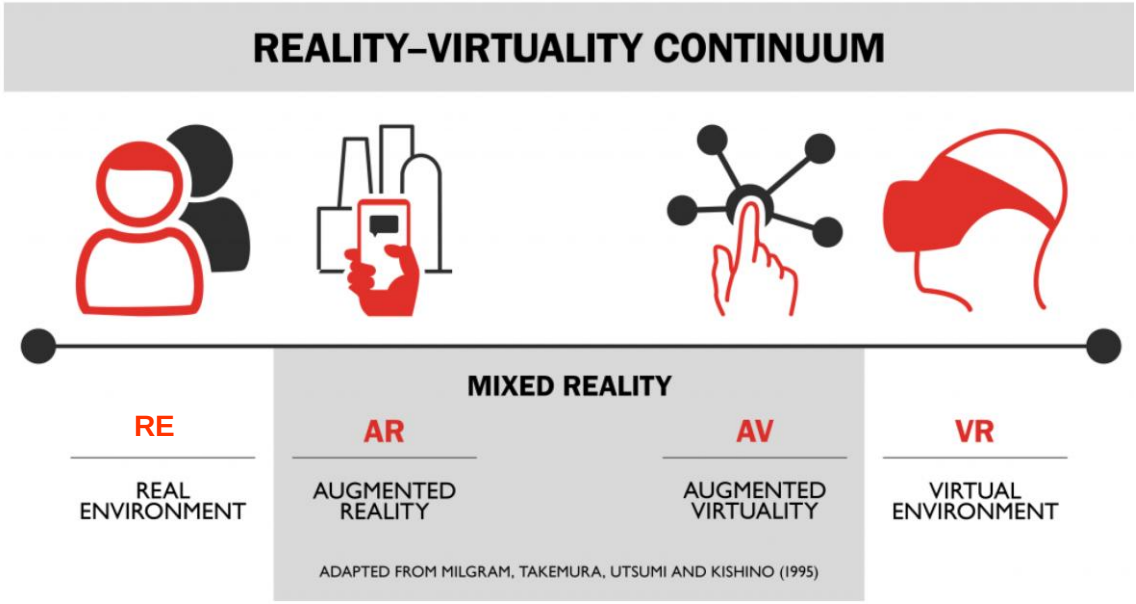


Figure 1. Reality-Virtuality continuum, based on Milgram, 1995

Source: Móga et al, 2021. (11) (© 2021 IEEE), initially found:(12).

Augmented Reality systems are defined by the International Organization for Standardization (ISO) as a "*type of mixed reality system in which virtual world data are embedded and/or registered with the representation of physical world data*" (13). This technology steps back a layer in the VR continuum with two major approaches of display methodologies: video-see-through (camera-captured images) and optical see-through (transparent surface) methods. Within the real world, additional augmentation of information is provided. According to *Schmalstieg et al.* (14), the display could be positioned in three localisations: head, body, and world.

The description of Mixed Reality (MR) may be confusing. According to *Milgram et al.* (9), it is a mixture of AR and VR hardware solutions to project generated virtual content into the real environment, which happens to behave and can be interacted with as tangible objects. Also, he defined the Augmented Virtuality class of displays, where "situations in which real objects, such as a user's hand, can be introduced into the otherwise principally graphic world, in order to point at, grab, or somehow otherwise manipulate something in the virtual scene" (9). Based on ISO: Mixed Reality is a "*system that uses a mixture of representations of physical world data and virtual world data as its presentation medium*" (13). The collective term of Extended Reality (XR) has been spreading recently, involving both VR, AR, and MR (15-17).

The popularity and spread of AR and MR technologies, as are Industry 4.0 technologies (18), are unquestionable and extend to medicine too (19). A wide variety of articles cover their medical-clinical applications, like general surgery, hepato-biliary surgery, robotic surgery, orthopaedics, traumatology, neuro and spine surgery, microsurgery, maxillofacial surgery, vascular surgery, pediatric surgery, obstetrics, and gynaecology, interventional radiology and cardiology. The applications focus on surgical trainee assessment, coaching and skill training, education, surgical team communication, telecommunication, and improving operating room efficiency, with navigation decreasing surgical errors and improving patient safety (20).

Different approaches and devices have been used for various medical specialities. Some of the most common HMD optics were diffractive and geometric waveguides, and for displaying techniques, LCD (liquid-crystal display), OLED (organic light-emitting diode), DLP (digital light processing), LCoS (liquid crystal on silicon), OLED-on-Si (OLED on silicon, micro-OLED), and mLED (micro light-emitting diode) could be identified. A prospective report in 2017 from *Dr. He* collected all the existing AR, VR, and MR display solutions, and forecasted their possible usages up to 2030 (12), which was updated and actualised in 2024 (21).

Based on the most recent articles, a list of devices used and the surgical approach's main specifications and features were collected in Table I and presented in Figure 2.

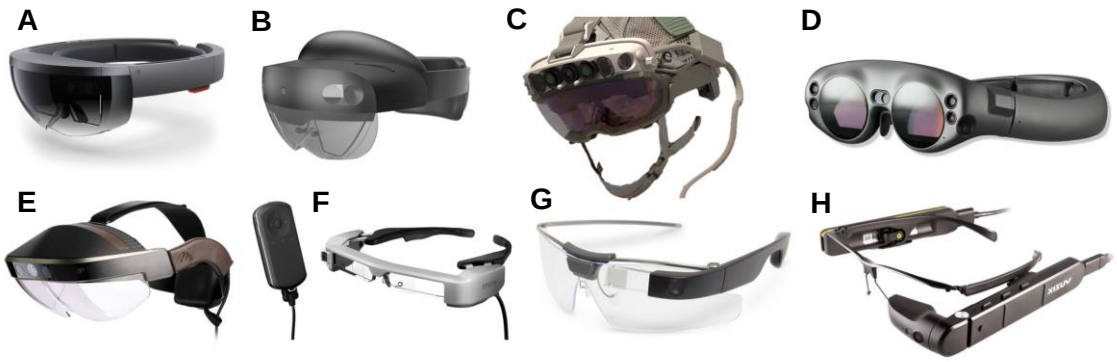


Figure 2. Medical AR/MR used devices

A - HoloLens, B - HoloLens 2, C - HoloLens IVAS CS3, D - Magic Leap 1, E - Meta 2, F - Epson Moverio BT300, G - Google Glass, H - Vizux M400.

Source: Móga et al, 2021. (11) (© 2021 IEEE).

Table I. Specifications of AR/MR devices mentioned in recent (2019-2025) medical articles

Source: The table was updated based on the published work in Móga et al, 2021. (11) (© 2021 IEEE).

Device name	Type	Release date	Field of view	Resolution / Display	Refresh rate	Sensors / Input	Battery life	Development status, price
Microsoft HoloLens (Model 1688)	MR	2016	34° (30°×17°)	2.3 Mp HD, 1280×720 per eye, holographic lenses (waveguides)	240 Hz 60 FPS	Hand, gesture, voice, eye tracking	2–3 hrs	Updated, $3000–5000 USD
Microsoft HoloLens 2 (Model 1855)	MR	2019	52° (43°×29°)	2k 3:2 light engines 1440×936 per eye (waveguides)	240 Hz 60 FPS	Hand, gesture, voice, eye tracking	2–3 hrs	$3500–4590 USD, discontinued 2024 (support till 2027)
Microsoft HoloLens IVAS CS3	MR (Military)	2022–2028	80°×40° (estimated)	n.a.	n.a.	Night vision, thermal, facial recognition, see-through walls	n.a.	In development (military, possible field triage)
Magic Leap 1	AR	2018	50° (40°×30°)	1.3 MP× per eye 1280×960, lightfield photonic display	120 Hz	Controller, hand, eye tracking, voice and haptic feedback	3.5 hrs	$2295–2995 USD
Magic Leap 2	AR	2022	~70° diagonal (≈45°×55°)	1440×1760 px per eye	120 Hz	Eye, hand, controller	3.5 hrs continuous 7 hrs sleep	$3299–4999 USD
META 2	AR	2016	90°	2560×1440 px LCD	60 Hz	Hand interactions, positional tracking	Power cable	Discontinued, $949–1495 USD
Epson Moverio BT300	AR	2020–2021	23° diagonal	Si-OLED, 16:9 1280×720	30 Hz	Controller	Power cable	$600–640 USD
Google Glass (2013–2017, v3.0)	AR	2013	13° (horizontal)	640×360 LCoS (prism projector)	n.a.	Touchpad	8 hrs	Discontinued, $1000 USD
Vuzix M400, M4000	AR	2019	16.8-28° (diagonal)	16:9 OLED 640×360 (waveguides)	n.a.	Touchpad, voice	2–12 hrs	$1000–2500 USD
Apple Vision Pro	MR	2024	~100°×73°	3660×3200 px per eye (dual OLED)	Up to 100 Hz	Eye, hand, voice tracking	2–2.5 hrs (external battery)	$3499 USD; used in surgery & training
Meta/Samsung XR Glasses (prototype)	AR	2025–26 (demo)	n.a.	Micro-displays AR	n.a.	n.a.	n.a.	Prototype; consumer release expected

1.3. AR/MR use in recent surgical literature

1.3.1. Education and skill training

Surgical training has always been linked to lifelong learning. With the 21st century's new technical possibilities, this practice and training are no longer limited to operating rooms and with live patients. Virtual Reality is applied to learn the basics, and AR/MR surgical simulators help young surgical trainees and residents become professional specialists.

The use of immersive displays with haptic feedback, with the first successful trials ongoing, provides experience for getting ready for the challenges of surgery (22).

1.3.2. Telemedicine, teleconsulting

A recent review described the positive aspects of AR/MR in the light of remote collaboration. AR/MR-based remote collaboration could be more natural than traditional camera-based ones, as intuitive interactions can be performed using this technology. Improved performance times were achieved, as the user experience is more engaging through shared non-verbal cues, emotions, gestures, empathy, and trust, which affect productivity and collaboration (23). Several collaborative working tool providers (like Microsoft and Facebook) have recently started to implement avatar-based communication support. Avatar-based communication in mixed reality environments enhances users' comfort in workplace communication (24).

1.3.3. Pre- and intra-operative use

Generally speaking, the most essential and engaging form of use of AR/MR in surgical practice is the planning phase and navigation through the surgical procedures. The essential step is acquiring a valid, reliable, and accurate 3D model from CT (Computer Tomography) or MRI (Magnetic Resonance Imaging) scans to add relevant anatomical or functional information to the static examination results, which helps 3D spatial understanding in open-, laparoscopic- or robot-assisted surgeries. Some research centres developed workflows for clinical use of the technique and real-time navigation, or deep learning and artificial intelligence for model segmentation (25-27).

1.4. Spine surgical uses of AR/MR

The technological revolution has also primarily influenced spine surgery (28). Annually, 266 million patients suffer from spinal diseases worldwide, which can significantly degrade their quality of life. Pedicle Screw Placement (PSP) is the most common orthopaedic procedure that is widely used for the stabilisation of potential (spinal stenosis

and degenerative diseases) and existing instabilities (pseudoarthrosis and post-laminectomy syndrome) of the vertebral column, spinal trauma and fractures, tumours and spinal deformities (scoliosis and kyphosis) (29).

The main steps of PSP in percutaneous screw fixation begin with the intra-operative posteroanterior and lateral fluoroscopy as image guidance. After planning the entry point of the vertebra, the pedicle entry site is decorticated, and, with the application of a curved or straight pedicle probe, a pathway is pre-formed for the screw through the cancellous bone. The pedicle screws are then inserted with extended tabs. Finally, the contoured fixation rod is inserted into the subfascial layer, and the screws are tightened on the rod (Figure 3 A–D) (30).

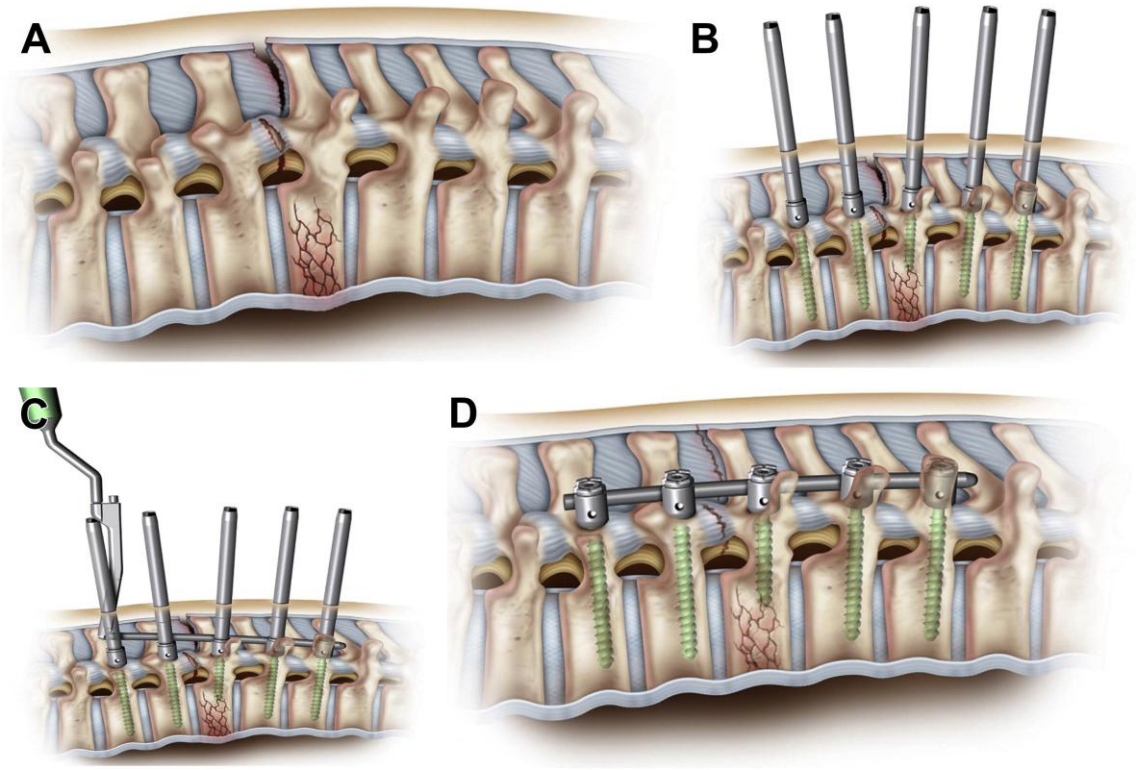


Figure 3. Steps for percutaneous pedicle screw fixation in the case of thoracolumbar fracture

(A) The patient is ready for the intra-operative posteroanterior and lateral fluoroscopy as image guidance. (B) Planning the entry point and a pathway is pre-formed for the screw through the cancellous bone. Pedicle screws are then inserted with extended tabs

(C)-(D) Contoured fixation rod is inserted into the subfascial layer, and the screws are tightened on the rod. Source: Móga et al, 2023. (31) (CC BY 4.0), initially found: (30).

Free-hand (FH) vertebral procedures are supported by fluoroscopic image-based navigation, aiming for precise operation next to the spinal cord. Fluoroscopy is time-consuming, presenting only 2D imaging information, and the surgical staff is exposed to radiation during the scan. Even with this utmost attention, a meta-analysis by *Staartjes et al*. discussed that 3.3% of FH spine fixation cases needed re-operation due to misplaced screw implantation. Preventing this would decrease patient morbidity and mortality, reducing the cost of care by $23,865–32,915 per revision surgery (32, 33).

In the past decade, different computer-integrated approaches have been invented to achieve minimal invasive PSP: robot-assisted surgery, computer-assisted navigation, image-guided surgery, and neurophysiological monitoring (34-36). More recently, the development of new spine surgery devices and other robotic services accelerated partially due to the complex effects of COVID-19 (37). Robotic spine surgery systems have been used for PSP since 2004, typically for positioning the screw-guiding tubes for manual insertion (38). The present PSP robotic systems can be divided into three different groups regarding their control strategies:

- the robot performs surgical sub-tasks or steps of the procedure autonomously in supervised mode,
- the surgeon remotely controls the robot,
- the surgeon simultaneously commands and controls the robot (39, 40).

The requirements for using RAS include high-quality preoperative imaging data, image–patient registration for navigation purposes (i.e., enabling anatomical landmarks or markers), and, if possible, real-time intra-operative tracking to monitor and compensate for additional patient movement (41).

The currently commercially available RAS systems display real-time patient and surgical tool movements using mechanical, electromagnetic or optical tracking. Though the cost of a RAS system is high, the studies report better clinical outcomes, more accurate screw insertions, fewer complications, and fewer revision surgeries, shorter operation times and smaller radiation dosages by employing robotic technologies (28, 42).

Computer-enhanced surgery provides many updates to the Operating Room (OR) workflow. Image-guided surgery incorporates the intra-operatively most relevant form of Extended Reality (XR): Augmented Reality (AR) for navigation in the field of spine surgery. ***The use of AR and MR navigation term is not coherent in the research field; in the following, both AR and AR/MR will be used as a collective term, including both XR solutions.***

The patient's preoperative CT/MRI diagnostic images or intra-operative cone-beam CT scans can be segmented automatically or semi-automatically into 3D models through image-computing platforms (e.g., 3D Slicer) (43-45). These models can be monitor-based, microscope-based or Head-Mounted Display (HMD) AR holograms. For image-patient registration, trackers are used together with anatomical or artificial landmarks (fiducials) (11).

Merging the two technologies, robotics and AR in surgery, can result in additional intraoperative and post-operative benefits, including patient accuracy and patient outcomes.

Among the many forms of XR, only a few specific modalities have found ways to test prototypes with clinical experts. The main advantage of using a see-through AR/MR-HMD navigation is the ability to project the holographic image into the surgeon's exact field of view without disrupting the workflow, as this would negatively impact the operation time, risking more prolonged anaesthesia. The effect of HMDs on the operation time (OT), the overall radiation dose, the PSP accuracy, and cost-benefit are not yet clarified, as only a few cases and heterogeneous data have been found in the literature. A previous systematic literature review by *Bursröm et al.* collected 28 studies focusing on AR in spine surgery in 2020, but included only limited clinical data sets (46).

1.5. Clinical Realism, Lack of Situation Awareness and Benchmarking

Defining clinical readiness or realism is not easy. Still, the US Military Health System gives one of the best definitions: "*Clinical practice that is relevant to a provider or provider team's ability to perform their assigned deployed role*" (47). In this dissertation, the clinical realism of use was put on a subject scale of 1–7, from theoretical use to routine

use for clinical or surgical application. The primary directive on this subject evaluation is how they can potentially be used from a specialist perspective.

1. Planning phase, virtual stage
2. Prototyping in research labs
3. Prototyping in production
4. Use in phantom models (organs, body cavities)
5. Use in cadaver/animal experiments
6. Human use for research purposes
7. Routine use in human clinical care

For surgically used HMDs, the most widespread one is Microsoft's HoloLens (first and second generations), as it is one of the first complex and versatile pioneers on the Mixed Reality device market, accessible relatively easily to date (48). Both examples are in the pre-clinical phantom model and day-to-day routine use. One of the principal explanations for its popularity is the wide variety of control use methods, such as gesture commands, voice commands or eye-tracking capabilities.

Next to it, Magic Leap is a similarly compound device with clinical relevance for future use in surgery. Epson Moverio, Apple Vision Pro and other manufacturers' prototypes, as a new possibility of AR-capable devices, should also be mentioned, as they might be popular in the research field. Three other widely used AR and MR display methods are monitor-, microscope- and projector-based; all require 3rd party devices and computers to be operated. As the technology is somewhat simpler (as they do not need to be ergonomically integrated and fitted on the head), broader examples of the use in neuro-, spine, and microsurgery are given. The collected and grouped screening results can be seen in Table II.

Table II. Identified AR/MR devices used in medical articles

Medical use has been proven. ** Clinical realism scale (1-7). * 0: clinically irrelevant; 1: small potential for clinical use; 2: great potential for clinical use.*

Source: The table was updated based on the published work in Móga et al, 2021. (11) (© 2021 IEEE).

DEVICE	USE CASES references	TECHNOLOGY READINESS LEVEL*	CLINICAL REALISM**	SUBJECT EVALUATION of valid clinical use***	IN USE CONTROL options	Situation Awareness, Benchmark
HoloLens (1st gen.)	General surgery, Da Vinci surgery, neuro and spine surgery	Medical use	4–7	2	Gesture	Spatial awareness, improved outcome, technical accuracy
HoloLens 2	General and orthopaedic surgery, medical training, 3D visualisation [multiple studies]	Medical use	6–7	2	Gesture, voice, eye tracking	Spatial awareness, improved navigation, widely validated; discontinued 2024, but in use
Magic Leap 1	General surgery, Da Vinci surgery	Medical use	6–7	2	Gesture	Spatial awareness, improved outcome
Magic Leap 2	Surgical planning, procedural training, neurosurgery exploration (enterprise trials)	Medical use (early adoption)	6–7	2	Gesture, controller, eye tracking	Enhanced contrast via dimming, better overlay accuracy, improved visualisation in OR
Apple Vision Pro	Laparoscopic and minimally invasive surgery (UC San Diego 2024), anatomy training	Medical use (pilot)	7	2	Eye, gesture, voice tracking	High-resolution overlays, hands-free use, high level of situational awareness
Epson Moverio	Neuro and spine surgery	Medical use	4	1	Controller, 3rd party device	Moderate accuracy, dependent on 3rd party integration
Monitor-based AR	Neuro and spine surgery	Medical use	4–6	1–2	3rd party device	High accuracy, acceptable time, reduced radiation, but difficulty in obese patients
Microscope-based AR	Microsurgery	Medical use	6	1–2	3rd party device	Reduced target registration error, improves anatomical accuracy
Projector-based AR	Neuro and spine surgery	Medical use	4–5	1	3rd party device	Improves anatomical orientation

The literature pre-screenings identified the lack of users' Situation Awareness (SA) research and specific benchmarking in the clinical-surgical usage of AR-HMDs.

Benchmarking is a process of identifying, collecting, and analysing data to determine and target future performance goals. The standardised AR/MR capable device is not yet granted in the surgical field. The aim of scaling and benchmarking AR-capable devices for medical use is not a novelty. A published paper in 2007 by *Behringer et al.* (49)

examined some usability issues of VR, AR and MR in the medical domain. The fusion of medical imaging data and AR/MR was described as possible support for surgery, diagnostics and medical training to improve the situation and spatial awareness of the medical staff. The user-centred (specialist, surgeon) development of devices was suggested, and a few examples were given to evaluate the usability of such applications, such as checklists of criteria for assessing design (50) or questionnaire audit (51).

While some general standardised tools have been applied to AR/MR systems, no widely adopted, domain-specific benchmarking checklist or questionnaire is designed exclusively for medical AR/MR devices. The requirement of presence and immersion of different virtual layers and elements is not to be questioned, as *Lombard* and *Ditton* defined: *"…the perceptual illusion of non-mediation [which] involves continuous ('realtime') responses of the human sensory, cognitive and effective processing systems*" (52).

Behringer et al. stated the required bullet points for the devices for extended usability, which can also be used in the medical field:

- be more accessible;
- usable for the non-technician end-users, who lack a deep knowledge of IT systems;
- follow the trends of pervasive and ubiquitous computing;
- implement the basic ideas of commonly used software (49).

More and more screened articles describe improved technical accuracy and outcomes within average operative time using AR-HMDs. Still, Situation Awareness testing and investigating methods are not standardised or given. It is well-known and discussed that target registration and navigation errors could lead to inaccuracies, especially in overweight patient populations (31).

This dissertation aims to collect and compare outcomes of pedicle screw placement operations with the most recent experience reported by studies exclusively on AR/MR-HMDs in spine surgery. The comparison highlighted that researchers have not commonly published standardised and validated data, used no standard methodology for measuring usability, or mentioned more than the fundamental aspects of situation awareness and benchmarking research.

2. OBJECTIVES

The following objectives were identified, aiming to answer the above-mentioned open research questions and gaps:

I. Usability, Situation Awareness and Benchmarking for Intraoperative AR/MR-HMDs

A) Collect and comprehensively analyse the clinical impact and usability of digital technology in recent spine surgery – Pedicle Screw Positioning (PS, PSP) studies using AR/MR-HMD technology for navigation and guidance.

B) Benchmarking situation awareness: highlight the importance of measuring and standardising SA in the research field, as it is still missing.

C) Offer a standardised scoring method with visualisation for comparing AR/MR-HMD surgical navigation and guidance tools.

D) Benchmarking AR/MR-HMDs: pointing out surgical benchmarks for future device research and development.

II. Workload analysis in the real clinical-surgical environment to prove the usability and clinical advantages of AR/MR-HMDs

My thesis work presents a use-case study to measure the effect of an AR/MR-HMD navigation and guidance system on users' workload and task load during its use in clinical/surgical environments.

A) Collect the overall reception and usability of end-users in the work environment.

B) Perform workload analysis to present the advantage of AR/MR-HMD navigation.

3. METHODS

3.1. Clinical Impact, Situation Awareness, Usability and Benchmarking

The research followed the PRISMA guidelines for systematic literature review (SLR) on the scientific topic of AR- and MR-guided spine surgeries (53). The systematic search was conducted for the last five years of PubMed, Scopus, Embase, Web of Science, IEEE Xplore, Cochrane Library and Google Scholar databases with the keywords „Augmented Reality" OR „Mixed Reality" AND „Spine Surgery". Inclusion criteria were defined as any article written in English, mentioning AR/MR navigation and guidance with Head-Mounted Displays in spine surgery, with case reports and accuracy data included. Next to human living patients, further analysis was performed on cadaver and phantom model studies, accessing a broader scope. Any reviews, systematic reviews, conference papers, abstracts, letters, opinions and papers describing educational, telemedicine and teleconsultation use of AR technology were excluded during the search. Duplications were removed with the Mendeley Desktop application (Mendeley Ltd., v.1.19.8).

3.1.1. Eligibility Criteria:

- Patients

Studies involving human patients with a minimum sample size of five patients or a minimum of five implanted screws were included. Articles discussing fewer interventions and cadaver and phantom model studies were separately analysed and referred to in the results section.

- Intervention

AR/MR-HMD navigation in PS spine surgery.

- Comparator

Any.

- Outcome

Outcomes comprise any clinical parameter (e.g., recovery rates, length of hospital stay, visual analogue scale and post-operative follow-up), accuracy data (linear tip error, angular trajectory error and Gertzbein–Robbins scale (54)), complications, procedural data (operating time and radiation dose) and user experience measured by any standard validated method.

- Setting

Both experimental and non-experimental settings.

- Publication Types

Articles written in English. Research papers describing teleconsultation, telemedicine and educational use of AR and MR technologies, as well as any abstracts, opinions, letters, reviews, SLRs, conference papers, and single case reports were excluded.

3.1.2. Screening:

Records identified during the search were screened in two steps. First, the hits were screened by their title and abstract by two reviewers (Kristóf Móga and Áron Hölgyesi) independently. Discussions solved disagreements, and a third researcher (Tamás Haidegger) was involved wherever needed. Second, the selected articles were downloaded in full text and screened by applying the eligibility criteria using the same independent review method.

3.1.3. Data Extraction:

Clinical, procedural and accuracy data from the included articles were collected, analysed, and narrated by the reviewers. Data extraction included the number of patients who underwent spine surgery, the number and accuracy of inserted pedicle screws, with accuracy data such as the Gertzbein–Robbins scale (GRS), linear tip error (LTE), angular trajectory error (ATE), operating time (OT) and surgical complications, such as the

occurrence of specific intra-operative and post-operative ones. Demographic data and the proportion of affected patients were analysed as well.
The user experience of surgeons and criticism about HMDs and AR/MR navigation were collected, highlighting the presence of situation awareness, technical challenges, and limitations of the different manufactured HMDs. As secondary data, where available, the region of spine operation (collar, thoracic, lumbar or sacral), the disease aetiology (trauma, compression fracture, and tumour) and the patient-reported outcomes were all gathered.

3.1.4. Analyses:

A descriptive analysis of the included studies was performed. Studies fulfilling our eligibility criteria in the systematic review by *Bursröm et al.* (46) were also analysed to assess whether those and the more recent studies were suitable for analysis. Multi-level Poisson and multi-level binomial models were used for this.
Statistical analysis and meta-analysis were performed with STATA 17 (StataCorp LLC, v.17). The SLR followed the Preferred Reporting Items on Systematic Reviews and Meta-Analysis (PRISMA) guidelines (53).

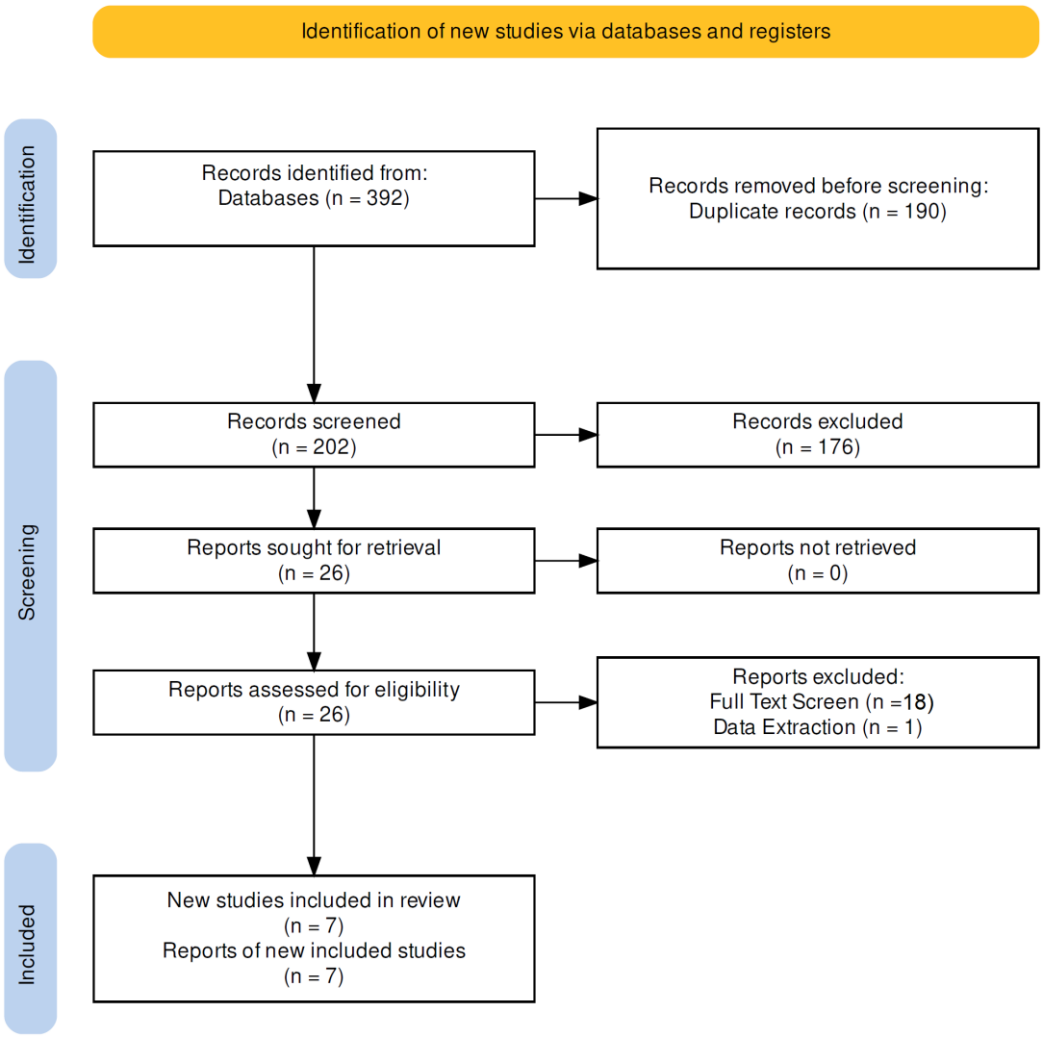


Figure 4: PRISMA chart of AR/MR guided Pedicle Screw implantation literature search

Source: Móga et al, 2023. (55) (CC BY 4.0).

Further screening search was performed in the literature with the inclusion criteria of Situation Awareness or Benchmarking (operating-room-specified, standardised AR/MR device development, production) mentioned on the topic of AR and MR used in surgery. The excluded articles' abstracts were re-screened to classify AR/MR display technology and devices. Then those that did not reveal information about the used ones with clinical relevance or use case control were excluded. The articles were reviewed with the help of the PRISMA methodology for transparency. The clinical realism of the use cases was scored with a subject scale of 1-7, with subject evaluation, taking into account how they can potentially be used in surgical environments (11):

1: Theoretical use, planning phase, virtual stage;

2: Prototyping in research labs;

3: Prototyping in production;

4: Use in phantom models - organs, body cavities;

5: Use in cadaver/animal experiments;

6: Human use for research purposes;

7: Routine use in Human Healthcare for surgical application.

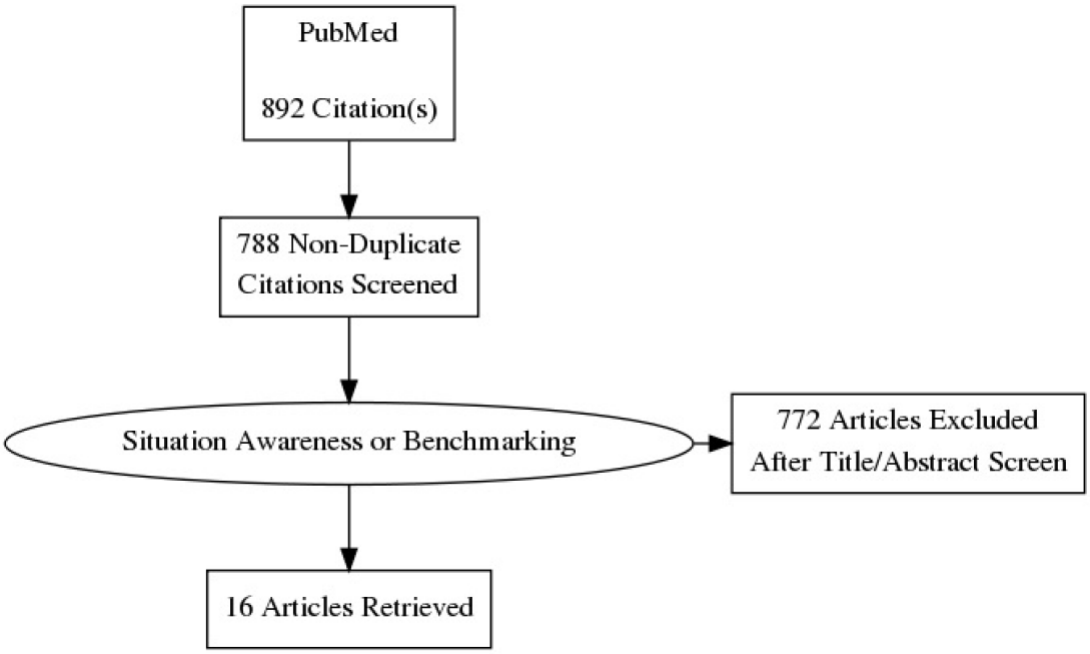


Figure 5: PRISMA chart of Situation Awareness or Benchmarking literature search

Source: Móga et al, 2021. (11) (© 2021 IEEE).

3.2. AR/MR-HMD usability and workload assessment of surgeons in Clinical Surgery

The usability and workload assessment anonymised data of AR/MR Head-Mounted Displays of surgeons were collected and analysed with the SURG-TLX (Surgery Task Load Index, based on NASA-Task Load Index) Assessment of Surgeon Workload standardised tool (56-59). The surgeon's feedback, HMD device evaluation and qualitative assessment of user experience were also gathered through the study. Altogether, 16 surgeons were involved in the study, from the ages of 25-65, subcategorised by their expertise in the surgical field as resident doctors (0-6 years), young specialists (7-15 years), and consultant surgeons (16+ years). In different preoperative and intraoperative surgical scenarios, HoloLens 2 HMD was used to project patient-specific 3D models onto the surgical workspace (operating room, table) for navigation and guidance purposes, generated from CT scans with Total Segmentation v2 AI (Artificial Intelligence model) tool in 3D Slicer (v.5.7.0) program (45, 60). A maximum level of effort of ethical considerations and criticism was given while using a non-medical device as a decision-making / helping tool through the surgeries with highly experienced surgeons as supervisors, paying attention to the utmost patient safety and care.

- Patients

  Surgeons with various experiences using AR/MR HMD for anonymised usability and workload assessment.

- Intervention

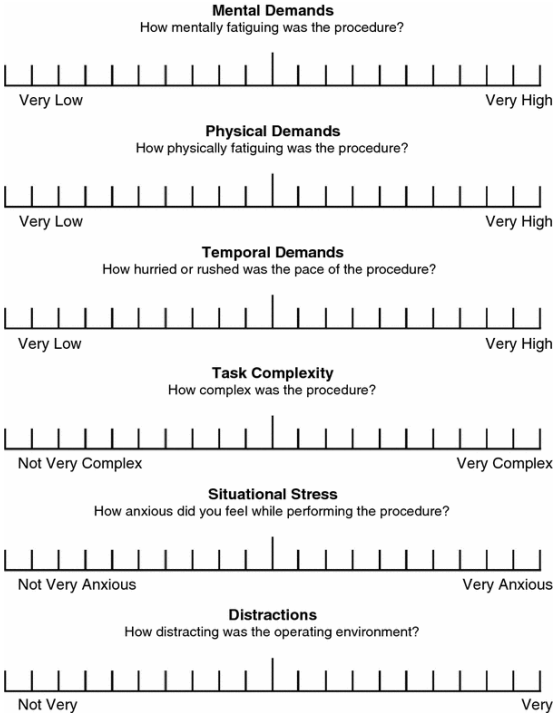
  HoloLens 2 HMD navigation and guidance for preoperative and intraoperative surgical scenarios.

- Comparator

  Any.

- Outcome

  Usability and workload assessment data collected and analysed with SURG-TLX. Any feedback, device evaluation and qualitative assessment of user experience were also gathered.



Task Complexity **Or** **Mental Demand**	**Distractions** **Or** **Situational Stress**	**Task Complexity** **Or** **Distractions**
Task Complexity **Or** **Temporal demand**	**Mental demand** **Or** **Situational Stress**	**Physical Demand** **Or** **Distractions**
Mental demand **Or** **Physical demand**	**Situational Stress** **Or** **Physical demand**	**Situational Stress** **Or** **Task Complexity**
Temporal demand **Or** **Mental demand**	**Distractions** **Or** **Mental demand**	**Temporal demand** **Or** **Distractions**
Physical demand **Or** **Temporal demand**	**Physical demand** **Or** **Task Complexity**	**Temporal demand** **Or** **Situational Stress**

Figure 6: Surgery-TLX sheets used for usability and workload assessment (59)

4. RESULTS

4.1. Usability, Situation Awareness and Benchmarking

4.1.1. Clinical Impact of AR/MR-HMDs in Spine Surgery

4.1.1.1. Pedicle Screw Accuracy and Surgical Outcomes

From the previous systematic literature review by *Bursröm et al.* (46), one article fulfilled the eligibility criteria (61). From that search time period, 392 publications were identified. With Mendeley Desktop (v.1.19.8) software, 190 duplicates were removed, and 202 records were screened by title and abstract. During this screening, 176 articles were excluded, as they were not eligible for the inclusion criteria. A full-text review was performed on the remaining 26 papers, where 19 publications were excluded, as they did not contain valid clinical outcome data on living patients. Altogether, seven articles were included in the analyses. Figure 4 contains the PRISMA flow diagram for the search and screening.

General Data and Patient Demographics

Table III contains the patients' demographic and general data. Overall, 272 patients (47% male and 53% female) underwent PS insertion using AR/MR-HMD navigation, with a weighted average age of 59.34 years. Patients' average body mass index (BMI) was reported as 30.34 kg/m^2. Only one article added information about patients' pre-operative conditions, such as the American Society of Anesthesiologists score (96.9% ASA 2 and 3) and the Charlson Comorbidity Index (CCI, 0.3 ± 0.5) (62). The disease aetiology and PS-insertion indications were the following:

Existing instability (pseudoarthrosis and post-laminectomy)	16 patients (6.15%)
Potential instability (Stenosis, Degenerative)	200 patients (76.92%)
Trauma, unstable fracture	8 patients (3.07%)
Tumour, infection	14 patients (5.38%)
Deformity	22 patients (8.46%)

Before the PS insertion, the most common to the rarest symptoms were:

- lumbar back pain,
- radicular pain,
- weakness of the limb,
- loss of sensation,
- and urinary retention.

The preoperative visual analogue scale (VAS) and Oswestry disability index (ODI) were measured as an average of VAS 6.7 ± 1.8 and ODI 82.7 ± 6.2. All the papers added information about the operation-performing surgeons: trauma/orthopaedics specialists or senior (chief) medical doctors. The pedicle screw fixations were performed on the lumbar, thoracic, thoracolumbar and sacral one vertebrae.

Table III: General and demographic data of patients with AR/MR-HMD navigation

Source: The table was updated based on the published work in Móga et al, 2023. (55) (CC BY 4.0).

Author	**Patients** Male / Female	**Age** Years, mean	**BMI** Mean	**Disease Etiology**	**Procedure**	**Surgeon**	**Spine Segment**	**Complication**
Li, J. et al. (63)	7 (N/A)	N/A	N/A	Fracture (n = 7, 100%)	PS, decompression, rod fixation	1 senior	Lumbar	0
Liu, A. et al. (64)	28 (11M/17F)	62.5 ± 13.8	31.8 ± 6.1	Degenerative (n = 12, 43%); Deformity (n = 12, 43%); Tumor (n = 3, 11%); Trauma (n = 1, 4%)	PS, osteotomy, discectomy, interbody placement, corpectomy, tumour resection	3 seniors	Thoracic (33%), Lumbar (55%), Sacral (13%)	0
Yahanda, A.T. et al. (65)	9 (5M/4F)	71.9 ± 11.5	27.3 ± 5.6	Spinal tumor (n = 4, 44.4%); Degenerative (n = 3, 33.3%); Deformity (n = 1, 11.1%); Deformity and infection (n = 1, 11.1%)	PS	1 specialist	Thoracic (50.8%), Lumbar (49.2%)	0
Bhatt, F.R. et al. (62)	32 (13M/19F)	50.9 ± 15.0	30.3 ± 4.9	Stenosis (n = 10, 31.3%); Post-laminectomy syndrome (n = 9, 28.1%); Deformity (n = 6, 18.8%); Instability (n = 5, 15.6%); Pseudarthrosis (n = 2, 6.3%)	PS and cortical screw insertion	3 specialists	Thoracic, Lumbar	0
Butler, A.J. et al. (66)	165 (83M/82F)	59.74	N/A	Degenerative (n = 156, 94.5%); Tumor (n = 6, 3.6%); Deformity (n = 3, 1.8%)	Transforaminal interbody fusion, lateral, anterior, combined lumbar interbody fusion, stabilisation	3 seniors	Lumbar (97.3%), Thoracic (2.7%)	0
Harel, R. et al. (67)	19 (8M/11F)	59.52 ± 12.49	26.98 ± 3.58	Spondylosis (n = 19, 100%)	PS	6 seniors	Lumbar, Sacral	0
Judy, B.F. et al. (68)	12 (N/A)	N/A	N/A	Deformity, degenerative disease, and tumour	S2AI screw placement	2 seniors	Sacral	0

Accuracy and Tracking

Remarkable data description heterogeneity was recognised through the literature review. A primary reason could be the lack of standardised requirements for data publication in the field of spine surgery or, more precisely, AR/MR-navigated spine surgery. On this basis, no information regarding LTA and ATE was identified in the analysed articles. However, the clinically widely used GRS classification score was presented in all of the screened papers based on intra-operative C-, O-arm or CT scans. GRS has five grades from A to E, based on the pedicle cortex breaching of the implanted screws (54):

- 'A' no breach detected in intrapedicular screw position;
- 'B' screw exceeding the pedicle cortex is a maximum of 2 mm;
- 'C' screw exceeding the pedicle cortex is 2–4 mm;
- 'D' screw exceeding the pedicle cortex is 4–6 mm;
- 'E' screw exceeds more than 6 mm or outside of the pedicle.

Only Grades A and B can be considered satisfactory operation results, as in cases C to E, mild-to-severe neurological symptoms could occur during the post-operative follow-up (69).

Table IV: Clinical accuracy of AR/MR-HMD navigated pedicle screw insertion

Source: The table was updated based on the published work in Móga et al, 2023. (55) (CC BY 4.0).

Author	Year	Use Case	Nr. Cases	Nr. Screws	Linear Tip Error (mm)	Angular Trajectory Error (°)	Gertzbein–Robbins Scale	Device
Li, J. et al. (63)	2021/03	in vivo	7	57	N/A	N/A	100% A	MITINS system (HoloLens)
Liu, A. et al. (64)	2021/10	in vivo	28	205	N/A	N/A	94% A, 4% B	xVision, Augmedics
Yahanda, A.T. et al. (65)	2021/08	in vivo	9	63	N/A	N/A	96.8% A, 3.2% B	xVision, Augmedics
Bhatt, F.R. et al. (62)	2022/01	in vivo	32	218	N/A	N/A	97.1% A and B *(4 misplaced screws revised intraoperatively)*	xVision, Augmedics
Butler, A.J. et al. (66)	2022/09	in vivo	165	606	N/A	N/A	99.51% A and B *(3 screws replaced intraoperatively)*	xVision, Augmedics
Harel, R. et al. (67)	2022/05	in vivo	19	86	N/A	N/A	97.7% A and B	xVision, Augmedics
Judy, B.F. et al. (68)	2023/01	in vivo	12	23	N/A	N/A	95.6% A and B	xVision, Augmedics

Table IV shows the clinical accuracy data of the reviewed publications. Altogether, 1258 pedicle and cortical screws were implanted into the 272 patients, an average of 4.6 screws/patient with a range of 2–15. The overall weighted average of GRS A and B was calculated as 98.69%. *Bhatt et al.* discussed 4, while *Butler et al.* discussed three misplaced screws out of their 218 (1.8%) and 606 (0.49%) placed ones through the PS implantation, which were identified, replaced and revised intra-operatively (62). Where the few GRS C or D PS grades were recognised through the post-operative follow-up (mainly in Lumbar 4 and 5 vertebrae), patients were 100% asymptomatic (64, 67).

Results from the **broader screenings** were also added and collected in Table V, with cadaver studies, phantom studies, and previously excluded in vivo studies. The 154 reported cases were found from the 14 articles in total: 34 phantom models, 17 cadavers, and 103 (103/154, 66.8%) living patients who underwent the PSP procedure. In these reports, altogether, 1189 screws were implanted, 849 (71.4%) of them into living or cadaveric thoracic, lumbar, and cervical vertebrae. Overall, 557 (557/1189, 46.8%) screws were inserted into 103 patients (31). The grouped average of LTE, ATE, and GRS A-B accuracy results was the following:

	LTE	**ATE**	**GRS A and B**
In vivo	2.78 mm (±2.00)	4.85° (±2.80)	97.1-100%
Cadaver	3.96 mm (±3.15)	5.02° (±3.79)	96–97.5%
Phantom	1.52 mm (±0.57)	ATE: 4.92° (±6.58)	94–98.4%

In vivo LTE and ATE were mentioned in only a few articles, but no intra- or postoperative surgical complications were described in the living cases; the detected misplaced screws were revised during surgery, and no secondary revision surgery was needed.

Table V: Publication list of search results for AR/MR-HMD navigated pedicle screw insertions

Source: The table was updated based on the published work in Móga et al, 2023. (31) (CC BY 4.0).

Author	Year	Use Case	Cases (nr.)	Screws (nr.)	Linear Tip Error (mm)	Angular Trajectory Error (°)	Gertzbein–Robbins Scale	Device
Molina, C.A. et al. (70)	2021/03	in vivo	1	6	2.07 (1.62-2.52)	2.41 (1.57-3.25)	100% A	xVision Augmedics
Li, J. et al. (63)	2021/03	in vivo	7	57	N/A	N/A	100% A	MITINS (HoloLens)
Liu, A. et al. (64)	2021/10	in vivo	28	205	N/A	N/A	94% A, 4% B	xVision Augmedics
Dennler, C. et al. (71)	2021/05	in vivo	25	N/A	N/A	N/A	„increased intraoperative accuracy”	HoloLens
Yahanda, A.T. et al. (65)	2021/08	in vivo	9	63	N/A	N/A	96.8% A, 3.2% B	xVision Augmedics
Farshad, M. et al. (72)	2021/05	in vivo	1	4	3.5±1.9	7.3±3.6	N/A	HoloLens 2
Bhatt, F.R. et al. (62)	2022/01	in vivo	32	218	N/A	N/A	97.1% A and B *(4 screws revised intraoperatively)*	xVision Augmedics
Felix, B. et al. (73)	2022/02	cadaver	7	124	1.9	2.4	96% A and B	HoloLens 2
Spirig, J.M. et al. (74)	2021/12	cadaver	2	8	5.99±3.60	5.88±3.69	N/A	HoloLens
Farshad, M. et al. (75)	2021/01	cadaver	8	160	4±2.7	6.8±3.9	97.5% A and B	HoloLens 2
Dennler, C. et al. (76)	2021/02	phantom	1	40	N/A	10.8±11.77	97.5% A	HoloLens
Frisk, H. et al. (77)	2022/01	phantom	4	48	1.4±0.8	3.0±1.4	94% A and B	MagicLeap
Yanni, D.S. et al. (78)	2021/08	phantom	24	192	N/A	5	98.4% A and B	SpineAR (prototype)
Urakov, T.M. (79)	2020/12	phantom	5	60	N/A	N/A	97% A	Caduceus AR

HMDs and Registration

According to the screened publications, in 6 out of 7 centres (85.7%), the xVision Spine AR navigation system (Augmedics Ltd., Chicago, IL, USA) was used, which is approved by the U.S. Food and Drug Administration (FDA) (80). One centre used an MR-based intra-operative three-dimensional image-guided navigation system (MITINS), including HoloLens by Microsoft (Redmond, WA, USA) (81). xVision uses registration clamps and infrared light-reflecting optical markers for image–patient registration, while MITNIS is based on electromagnetic tracking and navigation. No significant difference in clinical accuracy was recognised between the two systems.

From the **wider screened** and reviewed articles (cadaver studies, phantom studies, and previously excluded in vivo studies), also no significant accuracy difference could be found between HoloLens-based and xVision navigation systems (97.5–100% vs. 97.1–100% GRS A–B). For patient–image registration, sterile metal marker fiducials, registration clamps, optical/infrared reflective markers, and electromagnetic sensors were used in the screened cases. Registration accuracy and location errors were not reported in any cases (31).

4.1.1.2. Operating Time and Radiation Exposure

Only *Bhatt et al.* presented specific data about operation time (OT), averaging 3.6 ± 1.7 h. They added information on intra-operative blood loss of 224.0 ± 332.5 ml and the use of a mean of 3 packets of blood transfusion in their 32 patients. The length of hospital stay of their patients was 4.1 ± 1.6 days (62).

Specifically for PS insertion, *Butler et al.* added information about the average placement time of 3 min and 54 s per screw with a median of 4 min and 8 s (from 1 min 10 s to 6 min 30 s). They also measured the learning curve through their data collection, differentiating the experience of the first and the final 20 patients' procedures, where the mean insertion times were 4 min 1 s and 3 min 52 s per screw ($p = 0.48$) (66).

Bhatt et al. also added the mean total 3D imaging radiation dose for AR-navigated PS implantation, which was 576.8 ± 368.8 mGycm2, and the average fluoroscopy time was 25.7 ± 29.8 s with a mean radiation dose of 0.3 ± 0.4 mGym2 (62).

4.1.1.3. Complication Rates and Patient Safety, comprehensive analysis

None of the seven in-vivo study articles presented any surgical complications or the need for revision surgeries through the hospital stay or post-operative follow-up (from 2 weeks to 24 months). The clinical symptomatic reduction was noted in all patients through questionnaires by *Yahanda et al.*, as well as a significant improvement of VAS (18.4 ± 2.9) and ODI (16.4 ± 2.6) scores by *Li et al*. (63, 65).

From the previously published SLR by *Burström et al.*, only one article was identified as eligible for our inclusion criteria. *Abe et al*. published a cohort study on their vertebroplasty experience guided by the Epson Movero AR-headset in 2013 (46, 61). Altogether, five osteoporotic vertebra-fractured (Th. 12–L3) patients were operated on without any pedicular breach (100% GRS A), having only 2.09° ± 1.3° axial and 1.98° ± 1.8° sagittal trajectory error of the 10 implanted screws. No complication was identified through their follow-up.

As there were no complications described in either of the in vivo investigated articles, a statistical meta-analysis could not be performed. For further analysis, a multi-level Poisson model was used for the hypothesis, which measures and explains the incidence rate of "screw error odds" for every single screw insertion (82). The confidence intervals were calculated using the exact Poisson method (83). The examinations were made according to two scenarios, highlighting only GRS A class and GRS A and B rates together, as those are still clinically completely acceptable. Figure 7 contains the results of the analysis.

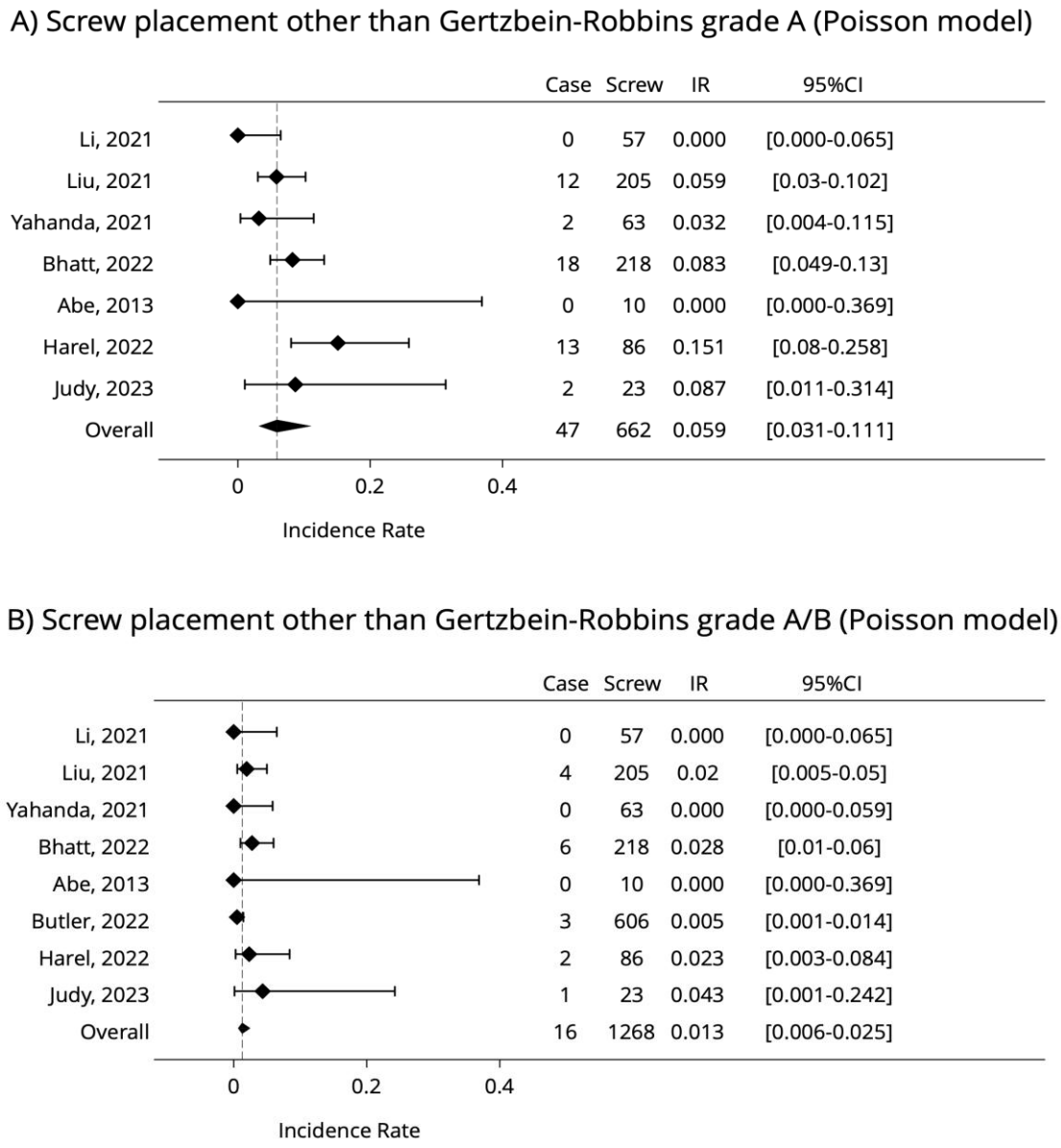


Figure 7: Outcomes of multi-level Poisson model analysis on the incidence rate of "screw error odds" (the lower, the better)

A: PSP other than GRS grade A.

B: PSP other than GRS grade A and B.

Source: Móga et al, 2023. (55) (CC BY 4.0), (61-68).

A multi-level binomial model was used in the ideology to measure the "screw error odds" for every single new screw insertion as a further attempt at error. The confidence intervals were calculated using the binomial (Clopper–Pearson) method (84). The same two scenarios were used as given above; Figure 8 shows the analysis results.

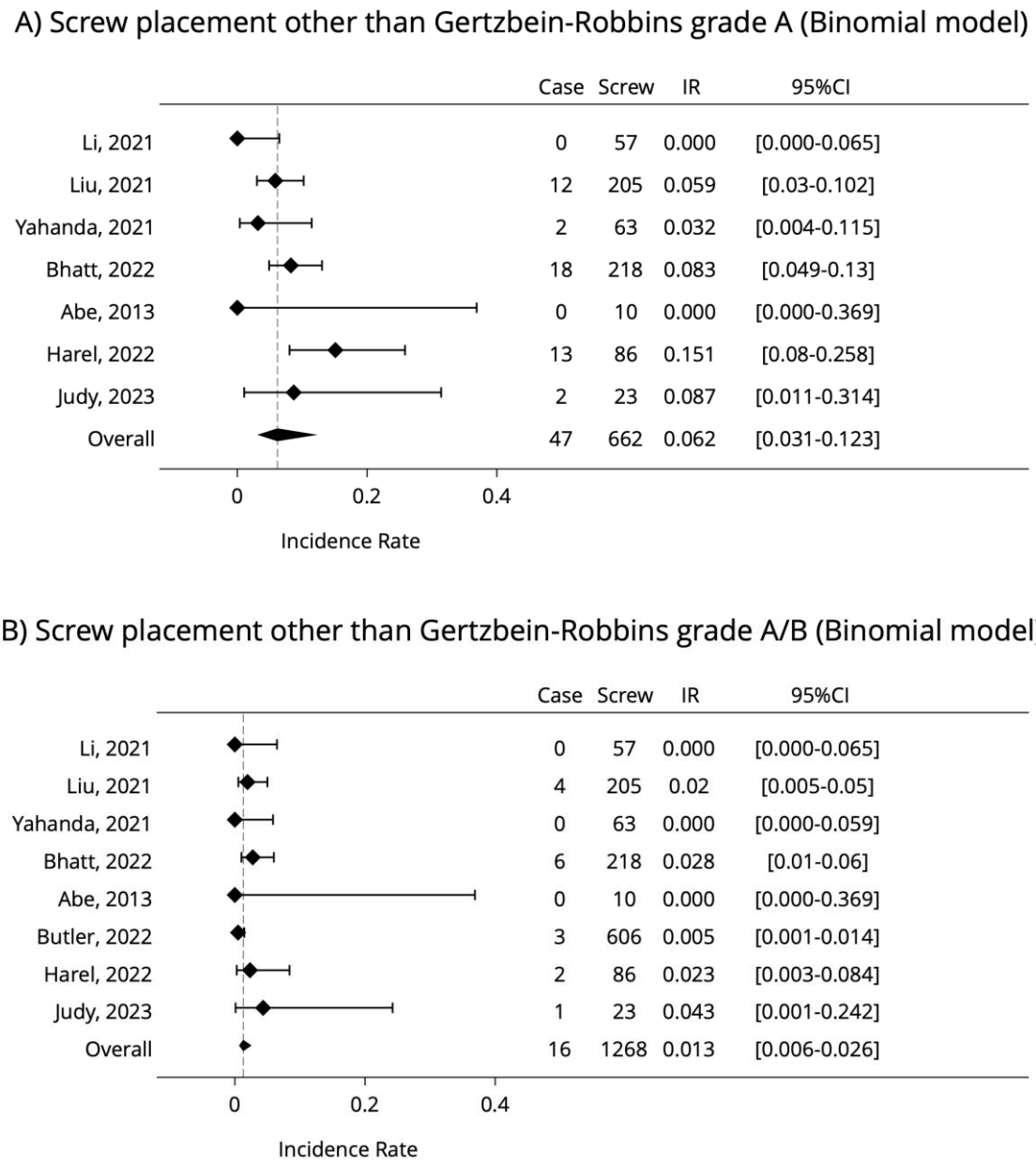


Figure 8: Outcomes of multi-level binomial (Clopper – Pearson) analysis method on the incidence rate of "screw error odds" (the lower, the better)

A: PSP other than GRS grade A.

B: PSP other than GRS grade A and B.

Source: Móga et al, 2023. (55) (CC BY 4.0), (61-68).

In conclusion, non-GRS A may occur in 6% (95%CI: 3–12%), non-GRS A and B (clinically unacceptable grades) may occur in 1.3% (95%CI: 0.6–2.6%) of all the screw implantations.

4.1.2. Situation Awareness, Usability

From the screened and grouped articles, a few precedents could only be found with some aspects of mentioning situation awareness. However, this is not a common topic to scale and determine the use of AR/MR in the medical (surgical) field. The definition of SA comes from *Endsley*, who described it as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (85). The theory was interpolated and applied to surgery as well by *Yule et al.* in 2006: "developing and maintaining a dynamic awareness of the situation in the operating room, based on assembling data from the environment (patient, team, time, displays, equipment), understanding what they mean, and thinking ahead about what may happen next" (86). SA can be measured by the surgeon's subjective assessments, observations, and interview techniques, including mental workload, stress, and communication (87). Methods can be applied, such as freeze probe technique, real-time probe technique, questionnaire, observation by an expert or individual operator, collecting psycho-physiological signs (eye tracking, EEG, cardiac activity, skin conductance, heart rate and heart rate variability) and behavioural measures (88).

It was clear from the screening that the standard measurements, data collections, and testing frameworks are missing for SA research in the application of AR/MR-HMDs in surgery or comparing any navigation systems, such as FH or RAS techniques. The methods mentioned above can offer objective insights into how AR/MR-HMDs affect cognitive load and environmental awareness through the procedures. Integrating SA testing frameworks will allow for better benchmarking of navigation systems beyond traditional accuracy and procedural metrics, ultimately leading to safer, more effective surgical workflows.

Despite its critical role in intraoperative decision-making and safety, most current publications rely solely on subjective feedback or qualitative expressions of enhanced situational understanding without applying validated assessment tools. A standardised data collection and publication methodology, including SA data analysis, is required in the future.

The following user observations and qualitative comments were added regarding the use of AR/MR-HMD navigation and guidance in SPS use cases:

4.1.2.1. Advantages

All the authors who added their criticism about AR/MR-HMDs agreed on the advantages and benefits of the operative outcomes while using the technology. Certain authors have also mentioned lower procedural costs and optimised ergonomics utilising the system (70, 78).

In the living patient group, *Molina* and *Dennler et al.* discussed increased accuracy and precision with no attention shift to remote screens and reduced radiation exposure (70, 71).

Yahanda et al. mentioned similar or superior accuracy compared to the most commonly used robotic systems (e.g., SpineAssist platform by Mazor, ExcelsiusGPS, ROSA or TianJi). He also added that the fluoroscopy time and radiation dose decreased through the surgeries (65).

According to *Bhatt et al.*, the technology is highly effective in real-world patient-care scenarios without needing a significant learning curve compared to robotic guided systems, where accuracy ranges from 97.9% to 100%, thereby suggesting that AR technology is a compelling alternative. He also added that, with just a limited disruption in workflow, AR-HMDs are simple to integrate, and with it, the implantation accuracy is elevated, and the overall radiation dose is decreased through the procedures compared with the FH technique (62).

Liu et al. added their accuracy and surgical workflow data to highlight the similarities with RASs (Mazor X, ROSA, TianJi). It was noted as a substantial benefit that using AR-HMDs minimises the attention shift, as the user can simultaneously visualise the operation field and the image guidance. With this, cognitive and motor task performances are increased. Another comment was that any instrument can be universally navigated with the AR-HMD system, causing only a minimum interruption in the line of sight. Additionally, the cost of technology is not prohibitive in patient distribution (64).

Harel et al. described the collected data on their user experience questionnaire (UEQ, 1–7 numbered scale on 26 clinical usability questions) regarding the xVision AR-HMD

system. All the scores were above six on average regarding clarity of navigation display, fit into the surgical workflow, and system reliability. The lowest score they noted was about the HMD ergonomics (5.9-point average) (67).

Regarding cadaver experience, *Farshad et al.* mentioned that surgeon satisfaction with AR-based screws and rod navigation was 5.38 ± 0.67 (on a scale of 1 to 6 in their questionnaire, where 6 was the best score) (75).

Using AR navigation on spine phantom models, the user experience corresponded with other groups' findings, and the published interpolating results showed that patient safety and outcomes could be increased with this technology (78). Table VI collects and displays the comments of authors.

Table VI: Mentioned advantages and limitations of using intraoperative AR/MR–HMD navigation in PSP

(+ positive feedback, - negative feedback, 0 limited to no effect)

Source: The table was created based on data published in Móga et al, 2023. (55) (CC BY 4.0). (61-68)

ADVANTAGES and LIMITATIONS	Accuracy Improvement	Attention Shift Decrement	Lower Radiation Exposure	Learning Curve	Workflow	Cost	Ergonomy	Visual Obstruction, Discomfort
Molina et al.	+	+	+			+	+	-
Denler et al.	+	+	+					
Yahanda et al.	+		+	-				-
Bhatt et al.	+		+	+	0			
Liu et al.	+	+			+	+	-	-
Harel et al.					+		+	+
Li et al.				0/-		-		-

4.1.2.2. Limitations

Only a few comments were provided about the disadvantages or difficulties of using the HMDs in living patient care. *Molina et al.* mentioned visual obstruction caused by the translucent data and experienced visual discomfort for one surgeon with the interpupillary distance calibration (70).

Li et al. drew up the cost of AR-HMD systems compared to traditional FH/MIS methods. Additionally, they felt the disturbing defects of soft tissue simulation, the contrast of AR models, and images were limited by bright light, and through their use, eye strain and visual discomfort may occur in the user, which requires training. For patients, unusually different lying postures were necessary on the OR table because of the use of EM trackers and navigation tools (63).

Liu et al. commented on the limited use of the technology on obese patients (as the image–patient registration marker clamps needed to be fixed on rigid locations were too short, causing four patients to be excluded from the use case study). The disadvantages of the HMDs were described as mechanical and visual discomfort, visual obstruction and sensory overload, and lastly, the prevention of using headlights through the procedures (64).

Yahanda et al. gave the same notes about the learning curve of visual discomfort and disorientation of HMDs and the difficulties caused by the patient's obesity (65).

4.1.2.3. Standardised measurement and visualisation

From all of the above, the main fields of advancement using AR/MR-HMDs compared to FH are the increased accuracy of PSP, decreased radiation time and exposure without disrupting operation time significantly with a more affordable price level, compared to RAS. As a new surgical navigation and guidance technology, it can be described as fulfilling technical, functional, and clinical requirements. The three attribution dimensions can be visualised in a 3D plot chart, giving end-users valuable and standard information, as seen below (Figure 9).

Accuracy can be measured, as previously highlighted, with standardised GRS (incidence of A, or A and B), or with the more precise and sophisticated ATE and LTE, and compared widely with different spine surgical navigation and guidance techniques. In broader clinical surgery, accuracy can also be extensively enumerated, such as the image-patient registration difference or errors and how the patient-specific virtual 3D organ models accurately represent the localisation and inner structure of the operating field and body. The acceptable boundary in the literature is under 4 mm, preferably under 2 mm, as higher inaccuracy could lead to severe organ, tissue, and body structure damage, leading to surgical complications.

In spine surgery and other traumatic, orthopaedic surgeries, the secondary main factor is the exposure to X-ray radiation (fluoroscopy) to the patient and the surgical staff. It is clear that the shorter this burden to the body is, the safer it is in the long term, lowering the risks of DNA structural harm and the risk of developing malignancies. This radiation dose needs to be collected, published and analysed in the broader surgical practice where applied.

Thirdly, operating time (OT) is also an essential and easy-to-follow, collect, and analyse data for end-users (surgical staff) concerning the newly implemented technology. It can be affected by a decrease or an increase in OT compared to conventional surgical approaches. Significance of +15% of cut-off time-value in minutes could also be disadvantageous to patient care (e.g., more prolonged anaesthesia involves a higher risk of medical complications) (31).

Presenting the three dimensions on a 3D chart makes an easy-to-understand and visually comforting standardised measurement tool compared to raw tables.

- Clinical Accuracy
    - GRS A, AB percentages, ATE, LTE scores, or
    - image-patient registration errors in mm-s
- Radiation Exposure (mGy/m2)
- Operation Time (minutes, ± 15%)

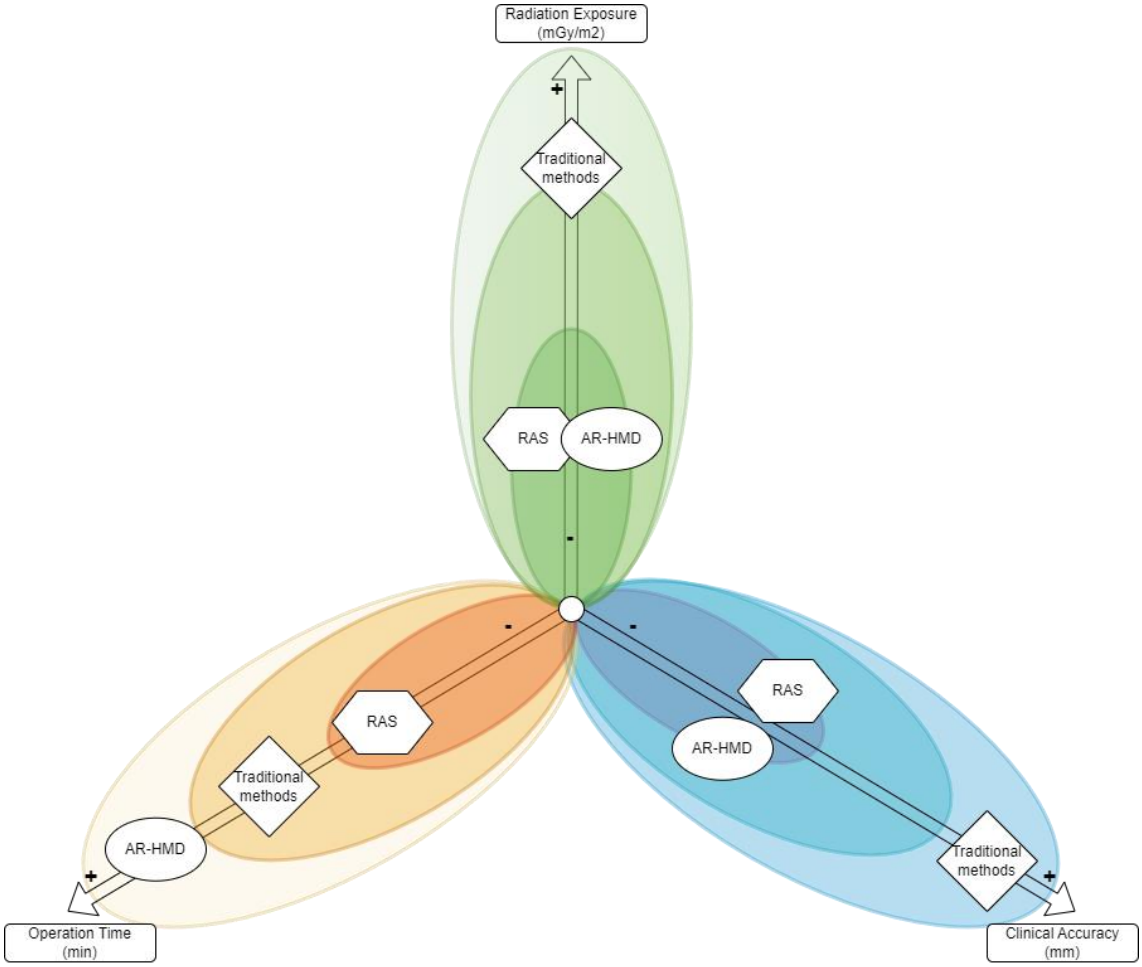


Figure 9: 3D plot chart presenting clinical accuracy, radiation dose and operation time of traditional FH, RAS and AR/MR-HMD navigated PSP

Source: The figure was created based on data published in Móga et al, 2023. (31) (CC BY 4.0).

Adding measured feasibility, usability, and User Satisfaction Scores (USS) as standard indexes (in a range of 0 to 1) about the used device's ergonomics, immersion, control, and interface will make a new visualisable dimension (Figure 10). Severe qualitative and quantitative tools, such as the Questionnaire for User Interaction Satisfaction (QUIS) or the Smart Glasses User Satisfaction (SGUS) questionnaire, can be used. A widely used one is the modified NASA-Task Load Index (NASA-TLX) implemented in surgical environments, namely SURG-TLX, measuring the impact and range of various stressors in the OR as a newly implemented device (58, 89-92).

From the perspective of the surgeon as an end-user, the ideal scoring tool needs to involve the following (see recommended calculation in the Supplement):

- Overall USS (0-1, mean score)
    - Ergonomics: weight and comfort in short (<60min) and long-term (>60min) usage (overheating, sweating, humidity as distracting factors).
    - Field of View (FoV): lack of visual obstruction, discomfort, no need for attention shift.
    - Image Quality: accuracy of image, contrasted visibility and immersion.
    - Control: precision and accuracy of gaze/motion/voice commands.
    - Usability, Setup: experience with the user interface from start-up, image-patient registration, and navigation.
    - Future-proof application.

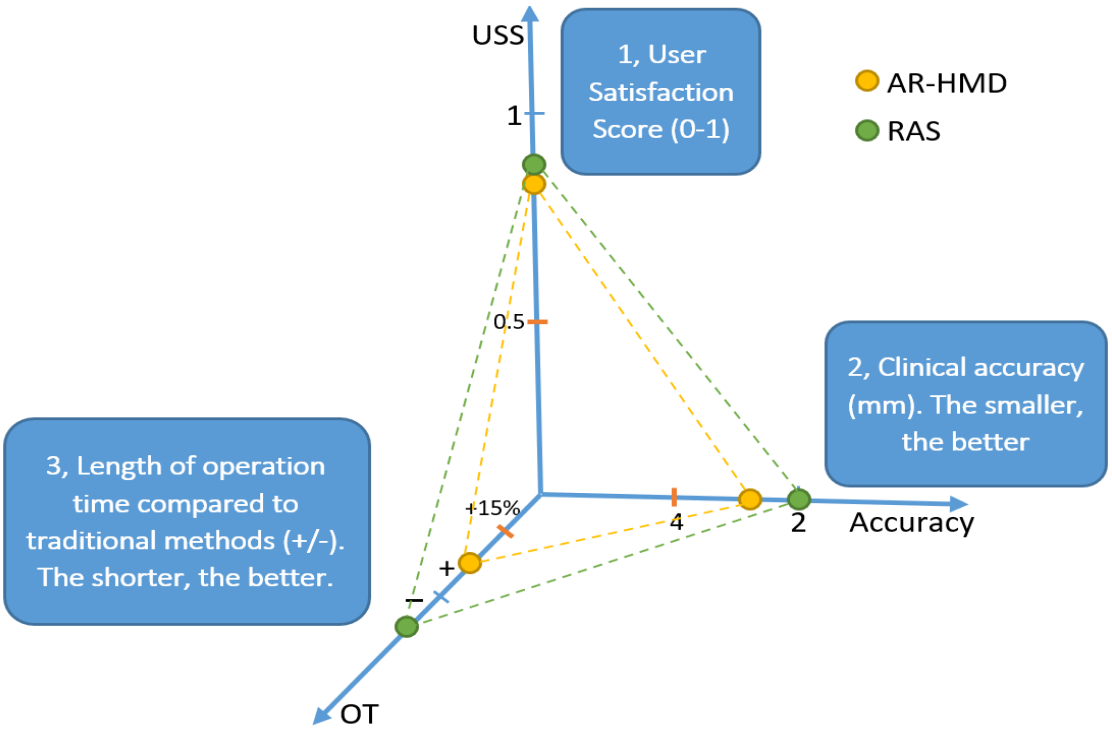


Figure 10: Visualisation example of the main attributes with minimum requirements of AR/MR-HMDs and RAS in spine surgery

User Satisfaction Score (USS), Clinical Accuracy and relative Operation Time (OT). The larger triangle area is better.

Source: Móga et al, 2023. (31) (CC BY 4.0).

4.1.3. Benchmarking

According to the recent literature, a generally deployable, universal OR-specific and proven AR/MR headset has not yet been marketed (11). One of the main required characteristics of an HMD is its easy and time-efficient integration into the surgical workflow.

As the surgeons would not necessarily have an IT or computing background and relevant skills, and generally, there would be no such qualified staff for these tasks, the system's calibration and patient–image registration need to be adequate for the situation. On the other hand, additional health-economy investigations are necessary to provide evidence of the benefits of the presumably longer operation times using AR/MR navigation (31).

Precise and accurate overlay, localisation of the projected image in the 3D space, the contrast and quality of holograms and immersion are indisputable key factors for a device's success. Wearing and using AR/MR-HMDs for more prolonged operations are likely to happen, highlighting the importance of ergonomics and the weight distribution of the device (89). As the surgeon's head and FoV are usually pointing downwards to the working area (operation field), an adjustable, personally aligned, and balanced device could only prevent short-term exhaustion and fatigue of the user (31).

The Benchmarking requirements of an ideal surgery-specified, OR used AR/MR-HMD grouping in two levels are (11):

1, Essential Requirements (HW = Hardware, SW = Software):

- Wide Field of View (FoV), humidity-free see-through display with dust, water and droplet resistance (IP65) – HW;
- Built-in-one computer with a high computing capacity – HW;
- Broadband Wireless connection – HW;
- Long battery life – HW;
- Ergonomics and balanced weight under 400 g – HW;
- Touchless and controller-free navigation with gaze/motion/voice commands – HW;
- Easy, understandable user interface – SW;
- Support for setup, registration and calibration – SW.

2, Additional Requirements:

- Automatic focus adjusting to pupils – HW;
- Haptic/vibration feedback – HW;
- Image, video, and audio capturing, streaming (teleconsulting) – SW;
- Environmental and/or situational awareness monitoring, tool tracking – SW.

4.2. AR-HMD usability and workload assessment of surgeons in Clinical Surgery

4.2.1. Overall reception and usability

Altogether, 16 doctors from two hospitals with expertise in various surgical fields, 6 female and 10 male doctors, completed the survey in 3 years-of-experience groups. The average of surgical experience was 11.9 ± 10.3 years. Preoperative and intraoperative user experience were collected and analysed altogether as the SURG-TLX questionnaire aimed to highlight mental and physical burden, the accumulated stress, and the disturbance caused by using the device. Table VII contains the participants' distribution.

Table VII: Distribution of participant surgeons

Experience in the field (years)	**Specialty**		
	General Surgery	Urology	Vascular Surgery
Group 1: 0-6	5	2	0
Group 2: 7-15	4	1	1
Group 3: 16+	1	1	1

The general acceptance and reception of the AR/MR-HMD navigation and guidance system were highly anticipated. There were no significant differences in the general acceptance in the various surgical fields. However, a higher level of subjective “wow factor” was recognised in the less-experienced, younger generation.

In a few cases, qualitative feedback was given discussing the system’s advantages and disadvantages on the used questionnaire:
In urology, kidney benign and malign diseases (cysts, malign tumours) for preoperative surgical approach planning and intraoperative environmental awareness enhancement, (total nephrectomy or partial, organ sparing method, orientation of organ vessels and vital neighbouring tissue structures) and projection of MR fusion images in prostate procedures (biopsy, resection) seemed to be useful where the robotic surgical approaches are not accessible.

In general surgery, hepatic (liver) surgery can be enhanced mainly in the preoperative planning phase, intraoperative tumour localisation, and guidance for anatomical or non-anatomical resections. It is added, however, that the procedures for soft tissue diseases (e.g., breast surgery, gastrointestinal and colorectal surgery) may not benefit entirely from using the AR/MR-HMD navigation system. In these cases, the projected 3D organ and tissue model structures may not wholly represent their original forms and localisations as they are flexible and mobile. Nevertheless, the immediate availability of projected images of patient-specific CT and MRI scans without disrupting and shifting attention brings several benefits to spatial understanding. Understanding and knowing complex anatomical structures, such as the hepatic-biliary track, thyroid glands, and inguinal canal anatomy, can be advantageous pre- and intraoperatively for young surgical residents.
Vascular surgery is another surgical sub-speciality, where digital subtracted angiography (DSA) scans and contrasted CT scan-based 3D vasculature models benefit surgical planning and intraoperative orientation.

4.2.2. Workload Analysis with SURG-TLX

The SURG-TLX calculates 6 weight dimensions of mental-, physical burden and stress (bipolar low to high 20-point Likert scale), followed by 15 paired comparisons. The highest weight contributes the most to the perceived workload (range of 0-5), and the workload score is calculated as this is multiplied by the scale rate (59). The total workload score is counted by adding these together.
Overall analysis of the results (Figure 11) shows that the most demanding dimension was Situation Stress, followed by Temporal Demand across the workload study. The workload/taskload dimensions are derived in order (lowest to highest):

- Mental Demands,
- Distraction,
- Physical Demands,
- Task Complexity,
- Temporal Demand,
- and Situation Stress.

The average total Workload score was 64.71 ± 23.86 in all experience groups. A significant difference was measured between the years-of-experience groups: in Group 1 (0-6 years), 50.71, in Group 2 (7-15 years), 70.2, and in Group 3 (16+ years), 100 (Figure 12).

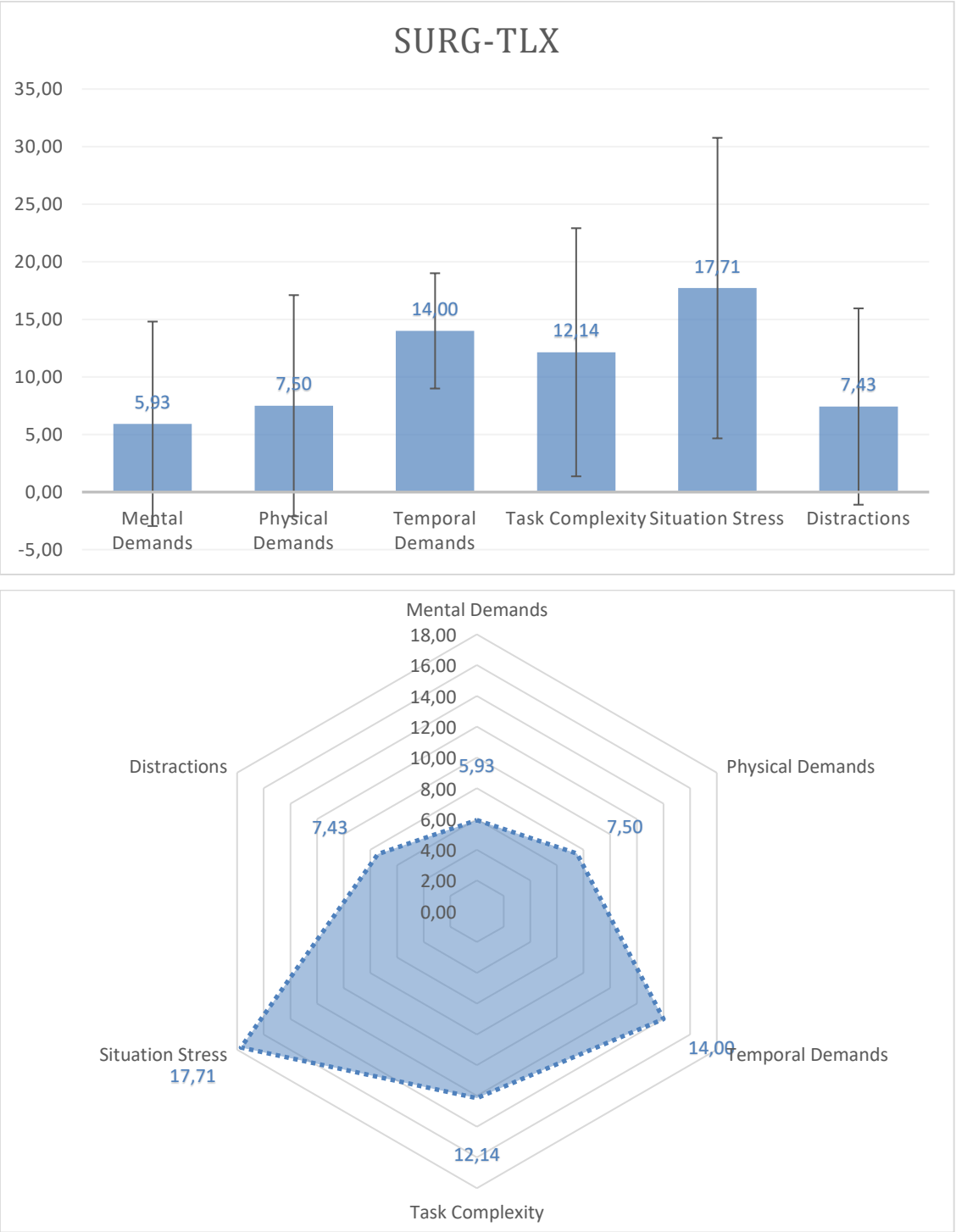


Figure 11: Average overall outcomes of SURG-TLX workload analysis using AR/MR-HMD surgical navigation and guidance system

Figure 12: Average outcomes of SURG-TLX workload analysis in different years of experience groups using AR-HMD surgical navigation and guidance system

5. DISCUSSION

Standardised studies and reports on AR/MR-HMDs navigated spine surgeries cannot yet be found in significant numbers; the identified articles are largely heterogeneous in describing objective outcomes. However, the authors presented evidence about the benefit of using the AR/MR navigation systems compared to free-handed pedicle screw insertion. The reviewed recent studies discussed low case numbers, and no exact complication rates could be collected. Meta-analysis can not be accurately performed without involving more surgical accuracy and complication data, without risking high bias.

The dissertation's findings suggest that future studies and reports on the topic should also be planned to gain standardised clinical, accurate, and procedural data for living patient care. Linear tip and angular trajectory errors were measured only in phantom and cadaver studies. Still, LTE and ATE would statistically include more objective values for in vivo use than the clinically used Gertzbein–Robbins Scale. It is also clear that the relevance of such measurements in the clinical field is not intense, as the main goal is to reduce pain and stabilise the spinal column or treat morbidity without causing any neurological complications. The conclusions of this thesis should encourage researchers to consider the following eligibility criteria in the design of future studies:

- clinical outcome (e.g., recovery rates, length of hospitalisation, visual analogue scale, and post-operative follow-up), complications;
- accuracy data (linear tip error, angular trajectory error and Gertzbein–Robbins scale);
- procedural data (operating time and radiation dose);
- user experience measured by a standard validated method.

Following the above-mentioned points as basic requirements when conducting clinical trials and observational studies in AR/MR-HMDs-guided spine surgery is desirable. In this dissertation, both of the multi-level statistical models resulted in approximately the same outcome, without significant differences from each other, showing that using AR/MR-HMD navigation and guidance for spine surgery has only a 1.3% GRS C-D-E

grade. This value supports the hypothesis that the technology reached a higher clinical readiness level, confirming the existing knowledge in this research topic. According to the reviewed articles, integrating this new technology was easy and time-efficient, did not disrupt the clinical workflow, and all the clinical outcomes were similar or better compared to robot-assisted spine surgeries, which might make the AR/MR-HMDs a cost-saving alternative method (31). Inserting into the surgical OR workflow, the AR/MR-HMD system would not significantly elongate the operation time, thus not elevating the surgical site infection risk (93). Some recommendations were made for future advancements: an HMD-built-in light source, a magnification lens, and complete system integration with robot-assisted surgical systems. User satisfaction was evident in real-world scenarios; the system increased the pedicle screw placement accuracy and decreased the overall radiation dose needed for screw implantation. However, visual discomfort and eye strain may occur through the use of AR/MR-HMD guidance, and it has limited possibilities for obese patients. Further development should consider the importance of ergonomics and the comfort of extended usability.

As seen across various domains, the recent pandemic has also accelerated the adoption of robotics in telemedicine and surgery (37). One significant advancement in the research topic was the FDA clearance of xVision by Augmedics (80). According to the FDA regulatory clearance, the requirement was a mean position accuracy error of under 2 mm and a mean trajectory error of 2° for the new system based on X-ray imaging. The investigation tested the overall system accuracy, the image registration accuracy, and the tracking accuracy. Technical performance, user need and software validation, electrical safety and electromagnetic compatibility, headset cleaning, disinfection, reusability, and biocompatibility tests were also passed (94). These preferences could be industry standards for developing future surgical-use AR/MR-HMDs.

Nevertheless, the ethical and regulatory aspects of the technology have to be managed in parallel to the technical advancement (95). Improving the transparency of the regulatory environment governing AR/MR and RAS in medicine, alongside the streamlining of standardisation frameworks and the promotion of societal acceptance, represents a critical priority for the safe and effective clinical integration of these technologies. At present, these objectives are supported by the IEEE 700X standards family (ethicsinaction.ieee.org/p7000, accessed 23.03.2025), most notably the IEEE 7000-2021:

Model Process for Addressing Ethical Concerns During System Design, which provides a structured approach to embedding ethical considerations into technological development from its earliest stages. In addition, the IEEE 7007-2021: Ontologies for Ethically Driven Robotics and Automation Systems introduces a formalised ontology framework that facilitates consistent communication, interoperability, and ethical compliance across diverse systems, with direct applicability to the emerging domain of digital surgery (96).

Beyond these standardisation efforts, alignment with existing international regulatory frameworks, such as the European Union Medical Device Regulation (MDR 2017/745), the U.S. Food and Drug Administration (FDA) guidelines for digital health and surgical navigation devices, and the International Organization for Standardization (ISO) standards, remains crucial. Such alignment ensures patient safety and device efficacy and fosters trust among healthcare professionals and patients, thereby accelerating societal acceptance. The convergence of IEEE ethical standards with formal regulatory mechanisms has the potential to establish a robust governance ecosystem, enabling responsible innovation and the sustainable adoption of AR//MR and RAS technologies in clinical practice.

5.1. Usability of AR-HMD navigation and guidance in spine surgery

Comparing this dissertation's findings based on 14 articles from the past years with the previous systematic literature review by *Bursröm et al.* shows that the Head-Mounted Augmented Reality navigation technology is slowly spreading, developing, and progressing in spine surgery.

The heterogeneity of the studies and reported data leads to difficulty in performing a meta-analysis. Not all the reviewed studies contained exact accuracy measurements, causing bias in comparison. As can be seen from the tendency in *Bursröm's* article, living patient-reported data is still not sufficient, as they reported only two publications using AR-HMD navigated procedures. At the same time, monitor-based and microscope AR were and still are the most popular methods (31, 46).

5.1.1. Accuracy and Clinical Outcomes

From the literature, it was clear that using AR/MR navigation increased the Pedicle Screw Placement accuracy compared to the Free-Hand fluoroscopy group's accuracy. *Elmi-Terander et al.* found higher accuracy in the AR vs. FH groups (AR: 93.9% vs. FH: 89.6%, $p < 0.05$) and increased GRS in a cadaveric setup (AR: 85% vs. FH: 64%, $p < 0.05$) (97). This dissertation shows clear advancement of the AR/MR-HMD accuracy results at living patients, cadaver and phantom model environments, too, with 94% to 100% GRS A and B grades and lower intra-operative screw misplacements and complications as well.

Molina et al. found that prior cadaveric data results were not significantly different from living patients' clinical results, that the absence of respiratory motion had a self-explanatory positive bias, that using low-resolution CT scans decreased the accuracy and that a one-pixel difference in the images may correspond to 0.5 mm in reality. They found the xVision HMD system to work better than other AR-HMDs and navigation systems, such as StealthStation (Medtronic Public Limited Company, Minneapolis, MN, USA) and ExcelsiusGPS (Globus Medical Inc., Audubon, PA, USA) (31, 70).

In spine surgery, different robot-assisted surgical systems are already in use, for example, the SpineAssist robot, which demonstrated a GRS A or B accuracy of 97.9% for 487 screws placed into 112 patients (98). Constant advancement and upgrades of SpineAssist increased the accuracy rate to 99% (99). In several studies, the ExcelsiusGPS system has shown GRS A-B between 96.6% and 100% (100, 101). A recent systematic literature review on the topic of spine robotic surgery by *IB Lopez et al.* (102) included the findings of 74 studies, where the evaluation clearly discusses the superiority of accurate screw trajectory by robots (up to 94-98% GRS A-B) compared to the FH technique. In revision surgery, PSP was shown to be more accurate than FH insertion (98.5% vs. 96.9%), with a lower number of superior facet joint violations (6.4% vs. 24.4%) (103). A significant reduction of operation and screw insertion times was experienced using robotic systems compared to FH ones (a mean time of 90 s/screw, overall 5.58 ± 1.22 min vs. 7.25 ± 0.84 min, $p < 0.001$) (104). Intra-operative blood loss and transfusion decreased according to a recent meta-analysis (105), and rarely occurring post-operative complications were published (106). According to a recent cost-analysis report, despite the higher application

costs, robotic surgery systems appear to be more cost-effective as a long-term strategy due to the shorter operation time, decreased need for revision surgery, lower infection rate, and reduced length of hospital stays (42, 107).

As the cost of an AR/MR-HMD system is significantly less than that of a robot-assisted spine surgical system (28) and their implantation accuracy and clinical outcomes are similar, the AR-HMD system appears to be a good alternative in spine surgery for countries with lower incomes or in the case of limited accessibility to robotic systems (108). *Yanni* highlighted the ultimate lowering of procedural costs using AR-HMD navigation (78). While preliminary reports suggest AR/MR-HMD navigation may lower costs compared with conventional navigation, robust, peer-reviewed cost-effectiveness analyses remain scarce and are explicitly identified as a research priority.

5.1.2. Patient–Image Registration and Surgical Workflow

In the literature, several patient–image registration techniques can be found relying on intra-operative tracker data (109-111). The two primary modalities are the optical and the electromagnetic trackers (112, 113). Most optical tracking systems use RGB cameras to capture movements from the reflected infrared light as they are in the field of view. Optical references include natural anatomic landmarks or artificial reflective markers (fiducials) attached to the patient's skin and surgical instruments and tools. Electromagnetic navigation has the advantage that maintaining the line of sight is not required for precise tracking; however, metals and other electric instruments used in OR can distort the magnetic field and, thus, diminish the accuracy of the tracked tool (31, 114).

Surgical workflow can be positively affected by the use of Head-Mounted Augmented Reality navigation (65). Studies reported a slight increase in the operation time, as more work was needed for patient–image registration. *Urakov* and *Bhatt et al.* mentioned more effective and shorter learning curves using AR navigation (62, 79).

Lowering the radiation dose affecting patients and medical staff has been reported by several studies; however, only *Bhatt et al.* added exact data, where the mean total fluoroscopy time was 25.7 ± 29.8 seconds, and the mean total fluoroscopy dose was 0.3

± 0.4 mGy/cm2 for their 19 out of 27 patients. The mean total 3D imaging dose was 576.8 ± 368.8 mGy/cm to place screws using AR-navigation for 27 patients (31, 62).

The previous systematic literature review in the field by *Burströmet al.* presented sporadic data about the decrease of radiation dose: one study reported a significantly lower fluoroscopy dose-area and time with the use of AR-HMD (182.6 ± 106.7 mGy/cm2 vs. FH: 367.8 ± 184.7 mGy/cm2 and 5.2 ± 2.6 seconds vs. FH: 10.4 ± 4.1 seconds). Regarding their findings, for patient registration, the effective mean dose of intra-operative CT or cone-beam CT was 0.22 ± 0.16 mSv (cervical) and 15.8 ± 1.8 mSv (thoracolumbar) (46). In the case of RAS spine procedures, *IB Lopez et al.* observed a faster learning curve, with significant accuracy elevation and decreased operating time following 10 to 30 cases. They reported 80% less "per screw" and 78.3% "overall" radiation reduction for patients and the surgical staff during the procedure, compared to the standard FH fluoroscopy-guided technique (102). The same overall results were found using Mazor Renaissance and the ROSA robotic systems (115). In their recent systematic literature review, *Luengo-Matos et al.* described a lower radiation dose requirement for RAS than FS based on their findings from four studies (31, 116).

5.1.3. Ergonomics and Further Perspectives

In the recent literature, a few authors added user experience regarding ergonomics and comfort regarding the AR/MR-HMD system employed. *Molina* mentioned that using the device added no mechanical discomfort during short-term use, with approximately 800~ grams of weight, and, with the glasses, there was no need to look at the remote screen for patient-specific data; however, the paper also discussed some visual discomfort and obstruction caused by translucent data (70). Meanwhile, *Yanni et al.* found that the headset's ergonomics were not yet optimised for the surgical suite (78).

As feedback is not yet widely provided by authors regarding the employed AR/MR-HMD systems, further improvements and device developments cannot yet be accurately specified and standardised for OR use (31). These AR/MR-HMD navigation and guidance system specifics should be the basic requirements for device development. Still, further ergonomic data collection with more detailed experiences from the surgeons (as

end-users) is needed. This could be the key to achieving specific AR/MR-HMD development (11). Examples can be found as user scoring systems for AR systems, measuring how easy and accessible the device is, showing their complexity, or whether any support was needed through their use, scoring the interface and confidentiality during their use, etc. (QUIS, SGUS, SURG-TLX…) (117).

5.1.4. Situation Awareness

Using an AR/MR-HMD helps in navigation and guidance for the surgeons with projected holograms, giving additional spatial and depth knowledge of the operation region and field. Currently, this function works passively, like using a map. Still, advancements in computer vision and pattern recognition, deep learning, and AI solutions could soon lead to different levels of advanced environmental and situational awareness. This would reform surgery by identifying objects, calling for attention to adverse events and giving additional recommendations on the subsequent surgical steps and approaches (31, 118). Understandably, while maintaining a high focus on certain surgical steps and situations, the peripheral awareness and attention of the surgeon may drop off throughout the procedure. In the case of adverse events, such as bleeding or surgical clip misplacement, a discrete, not intrusive, alarming sound and/or visual signal could help identify and localise the problem. Virtual labels added to certain vessels, ducts or tissue structures could prevent unintentional injuries (119). The holographic guidance of specific surgical steps, such as an ideal incision or skin closure and suturing technique, may help achieve better wound-healing outcomes with less deformation and scars. Through the evolution of deep learning and computer vision algorithms, further environmental and situational awareness could be achieved in surgery (120). It is not an exaggerated thought that the used HMDs could collect and analyse various situation awareness data, helping with a better and safer understanding of the whole operative environment. A future problem to be solved is the idea of using the registration and navigation purpose in acute, urgent care situations, for example, polytraumatic injuries, when time is the most relevant factor in short-term mortality and long-term morbidity. The holographic 3D models in these

immediate conditions must be prepared urgently while maintaining the highest accuracy for safe and efficient work (31).

5.1.5. User Satisfaction Score and Comparison Tool of Surgical Navigation and Guidance Devices

Surgical AR navigation and guidance technologies fulfil technical, functional, and clinical requirements. Collecting, reviewing and representing all the primary attribution data through a standardised methodology could give the end-users valuable information regarding AR/MR-HMDs. As end-users, surgeons should provide reliable and standardised feedback about their experience of AR/MR-HMDs that they use during clinical procedures. Collecting these data would help systematically evaluate these valuable responses so that manufacturers can improve and upgrade their devices (31).
A recent overview by *Bitkina et al.* collected qualitative and quantitative evaluation possibilities of medical device usability and user experience (89):

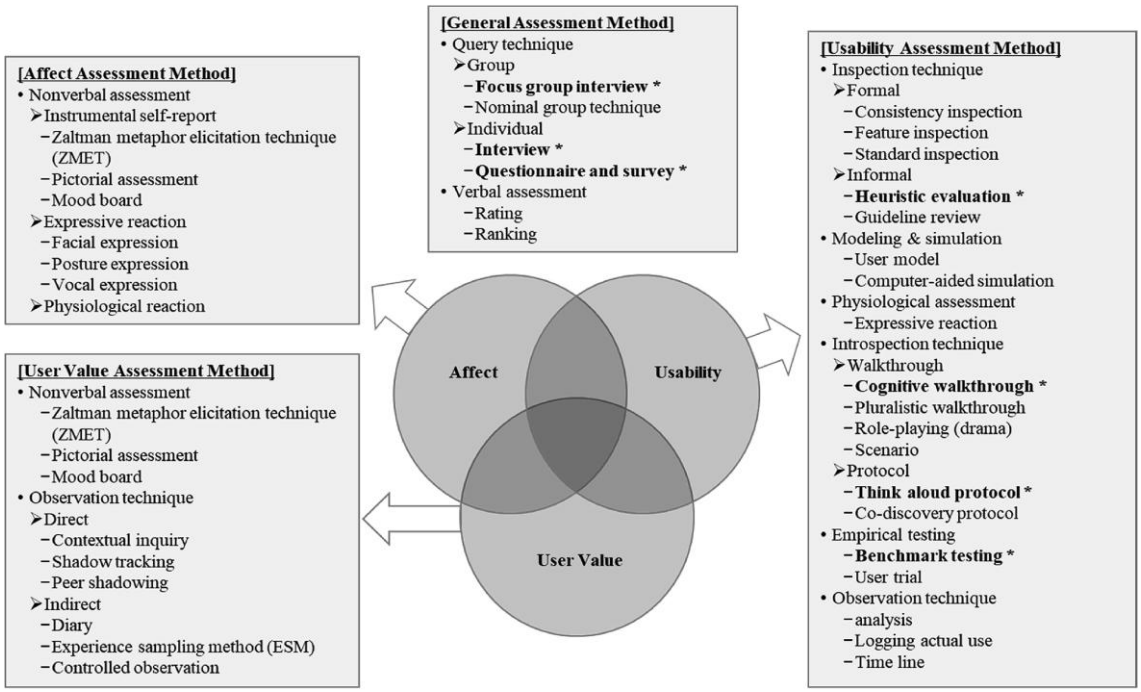


Figure 13: Usability scoring and evaluation options for medical devices collected by *Bitkina et al.* (89)

Source: Móga et al, 2023. (31) (CC BY 4.0).

Several evaluation possibilities regarding AR use in surgery can be found in the literature. *Dennler et al.* shared their results concerning a standardised 58-item electronic survey questionnaire using the REDCap data capture tool. Their survey was based on a 100-point scale with anchor statements (1: not useful at all, to 100: very useful), asking about the usability of the AR technology and the device. The overall satisfaction with the application of AR technology in spine surgery was described (78 ± 20 points), and future access was demanded (75 ± 22 points). The main valued aspect can be seen in Figure 14 (31, 71).

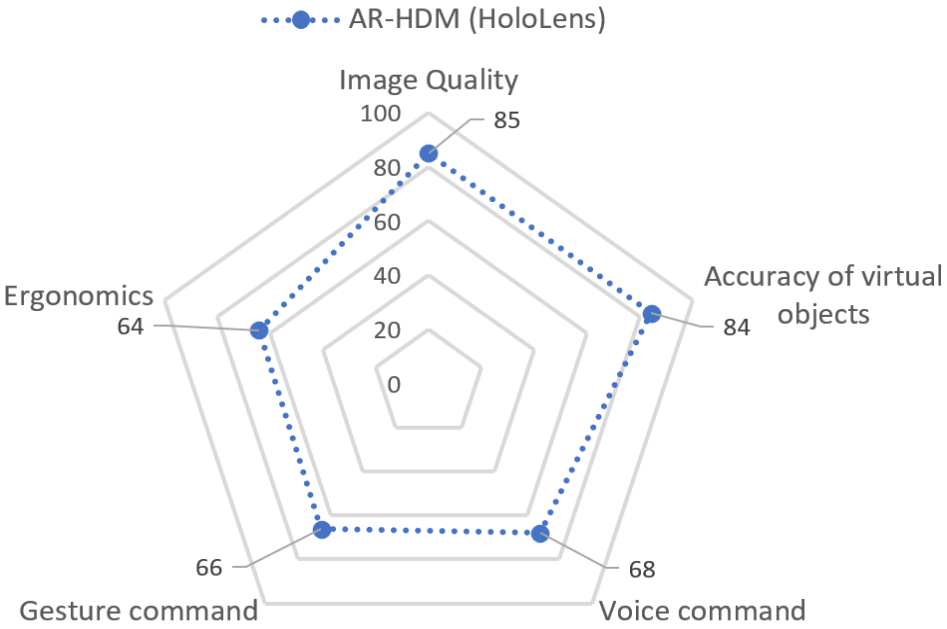


Figure 14: Example of scoring AR spine surgery systems based on the article by *Dennler et al.* (71)

Source: Móga et al, 2023. (31) (CC BY 4.0).

Other User Satisfaction Scoring systems can be found, such as the Questionnaire for User Interaction Satisfaction (QUIS) tool - which was designed to score human-computer interfaces (regarding the screen, system feedback, learning curve, technical capabilities and manuals, tutorials and multimedia use) and the Smart Glasses User Satisfaction (SGUS) questionnaire (regarding the AR perception, interaction, location and object awareness and AR content) (90, 91). These tools are only sporadically used and do not seem well-known in the surgical data publishing practice. It is essential to highlight the importance of using a standardised methodology for scoring different aspects of AR/MR-HMDs, similar to surgical robotics, for large-scale clinical trials (92).

5.1.6. Technical Requirements and Benchmarking

The most recent studies discuss four AR-display methods: HMD, monitor, projector and microscope-based. The most feasible way for the perioperative period of surgical procedures and interventions could be HMDs, as the all-in-one built computer and display instrument holds the best mobility and controllability with the least disturbance to the operative environment or affecting patient safety. Different requirements, specifications and functions are needed for various applications and adaptations.

Precise and accurate patient-specific model overlay, localisation in the 3D space of the projected image, the contrast and quality of holograms and immersion are indisputable key factors for a device's success. Wearing and using AR/MR-HMDs for more prolonged operations are likely to happen, highlighting the importance of ergonomics and the weight distribution of the device (89). As the surgeon's head and FoV are usually pointing downwards to the working area (operation field), an adjustable, personally aligned, and balanced device could only prevent short-term exhaustion and fatigue of the user (31).

As AR/MR-HDMs are built-in-one computers, the fundamental aspects for functionality are see-through displays with a wide FoV (Field of View), FHD (Full High Definition) resolution and HDR (High Dynamic Range) colour support, high-contrast imaging, high computing capacity with efficient cooling, image and video capturing with tracking sensors (gesture, voice, gaze - pupil, corneal reflection or blink patterns) and depth perception. As a power supply has to be guaranteed even for more prolonged procedures, sufficiently large and/or interchangeable batteries combined with energy-saving and/or recuperation strategies are crucial. Integrated communication capability is also required if further participants (physical or virtual) need to be involved in the AR/MR experience. Finally, ergonomics (wireless device, low weight, compact size, compliant with occupational safety, and workplace ergonomics regulations, i.e., shatterproof lenses and face shields) and effortless usability and time-efficient integration into the surgical workflow could be mentioned without explanation.

Surgeons do not necessarily have sufficient IT knowledge or computing background and relevant skills, and in the OR, qualified staff for these tasks may not be present. Thus, the system‘s calibration and patient–image registration must be adequate for the situation.

Using the devices in sterile conditions but within a high-humidity and direct fluid droplets environment, dust and water resistance are also absolute requirements, affecting their lifetime. For the best operative navigation and guidance experience, environment and tool recognition could also be a standard. The above-suggested minimum technical requirements are proposed for the specified OR-use device development. Without meeting these requirements, the intended use of surgical sites or procedures is nearly impossible in the long term.

In conclusion, similar to any medical device to be used in the surgical field, especially in operating room environments, HMDs must accomplish these Benchmarking requirements:

- Technical:
  High processing and computing capacity;
  Transparent full HD, HDR coloured, large field of view display;
  IP65 (dust, water resistant).
- Clinical:
  Gesture, voice control;
  Image, video and audio capturing, sharing;
  Tool tracking, environment and situation awareness;
  Sterilisable, cleanable.
- Usability:
  Ergonomics (e.g., weight, compliant with regulations);
  Wireless, integrated battery;
  Easy to use.

One of the most interesting HMD-related development areas is how the system could be merged with RAS or how haptic feedback could be integrated into its use. A recent patent describes that haptics function may include vibratory motors, electro-active polymers, piezoelectric devices, electrostatic devices, subsonic audio wave surface actuation devices, and reverse-electro-vibration to provide tactile feedback to the user (121). On the other hand, additional health-economy investigations are needed to provide evidence of the benefits of the presumably longer operation times using AR navigation.

5.1.7. Challenges in Implementing AR/MR in Surgery and Potential for Future Development

AR-HDMs use different visual representation methods, but all must comfortably accomplish immersion. Using specific waveguide technology, such as Microsoft's HoloLenses, has helped with the limitations of sight focus. However, the more extensive the planned field of view is, the more challenging it is for opticians to develop these for AR/MR. These challenges include the FoV/eye-box trade-off, angular resolution/optical foveation, achieving HDR, hard edge occlusion, and Vergence-Accommodation Conflict (VAC). Different imaging for central and peripheral sight is also a future challenge to be solved (122). Until this day, no researchers have investigated the effect of different vision defects, such as myopia, hypermetropia, astigmatism and dyschromatopsia, nor left/right-handedness on the surgical usability of AR/MR devices.

Other technological limitations, such as the device's ergonomics and weight, should be highlighted, as the user's discomfort could cause mental and physical burdens, especially in prolonged procedures, which may affect surgical outcomes. Moreover, the effortless integration into the existing surgical workflow must be mentioned again. Future research directions are visible: apply haptic feedback to the AR guidance system and AI integration to enhance situation awareness and real-time, patient-specific 3D modelling and projection overlay. Last, merging existing and new innovative technologies into medicine and surgery and applying step-by-step semi- or fully autonomous AI integration would lead to the next medical reform.

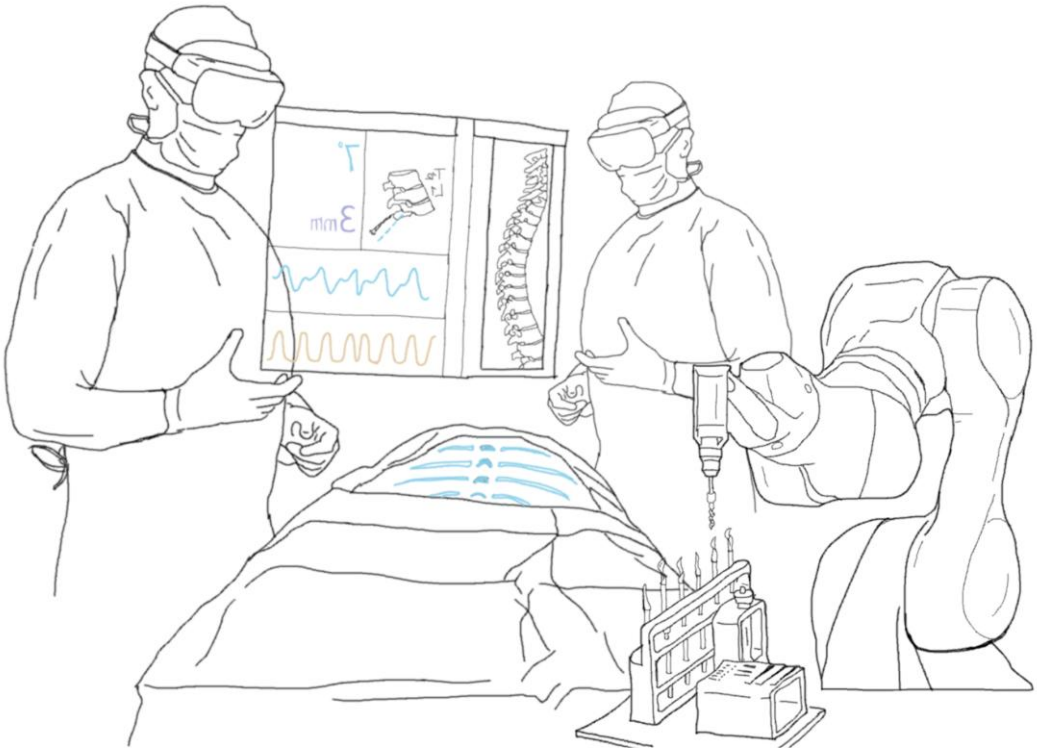


Figure 15: Vision concept about merging AR-HMD guidance with autonomous RAS technologies

Source: Móga et al, 2023. (55) (CC BY 4.0).

5.2. Workload study of AR-HMD navigation and guidance in surgery

The workload and taskload study's findings suggest that AR/MR-HMDs, when ergonomically optimised and accompanied by proper training, have significantly reduced workload burden, particularly for junior staff. However, the occurring situational stress through the use of the system and its physical demands requires design improvements, especially for long-duration procedures. Enhanced integration of intuitive controls, voice commands, and improved weight distribution would address these barriers.

Interestingly, less experienced surgeons (0–6 years) reported a higher subjective benefit and lower stress levels, indicating a higher adaptability to AR interfaces. Experienced surgeons (16+ years) expressed more concern over prolonged use and visual discomfort.

The overall workload study using the SURG-TLX framework highlighted critical usability parameters of AR/MR-HMD systems across different surgical specialities. Mental demand and distraction were rated lowest among the six domains, suggesting cognitive offloading through in-view data. However, there was a significant difference between Groups: the more experienced group (Group 3) showed more mental demand using the system, which can be explained by lower familiarity with newer technologies. Task complexity and physical demand were moderate, while temporal demand and situational stress were highest. This suggests that while AR/MR-HMDs streamline perception and spatial awareness, they may introduce novel ergonomic and attention-related stressors.

6. CONCLUSIONS

Digital health has emerged as one of the most critical areas in modern medicine, fundamentally transforming surgical practice through new technologies such as augmented and mixed reality systems. Its importance lies not only in its clinical relevance and the potential to improve patient safety and surgical precision, but also in its timeliness as healthcare systems worldwide seek efficient, cost-effective, and minimally invasive solutions. The scientific value of this field is underscored by the need for rigorous and standardised research methods to objectively evaluate usability, accuracy, workload, and situation awareness in real surgical environments. Therefore, this dissertation applied structured, validated methodologies to investigate AR/MR-HMDs in surgery, ensuring the findings are reproducible, comparable, and meaningful to digital surgery's evidence base.

My dissertation focused on and answered the following research questions:

I. Usability, Situation Awareness and Benchmarking for Intraoperative AR/MR HMDs

A) Collect and comprehensively analyse the clinical impact and usability in recent spine surgery – Pedicle Screw Positioning (PS, PSP) studies using AR/MR-HMD technology as navigation and guidance:

A comprehensive analysis was given, showing that AR/MR-HMD systems such as xVision and HoloLens showed high clinical usability with pedicle screw accuracy exceeding 98.5% (GRS A+B) in PSP. The navigation and guidance system enhanced surgical workflow without disrupting the operation, reduced radiation exposure, and was especially well-received by junior surgeons due to its intuitive visual guidance and minimal learning curve.

B) Situation Awareness: highlight the importance of measuring and standardising situation awareness in the research field, as it is still missing:

My study confirmed that SA is still underrepresented in AR/MR-HMD surgical research. While most authors mentioned subjective improvements in spatial understanding, no studies applied standardised SA metrics (e.g., SAGAT - Situation Awareness Global Assessment Technique, real-time probe, or EEG-based assessment). The dissertation emphasises that structured, validated tools must be adopted in future trials for objective comparison.

C) Offer a Standardised Scoring method with visualisation for comparing AR/MR-HMD surgical navigation and guidance tools:

A novel 3D scoring model was developed and presented, incorporating clinical accuracy, operation time, and radiation exposure, which allows meaningful comparisons between AR/MR-HMDs and other navigation systems (e.g., RAS or FH). The model was further enhanced by introducing a User Satisfaction Score (USS) to integrate subjective usability.

D) Benchmarking: pointing out surgical AR/MR-HMDs benchmarks for future research and development:

My dissertation proposed a two-level benchmarking framework:

- Essential requirements: wide FoV, IP65 protection, ergonomic sub-400g weight, wireless communication, and intuitive gaze/voice control.
- Additional features: haptic feedback, live streaming, AI-enhanced SA, and universal OR integration.

This framework supports future hardware and software development toward clinical-grade AR/MR-HMDs.

II. Workload analysis in the real clinical-surgical environment to prove the usability advantages of AR-HMDs

Present a use-case study to measure the effect of an AR-HMD navigation and guidance system on users' workload and task load during its use in clinical and surgical environments:

Using the SURG-TLX framework, this dissertation's workload study revealed:

- Lower mental demand and distraction among junior surgeons (0–6 years).
- Moderate-to-high stress and physical burden for mid to senior surgeons, highlighting ergonomic challenges.
- The overall workload score was lowest in the youngest group (≈50), increasing with experience, peaking in the senior group (100).
- This confirms that AR-HMDs offer substantial usability benefits, especially for younger or less experienced surgeons, while also identifying areas for ergonomic improvement in experienced users.

7. SUMMARY

This dissertation presented the latest results regarding implementing Augmented and Mixed Reality Head-Mounted Displays in spine surgery, focusing on clinical accuracy, surgeon workload, usability, and benchmarking metrics. Through a systematic review and workload analysis study, AR/MR-HMD systems like xVision or HoloLens-based navigation and guidance demonstrated high pedicle screw placement accuracy (GRS A+B > 98.5%), reduced radiation exposure, and minimal disruption to surgical workflow.

Usability studies using the SURG-TLX framework revealed that AR/MR-HMDs decrease mental demand and distraction, especially among junior surgeons (0–6 years), who reported the lowest overall workload scores. Surgeons in mid (7–15 years) and senior (16+ years) experience groups perceived higher physical and temporal burden, indicating the need for ergonomic refinements and adaptive interfaces. Cross-speciality use further validated the broad applicability of AR/MR-HMDs in digital surgery.

The dissertation also introduced a novel multidimensional 3D benchmarking framework combining accuracy, operative time, and radiation exposure. This was enhanced by a proposed User Satisfaction Score (USS), enabling comprehensive evaluation of surgical navigation systems. Additionally, the research highlighted the underdeveloped field of Situation Awareness in AR/MR-HMD literature and advocated for integrating standardised measurement tools like SAGAT, eye-tracking, or EEG.

Technical benchmarks for clinical-grade AR/MR-HMDs were outlined, recommending features such as a wide field of view, IP65-rated casings, ergonomic design, and seamless OR integration. These findings guide future development, validation, and regulation of AR/MR-HMD platforms in surgery.

In conclusion, AR/MR-HMD systems present a viable, scalable alternative to robotic guidance. They offer ergonomic, cognitive, and cost-efficiency benefits and are particularly beneficial for younger surgeons. With continued development in SA metrics and interface design, AR/MR-HMDs are positioned to become key tools in the future of digital and minimally invasive surgery.

8. REFERENCES

1. Taylor RH, Simaan N, Menciassi A, Yang GZ. (2022) Surgical Robotics and Computer-Integrated Interventional Medicine [Scanning the Issue]. Proceedings of the IEEE, 110: 823-834.
2. Georgiou KE, Georgiou E, Satava RM. (2021) 5G Use in Healthcare: The Future is Present. Journal of the Society of Laparoendoscopic Surgeons, 25.
3. Li T, Badre A, Alambeigi F, Tavakoli M. (2023) Robotic Systems and Navigation Techniques in Orthopedics: A Historical Review. Applied Sciences, 13: 9768.
4. Begley SL, McBriar JD, Pelcher I, Schulder M. (2024) Intraoperative MRI: A Review of Applications Across Neurosurgical Specialties. Neurosurgery, 95: 527-536.
5. Vahrmeijer AL, Hutteman M, van der Vorst JR, van de Velde CJ, Frangioni JV. (2013) Image-guided cancer surgery using near-infrared fluorescence. Nat Rev Clin Oncol, 10: 507-518.
6. Takács K, Haidegger T. (2024) Eye gaze tracking in robot-assisted minimally invasive surgery: a systematic review of recent advances and applications. Acta Polytechnica Hungarica, 21: 393-411.
7. Takács K, Lukács E, Levendovics R, Pekli D, Szijártó A, Haidegger T. (2024) Assessment of Surgeons' Stress Levels with Digital Sensors during Robot-Assisted Surgery: An Experimental Study. Sensors, 24: 2915.
8. Hayashi Y, Misawa K, Oda M, Hawkes DJ, Mori K. (2016) Clinical application of a surgical navigation system based on virtual laparoscopy in laparoscopic gastrectomy for gastric cancer. Int J Comput Assist Radiol Surg, 11: 827-836.
9. Milgram P, Takemura H, Utsumi A, Kishino F. (1995) Augmented reality: a class of displays on the reality-virtuality continuum. In: Proceedings of SPIE - The International Society for Optical Engineering Vol. 2351 pp. 282-292
10. IEC. (2019) VR definition. Vol. 2024
11. Moga K, Boesl DBO, Haidegger T. (2021) Augmented/Mixed Reality Technologies Supporting Digital Surgery. In: SISY 2021 - IEEE 19th International Symposium on Intelligent Systems and Informatics, Proceedings,

doi:10.1109/SISY52375.2021.9582533 pp. 183-189, Institute of Electrical and Electronics Engineers Inc.

12. Casa ED. (2018) The Milgram Continuum at a glance. Vol. 2024, 01. 25. 2018
13. ISO. (2019) ISO/IEC 18039:2019. AR and MR definition. Vol. 2024
14. Schmalstieg D, Hollerer T. (2017) Augmented reality: Principles and practice. In: Proceedings - IEEE Virtual Reality, doi:10.1109/VR.2017.7892358 pp. 425-426
15. Schäfer A, Reis G, Stricker D. (2022) A Survey on Synchronous Augmented, Virtual and Mixed Reality Remote Collaboration Systems. ACM Comput. Surv, 1: 23.
16. Galambos P, Csapó Á, Zentay P, Fülöp IM, Haidegger T, Baranyi P, Rudas IJ. (2015) Design, programming and orchestration of heterogeneous manufacturing systems through VR-powered remote collaboration. Robotics and Computer-Integrated Manufacturing, 33: 68-77.
17. Guzsvinecz T, Orbán-Mihálykó É, Perge E, Sik-Lányi C. (2020) Analyzing the spatial skills of university students with a virtual reality application using a desktop display and the gear VR. Acta Polytechnica Hungarica, 17: 35-56.
18. Haidegger T, Galambos P, Rudas IJ. (2019) Robotics 4.0 - Are we there yet? In: INES 2019 - IEEE 23rd International Conference on Intelligent Engineering Systems, Proceedings, doi:10.1109/INES46365.2019.9109492 pp. 117-124
19. Gibaud B, Forestier G, Feldmann C, Ferrigno G, Gonçalves P, Haidegger T, Julliard C, Katić D, Kenngott H, Maier-Hein L, März K, de Momi E, Nagy DÁ, Nakawala H, Neumann J, Neumuth T, Rojas Balderrama J, Speidel S, Wagner M, Jannin P. (2018) Toward a standard ontology of surgical process models. International Journal of Computer Assisted Radiology and Surgery, 13: 1397-1408.
20. Kan Yeung AW, Tosevska A, Klager E, Eibensteiner F, Laxar D, Stoyanov J, Glisic M, Zeiner S, Kulnik ST, Crutzen R, Kimberger O, Kletecka-Pulker M, Atanasov AG, Willschke H. (2021) Virtual and augmented reality applications in medicine: Analysis of the scientific literature. Journal of Medical Internet Research, 23.
21. Dale S. (2024) Optics for Virtual, Augmented and Mixed Reality 2024-2034: Technologies, Players and Markets.

22. Topsakal O, Mazhar Çelikoyar M. (2019) Surgeries that would benefit from augmented reality and their unified user interface. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) Vol. 11613 LNCS pp. 303-312
23. Wang P, Bai X, Billinghurst M, Zhang S, Zhang X, Wang S, He W, Yan Y, Ji H. (2021) AR/MR Remote Collaboration on Physical Tasks: A Review. Robotics and Computer-Integrated Manufacturing, 72.
24. Inkpen KM, Sedlins M. (2011) Me and my avatar: Exploring users' comfort with avatars for workplace communication. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW, doi:10.1145/1958824.1958883 pp. 383-386
25. Kumar RP, Pelanis E, Bugge R, Brun H, Palomar R, Aghayan DL, Fretland ÅA, Edwin B, Elle OJ. (2020) Use of mixed reality for surgery planning: Assessment and development workflow. Journal of Biomedical Informatics: X, 8.
26. Pelanis E, Teatini A, Eigl B, Regensburger A, Alzaga A, Kumar RP, Rudolph T, Aghayan DL, Riediger C, Kvarnström N, Elle OJ, Edwin B. (2021) Evaluation of a novel navigation platform for laparoscopic liver surgery with organ deformation compensation using injected fiducials. Medical Image Analysis, 69.
27. Survarachakan S, Pelanis E, Khan ZA, Kumar RP, Edwin B, Lindseth F. (2021) Effects of enhancement on deep learning based hepatic vessel segmentation. Electronics (Switzerland), 10.
28. Fichtinger G, Troccaz J, Haidegger T. (2022) Image-Guided Interventional Robotics: Lost in Translation? Proceedings of the IEEE, 110: 932-950.
29. Ravindra VM, Senglaub SS, Rattani A, Dewan MC, Härtl R, Bisson E, Park KB, Shrime MG. (2018) Degenerative Lumbar Spine Disease: Estimating Global Incidence and Worldwide Volume. Global Spine Journal, 8: 784-794.
30. Dahdaleh NS, Smith ZA, Hitchon PW. (2014) Percutaneous pedicle screw fixation for thoracolumbar fractures. Neurosurgery Clinics of North America, 25: 337-346.
31. Móga K, Ferencz A, Haidegger T. (2023) What Is Next in Computer-Assisted Spine Surgery? Advances in Image-Guided Robotics and Extended Reality. Robotics, 12.

32. Staartjes VE, Klukowska AM, Schröder ML. (2018) Pedicle Screw Revision in Robot-Guided, Navigated, and Freehand Thoracolumbar Instrumentation: A Systematic Review and Meta-Analysis. World Neurosurgery, 116: 433-443.e438.
33. Adogwa O, Parker SL, Shau D, Mendelhall SK, Aaronson O, Cheng J, Devin CJ, McGirt MJ. (2015) Cost per quality-adjusted life year gained of revision fusion for lumbar pseudoarthrosis. Journal of Spinal Disorders and Techniques, 28: 101-105.
34. Amato V, Giannachi L, Irace C, Corona C. (2010) Accuracy of pedicle screw placement in the lumbosacral spine using conventional technique: Computed tomography postoperative assessment in 102 consecutive patients - Clinical article. Journal of Neurosurgery: Spine, 12: 306-313.
35. Du JP, Wang DH, Zhang J, Fan Y, Wu QN, Hao DJ. (2018) Accuracy of Pedicle Screw Insertion Among 3 Image-Guided Navigation Systems: Systematic Review and Meta-Analysis. World Neurosurgery, 109: 24-30.
36. Siller S, Raith C, Zausinger S, Tonn JC, Szelenyi A. (2019) Indication and technical implementation of the intraoperative neurophysiological monitoring during spine surgeries—a transnational survey in the German-speaking countries. Acta Neurochirurgica, 161: 1865-1875.
37. Khamis A, Meng J, Wang J, Azar AT, Prestes E, Li H, Hameed IA, Takács Á, Rudas IJ, Haidegger T. (2021) Robotics and intelligent systems against a pandemic. Acta Polytechnica Hungarica, 18: 13-35.
38. Smith AD, Chapin J, Birinyi PV, Bhagvath PV, Hall AF. (2021) Automated Polyaxial Screw Placement Using a Commercial-Robot-Based, Image-Guided Spine Surgery System. IEEE Transactions on Medical Robotics and Bionics, 3: 74-84.
39. Overley SC, Cho SK, Mehta AI, Arnold PM. (2017) Navigation and robotics in spinal surgery: Where are we now? Clinical Neurosurgery, 80: S86-S99.
40. Haidegger T. (2019) Autonomy for surgical robots: Concepts and paradigms. IEEE Transactions on Medical Robotics and Bionics, 1: 65-76.
41. Haidegger T. (2019) Probabilistic Method to Improve the Accuracy of Computer-Integrated Surgical Systems. Acta Polytechnica Hungarica, 16: 112.

42. Menger RP, Savardekar AR, Farokhi F, Sin A. (2018) A cost-effectiveness analysis of the integration of robotic spine technology in spine surgery. Neurospine, 15: 216-224.
43. 3D Slicer image computing platform. Vol. 2024.
44. TotalSegmentator AI model Vol. 2024
45. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller JV, Pieper S, Kikinis R. (2012) 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging, 30: 1323-1341.
46. Burström G, Persson O, Edström E, Elmi-Terander A. (2021) Augmented reality navigation in spine surgery: a systematic review. Acta Neurochirurgica, 163: 843-852.
47. DoD. (2020) USA Department of Defense Instruction 6000.19. Military Medical Treatment Facility Support of Medical Readiness Skills of Health Care Providers. Vol. 2024
48. Park S, Bokijonov S, Choi Y. (2021) Review of microsoft hololens applications over the past five years. Applied Sciences (Switzerland), 11.
49. Behringer R, Christian J, Holzinger A, Wilkinson S. (2007) Some usability issues of augmented and mixed reality for e-Health applications in the medical domain. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) Vol. 4799 LNCS pp. 255-266
50. Gabbard JL, Hix D, Swan Ii JE. (1999) User-centered design and evaluation of virtual environments. IEEE Computer Graphics and Applications, 19: 51-59.
51. Kalawsky RS. (1999) VRUSE - A computerised diagnostic tool: For usability evaluation of virtual/synthetic environment systems. Applied Ergonomics, 30: 11-25.
52. Lombard M, Ditton T. (1997) At the heart of it all: The concept of presence. Journal of Computer-Mediated Communication, 3.
53. (2020) Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA). Vol. 2024

54. Gertzbein S, Robbins S. (1990) Accuracy of pedicular screw placement in vivo. Spine, 15: 11-14.
55. Móga K, Hölgyesi Á, Zrubka Z, Péntek M, Haidegger T. (2023) Augmented or Mixed Reality Enhanced Head-Mounted Display Navigation for In Vivo Spine Surgery: A Systematic Review of Clinical Outcomes. J Clin Med, 12.
56. Hart SG. (2006) NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the Human Factors and Ergonomics Society, pp. 904-908
57. Wilson MR, Poolton JM, Malhotra N, Ngo K, Bright E, Masters RS. (2011) Development and validation of a surgical workload measure: the surgery task load index (SURG-TLX). World J Surg, 35: 1961-1969.
58. Lowndes BR, Forsyth KL, Blocker RC, Dean PG, Truty MJ, Heller SF, Blackmon S, Hallbeck MS, Nelson H. (2020) NASA-TLX Assessment of Surgeon Workload Variation Across Specialties. Ann Surg, 271: 686-692.
59. Wilson MR, Poolton JM, Malhotra N, Ngo K, Bright E, Masters RSW. (2011) Development and validation of a surgical workload measure: The surgery task load index (SURG-TLX). World Journal of Surgery, 35: 1961-1969.
60. Ungi T, Lasso A, Fichtinger G. (2016) Open-source platforms for navigated image-guided interventions. Med Image Anal, 33: 181-186.
61. Abe Y, Sato S, Kato K, Hyakumachi T, Yanagibashi Y, Ito M, Abumi K. (2013) A novel 3D guidance system using augmented reality for percutaneous vertebroplasty. Journal of Neurosurgery: Spine, 19: 492-501.
62. Bhatt FR, Orosz LD, Tewari A, Boyd D, Roy R, Good CR, Schuler TC, Haines CM, Jazini E. (2023) Augmented Reality-Assisted Spine Surgery: An Early Experience Demonstrating Safety and Accuracy with 218 Screws. Global Spine Journal, 13: 2047-2052.
63. Li J, Zhang H, Li Q, Yu S, Chen W, Wan S, Chen D, Liu R, Ding F. (2021) Treating lumbar fracture using the mixed reality technique. BioMed Research International, 2021.
64. Liu A, Jin Y, Cottrill E, Khan M, Westbroek E, Ehresman J, Pennington Z, Lo SFL, Sciubba DM, Molina CA, Witham TF. (2022) Clinical accuracy and initial experience with augmented reality-assisted pedicle screw placement: the first 205 screws. Journal of Neurosurgery: Spine, 36: 351-357.

65. Yahanda AT, Moore E, Ray WZ, Pennicooke B, Jennings JW, Molina CA. (2021) First in-human report of the clinical accuracy of thoracolumbar percutaneous pedicle screw placement using augmented reality guidance. Neurosurgical Focus, 51: 1-8.
66. Butler AJ, Colman MW, Lynch J, Phillips FM. (2023) Augmented reality in minimally invasive spine surgery: early efficiency and complications of percutaneous pedicle screw instrumentation. Spine Journal, 23: 27-33.
67. Harel R, Anekstein Y, Raichel M, Molina CA, Ruiz-Cardozo MA, Orrú E, Khan M, Mirovsky Y, Smorgick Y. (2022) The XVS System During Open Spinal Fixation Procedures in Patients Requiring Pedicle Screw Placement in the Lumbosacral Spine. World Neurosurgery, 164: e1226-e1232.
68. Judy BF, Liu A, Jin Y, Ronkon C, Khan M, Cottrill E, Ehresman J, Pennington Z, Bydon A, Lo SFL, Sciubba DM, Molina CA, Witham TF. (2023) In-Human Report of S2 Alar-Iliac Screw Placement Using Augmented Reality Assistance. Operative Neurosurgery, 24: 68-73.
69. Zhang JN, Fan Y, Hao DJ. (2019) Risk factors for robot-assisted spinal pedicle screw malposition. Scientific Reports, 9.
70. Molina CA, Sciubba DM, Greenberg JK, Khan M, Witham T. (2021) Clinical Accuracy, Technical Precision, and Workflow of the First in Human Use of an Augmented-Reality Head-Mounted Display Stereotactic Navigation System for Spine Surgery. Operative Neurosurgery, 20: 300-309.
71. Dennler C, Bauer DE, Scheibler AG, Spirig J, Götschi T, Fürnstahl P, Farshad M. (2021) Augmented reality in the operating room: a clinical feasibility study. BMC Musculoskeletal Disorders, 22.
72. Farshad M, Fürnstahl P, Spirig JM. (2021) First in man in-situ augmented reality pedicle screw navigation. North American Spine Society Journal, 6.
73. Felix B, Kalatar SB, Moatz B, Hofstetter C, Karsy M, Parr R, Gibby W. (2022) Augmented Reality Spine Surgery Navigation Increasing Pedicle Screw : Insertion Accuracy for Both Open and Minimally Invasive S Surgeries. Spine, 47: 865-872.

74. Spirig JM, Roner S, Liebmann F, Fürnstahl P, Farshad M. (2021) Augmented reality-navigated pedicle screw placement: a cadaveric pilot study. European Spine Journal, 30: 3731-3737.
75. Farshad M, Spirig JM, Suter D, Hoch A, Burkhard MD, Liebmann F, Farshad-Amacker NA, Fürnstahl P. (2021) Operator independent reliability of direct augmented reality navigated pedicle screw placement and rod bending. North American Spine Society Journal, 8.
76. Dennler C, Safa NA, Bauer DE, Wanivenhaus F, Liebmann F, Götschi T, Farshad M. (2021) Augmented reality navigated sacral-alar-iliac screw insertion. International Journal of Spine Surgery, 15: 161-168.
77. Frisk H, Lindqvist E, Persson O, Weinzierl J, Bruetzel LK, Cewe P, Burström G, Edström E, Elmi-Terander A. (2022) Feasibility and Accuracy of Thoracolumbar Pedicle Screw Placement Using an Augmented Reality Head Mounted Device. Sensors, 22.
78. Yanni DS, Ozgur BM, Louis RG, Shekhtman Y, Iyer RR, Boddapati V, Iyer A, Patel PD, Jani R, Cummock M, Herur-Raman A, Dang P, Goldstein IM, Brant-Zawadzki M, Steineke T, Lenke LG. (2021) Real-time navigation guidance with intraoperative CT imaging for pedicle screw placement using an augmented reality head-mounted display: a proof-of-concept study. Neurosurgical Focus, 51: 1-8.
79. Urakov TM. (2020) Augmented Reality-assisted Pedicle Instrumentation: Versatility Across Major Instrumentation Sets. Spine, 45: E1622-E1626.
80. xVision. Vol. 2024, Augmedics
81. Hololens. Vol. 2024, Microsoft
82. Spittal MJ, Pirkis J, Gurrin LC. (2015) Meta-analysis of incidence rate data in the presence of zero events. BMC Medical Research Methodology, 15.
83. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Wolters Kluwer Health/Lippincott Williams & Wilkins2008
84. Clopper CJ, Pearson ES. (1934) The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika, 26: 404-413.
85. Endsley MR. (1995) Measurement of situation awareness in dynamic systems. Human Factors, 37: 65-84.

86. Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D. (2006) Development of a rating system for surgeons' non-technical skills. Medical Education, 40: 1098-1104.
87. Drexler DA, Takács Á, Nagy TD, Haidegger T. (2019) Handover process of autonomous vehicles – technology and application challenges. Acta Polytechnica Hungarica, 16: 235-255.
88. Lechappe A, Chollet M, Rigaud J, Cao CGL. (2022) Using Multimodal Data to Predict Surgeon Situation Awareness. In: Lecture Notes in Networks and Systems Vol. 223 LNNS pp. 308-316
89. Bitkina OV, Kim HK, Park J. (2020) Usability and user experience of medical devices: An overview of the current state, analysis methodologies, and future challenges. International Journal of Industrial Ergonomics, 76.
90. Questionnaire for User Interface Satisfaction (QUIS).
91. Olsson T. Concepts and Subjective Measures for Evaluating User Experience of Mobile Augmented Reality Services. In:Huang W, Alem L ,Livingston MA (szerk.), Human Factors in Augmented Reality Environments. Springer New York, New York, NY, 2013: 203-232.
92. Haidegger T, Kazanzides P, Rudas I, Benyó B, Benyó Z. (2010) The importance of accuracy measurement standards for computer-integrated interventional systems. Proceedings of the EURON GEM Sig Workshop on The Role of Experiments in Robotics Research at IEEE ICRA: 19-24.
93. Berman D, Eleswarapu A, Krystal J, Hoang H. (2022) Is the Use of Intraoperative 3D Navigation for Thoracolumbar Spine Surgery a Risk Factor for Post-Operative Infection? Journal of Clinical Medicine, 11.
94. Administration FUSFaD. (2021) Augmedics' xVision Spine system - FDA Clearance. Vol. 2025
95. Pentek M, Haidegger T, Czere JT, Kovacs L, Zrubka Z, Gulacsi L. (2023) EQ-5D studies in robotic surgery: a mini-review. In: SACI 2023 - IEEE 17th International Symposium on Applied Computational Intelligence and Informatics, Proceedings, doi:10.1109/SACI58269.2023.10158630 pp. 519-524
96. Prestes E, Houghtaling MA, Goncalves PJS, Fabiano N, Ulgen O, Fiorini SR, Murahwi Z, Olszewska JI, Haidegger T. (2021) The First Global Ontological

Standard for Ethically Driven Robotics and Automation Systems [Standards]. IEEE Robotics and Automation Magazine, 28: 120-124.

97. Elmi-Terander A, Burström G, Nachabé R, Fagerlund M, Ståhl F, Charalampidis A, Edström E, Gerdhem P. (2020) Augmented reality navigation with intraoperative 3D imaging vs fluoroscopy-assisted free-hand surgery for spine fixation surgery: a matched-control study comparing accuracy. Scientific Reports, 10.

98. Hoeckelmann M, Rudas IJ, Fiorini P, Kirchner F, Haidegger T. (2015) Current capabilities and development potential in surgical robotics. International Journal of Advanced Robotic Systems, 12.

99. Van Dijk JD, Van Den Ende RPJ, Stramigioli S, Köchling M, Höss N. (2015) Clinical pedicle screw accuracy and deviation from planning in robot-guided spine surgery: Robot-guided pedicle screw accuracy. Spine, 40: E986-E991.

100. Khan A, Meyers JE, Siasios I, Pollina J. (2019) Next-Generation Robotic Spine Surgery: First Report on Feasibility, Safety, and Learning Curve. Operative Neurosurgery, 17: 61-68.

101. Jiang B, Pennington Z, Zhu A, Matsoukas S, Karim Ahmed A, Ehresman J, Mahapatra S, Cottrill E, Sheppell H, Manbachi A, Crawford N, Theodore N. (2020) Three-dimensional assessment of robot-assisted pedicle screw placement accuracy and instrumentation reliability based on a preplanned trajectory. Journal of Neurosurgery: Spine, 33: 519-528.

102. Lopez IB, Benzakour A, Mavrogenis A, Benzakour T, Ahmad A, Lemée JM. (2023) Robotics in spine surgery: systematic review of literature. International Orthopaedics, 47: 447-456.

103. Zhang JN, Fan Y, He X, Liu TJ, Hao DJ. (2021) Comparison of robot-assisted and freehand pedicle screw placement for lumbar revision surgery. International Orthopaedics, 45: 1531-1538.

104. Yson SC, Sembrano JN, Sanders PC, Santos ERG, Ledonio CGT, Polly Jr DW. (2013) Comparison of cranial facet joint violation rates between open and percutaneous pedicle screw placement using intraoperative 3-D CT (o-arm) computer navigation. Spine, 38: E251-E258.

105. Tovar MA, Dowlati E, Zhao DY, Khan Z, Pasko KBD, Sandhu FA, Voyadzis JM. (2022) Robot-assisted and augmented reality–assisted spinal instrumentation: a systematic review and meta-analysis of screw accuracy and outcomes over the last decade. Journal of Neurosurgery: Spine, 37: 299-314.
106. Zhang RJ, Zhou LP, Zhang HQ, Ge P, Jia CY, Shen CL. (2022) Rates and risk factors of intrapedicular accuracy and cranial facet joint violation among robot-assisted, fluoroscopy-guided percutaneous, and freehand techniques in pedicle screw fixation of thoracolumbar fractures: a comparative cohort study. BMC Surgery, 22.
107. Garcia D, Akinduro OO, De Biase G, Sousa-Pinto B, Jerreld DJ, Dholakia R, Borah B, Nottmeier E, Deen HG, Fox WC, Bydon M, Chen S, Quinones-Hinojosa A, Abode-Iyamah K. (2022) Robotic-Assisted vs Nonrobotic-Assisted Minimally Invasive Transforaminal Lumbar Interbody Fusion: A Cost-Utility Analysis. Neurosurgery, 90: 192-198.
108. Takacs B, Haidegger T. (2021) Fasttracking Technology Transfer in Medical Robotics. In: 21st IEEE International Symposium on Computational Intelligence and Informatics, CINTI 2021 - Proceedings, doi:10.1109/CINTI53070.2021.9668590 pp. 61-66
109. Malham GM, Munday NR. (2022) Comparison of novel machine vision spinal image guidance system with existing 3D fluoroscopy-based navigation system: a randomized prospective study. Spine Journal, 22: 561-569.
110. Ahmad HS, Yoon JW. (2022) Intra-operative wearable visualization in spine surgery: past, present, and future. Journal of Spine Surgery, 8: 132-138.
111. Tu P, Qin C, Guo Y, Li D, Lungu AJ, Wang H, Chen X. (2022) Ultrasound Image Guided and Mixed Reality-Based Surgical System With Real-Time Soft Tissue Deformation Computing for Robotic Cervical Pedicle Screw Placement. IEEE Transactions on Biomedical Engineering, 69: 2593-2603.
112. Xu DR, Luan LR, Ma XX, Cong ZC, Zhou CL. (2022) Correction: Comparison of electromagnetic and optical navigation assisted Endo-TLIF in the treatment of lumbar spondylolisthesis (BMC Musculoskeletal Disorders, (2022), 23, 1, (522), 10.1186/s12891-022-05443-1). BMC Musculoskeletal Disorders, 23.

113. Franz AM, Haidegger T, Birkfellner W, Cleary K, Peters TM, Maier-Hein L. (2014) Electromagnetic tracking in medicine -A review of technology, validation, and applications. IEEE Transactions on Medical Imaging, 33: 1702-1725.
114. Sirokai B, Kiss M, Kovács L, Benyó B, Benyó Z, Haidegger T. (2012) Best practices in electromagnetic tracking system assessment. Best Practices in Electromagnetic Tracking System Assessment: 1-4.
115. Vasan N, Scherman DB, Kam A. (2022) A tale of two robots: Operating times and learning curves in robot-assisted lumbar fusion. Journal of Clinical Neuroscience, 97: 42-48.
116. Luengo-Matos S, Sánchez-Gómez LM, Hijas-Gómez AI, García-Carpintero EE, Ballesteros-Massó R, Polo-deSantos M. (2022) Efficacy and safety of robotic spine surgery: systematic review and meta-analysis. Journal of Orthopaedics and Traumatology, 23.
117. Cofano F, Di Perna G, Bozzaro M, Longo A, Marengo N, Zenga F, Zullo N, Cavalieri M, Damiani L, Boges DJ, Agus M, Garbossa D, Calì C. (2021) Augmented Reality in Medical Practice: From Spine Surgery to Remote Assistance. Frontiers in Surgery, 8.
118. Nagy TD, Ukhrenkov N, Drexler DA, Takacs A, Haidegger T. (2019) Enabling quantitative analysis of situation awareness: System architecture for autonomous vehicle handover studies. In: Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics Vol. 2019-October pp. 904-908
119. Woodward J, Ruiz J. (2023) Analytic Review of Using Augmented Reality for Situational Awareness. IEEE Transactions on Visualization and Computer Graphics, 29: 2166-2183.
120. Gnjatović M, Maček N, Adamović S. (2020) Putting humans back in the loop: A study in human-machine cooperative learning. Acta Polytechnica Hungarica, 17: 191-210.
121. Meglan D. (2021) Operating room and surgical site awareness. U.S. Patent US 2021/0 157 403 A1, May 27.
122. ZEISS. (2018) ZEISS Colloquium - Innovation Talk. In: Optical challenges for Mixed Reality,

9. BIBLIOGRAPHY

Publications related to the thesis:

Móga K, Hölgyesi Á, Zrubka Z, Péntek M, Haidegger T. Augmented or Mixed Reality Enhanced Head-Mounted Display Navigation for In Vivo Spine Surgery: A Systematic Review of Clinical Outcomes. Journal of Clinical Medicine. 2023; 12(11):3788.
DOI: 10.3390/jcm12113788
IF: 3.0 (Q1, 2023)

Móga K, Ferencz A, Haidegger T. What Is Next in Computer-Assisted Spine Surgery? Advances in Image-Guided Robotics and Extended Reality. Robotics. 2023; 12(1):1.
DOI: 10.3390/robotics12010001
IF: 2.9 (Q1, 2023)

Móga K, Boesl DBO, Haidegger T. Augmented/Mixed Reality Technologies Supporting Digital Surgery. 2021 IEEE 19th International Symposium on Intelligent Systems and Informatics (SISY), Subotica, Serbia, 2021, 183-189.
DOI: 10.1109/SISY52375.2021.9582533.

IF: Conference

Publications not related to the thesis:

Takács K, Takács B, Garamvölgyi T, Tarsoly S, Alexy M, Móga K, Rudas IJ, Galambos P, Haidegger T. Sensor-Enhanced Smart Gripper Development for Automated Meat Processing. Sensors. 2024; 24(14):4631.
DOI: 10.3390/s24144631
IF: 3.4 (Q1, 2023)

Cernelev PD, Móga K, Groves L, Haidegger T, Fichtinger G, Ungi T. Determining boundaries of accurate tracking for electromagnetic sensors. Proc. SPIE 12466, Medical Imaging 2023: Image-Guided Procedures, Robotic Interventions, and Modeling, 124661L (3 April 2023).
DOI: 10.1117/12.2654428

IF: Conference

Elek RN, Mach B, Móga K, Ladi A, Haidegger T. Autonomous Non-Technical Surgical Skill Assessment and Workload Analysis in Laparoscopic Cholecystectomy Training. 2022 IEEE 26th International Conference on Intelligent Engineering Systems (INES), Georgioupolis Chania, Greece, 2022, pp. 000033-000040.
DOI: 10.1109/INES56734.2022.9922657.

IF: Conference

Takács K, Móga K, Haidegger T. Sensorized psychomotor skill assessment platform built on a robotic surgery phantom. In 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI) (pp. 95-100). IEEE.
DOI: 10.1109/SAMI48414.2020.9108730

IF: Conference

∑IF: 5.9

10. ACKNOWLEDGEMENTS

Hereby, I address my gratitude to my:

Topic leaders: Dr. Andrea Ferencz, PhD habil. and Tamás Haidegger, PhD habil.

Leaders and colleagues of the Surgery Departments at Semmelweis University, the Medical Centre of the National Defence Forces, Péterfy Sándor Street Hospital and the Doctoral School of Military Sciences at Ludovica University of Public Service.

Colleagues at Semmelweis and Óbuda Universities (OE EKIK).

Colleagues at PerkLab, Queen's University: Prof. Gábor Fichtinger, Dr. Tamás Ungi, PhD, András Lassó and more.

Co-authors, namely: Prof. Dr. Márta Péntek, Prof. Dr. László Gulácsi, Dr. Zsombor Zrubka, PhD and all others.

Family.

11. SUPPLEMENT

1. User Satisfaction Score (USS) recommended calculation

1.1. USS formula (overall result ranged 0-1)

$$\text{USS} = \sum_{i=1}^{n} w_i \cdot S_i \qquad \text{where} \qquad \sum_{i=1}^{n} w_i = 1$$

S_i - normalised sub-scores $S_i \in [0,1]$ in each subcategory

w_i – recommended weights for each subcategory

1.2. Measurements of sub-categories with recommended weights

#	**Sub-category**	**How to measure**	**Normalise → (S_i)**	**Recommended weight (w_i)**
1	Ergonomics	Subjective Likert (1–7) for short use + long use + score for overheating/sweat Optionally, objective mass (g) as a covariate	average Likert → divide by 7 → S_{erg}	0.2
2	Field of View (FoV)	Likert 1–7 on obstruction/need to shift gaze Objective: horizontal FoV degree normalised vs target (e.g. 100°)	Likert / 7 (or combined with normalised FoVdeg) → S_{fov}	0.15
3	Image Quality	Likert 1–7 on contrast, clarity, registration accuracy Optionally include measured TRE (target registration error) inverted	Likert/7 $1-(TRE/TRE_{max})$ → S_{img}	0.2
4	Control	Likert 1–7 on precision (gaze/gesture/voice) Objective: command success rate (%)	convert to 0–1 (Likert/7 or success rate/100) → S_{ctrl}	0.15
5	Usability & Setup	Combination: SUS score (0–100 → /100) and registration time normalized → composite	$0.7*(SUS/100) + 0.3*(1 - t_{reg}/t_{max})$ → S_{use}	0.18
6	Future-proof, Extensibility Reliability & Safety	Likert 1–7 on API, modularity, upgradability, compatibility with OR workflows Objective failure rate, system downtime, or subjective confidence 1–7	$0.6*(1-\text{fail rate}) + 0.4*(\text{Likert}/7)$ → S_{future}	0.12

Weights sum: 0.2 + 0.15 + 0.2 + 0.15 + 0.18 + 0.12 = 1.00

Higher weight was given to Ergonomics and Image Quality because they strongly determine clinician acceptance in AR/MR-HMD contexts. Usability and Control dimensions are also important.

1.3. How to calculate each sub-category

1.3.1. Ergonomics (S_{erg})

- Collect Likert scores for (the higher the better):
    - short comfort (1–7), long comfort (1–7), overheating/sweat (1–7, 1 = worse, 7 = no complaint).
- Compute the average, then divide by 7.

Example: short = 6, long = 5, overheating = 6

Average=(6+5+6)/3=5.67 → S_{erg}=5.67/7 = 0.81.

1.3.2. Field of View (S_{fov})

- If only subjective: Likert score divided by 7.
- If objective FoV degree is available: normalise as $FoV_{score}=\min(FoV_{deg}/FoV_{target}, 1)$.
- Combine with subjective: $0.6*Likert/7+0.4*FoV_{score}$

1.3.3. Image Quality (S_{img})

- Subjective: average of contrast, clarity, immersion Likert items → divide by 7.
- If Target-Registration Error (TRE in mm) measured: define $S_{TRE}=\max(0, 1-TRE/TRE_{max})$, example: TRE=4 mm, TRE_{max}=8 mm.
- Combine with Likert: $0.6*Likert/7+0.4*S_{TRE}$

1.3.4. Control (S_{ctrl})

- Subjective: gaze, gesture, voice command average Likert scores (1-7), divided by 7.
- If an objective command success rate exists (%): divide by 100 (result: 0–1).

1.3.5. Usability & Setup (S_{use})

- Primary: System Usability Scale (SUS, 5-point Likert scale in 10 statements by John Brooke, 1986) score (0–100 → /100).
- Secondary: registration time t_{reg}: normalise as $1-\min(t_{reg}/t_{max}, 1)$, example: t_{max} = 600 sec (10 min).
- Combine: $S_{use} = 0.7*(SUS/100)+0.3*(1-t_{reg}/t_{max})$

1.3.6. Future-proof (S_{future})

- Likert score average based on the system's APIs, modularity, OR workflow compatibility and update frequency divided by 7.
- Reliability & Safety with fail rate (fraction of sessions with issues, example: 1/10)
- Combine: $S_{future}=0.6*(1-\text{fail rate})+0.4*(\text{Likert}/7)$

1.4 Labelling of AR/MR-HMDs based on USS Range

USS Range	**Interpretation**	**Suggested Label**
0.00–0.40	Low satisfaction	"Needs improvement"
0.41–0.70	Moderate satisfaction	"Acceptable"
0.71–1.00	High satisfaction	"Optimal usability"