

SUPPLEMENTAL DIGITAL CONTENT

Supplemental Table S1
SUMMARY OF XR+ DEVICES, USE CASES, AND ASSESSMENTS FROM ANALYZED ARTICLES

Reference	Specialty	Vis. Type	Display(s)	Task	# of Subs (F)	Procedure Space	Assessment
(Acker et al. 2018)	Neuro	Stereo	DT+glasses	Pl	26 (NR)	N/A	Interrater agreement of feature detection; Detection rates; Relationship of detection rate, experience, and visualization type
(Alexander et al. 2020)	Ortho	AR	DT	Pr	8 (NR)	Phantom	Task error, Task time, Fluoroscopy time, System Usability Scale (SUS), Surg-TLX
(Al-Saud et al. 2017)	O/M	VR	DT+glasses	Pr	16 (13)	Virtual	Error scores, Task completion percentage, Refresh rate
(Allgaier et al. 2022)	General	VR	HMD	Pl	22 (8F, 3NB)	N/A	NASA TLX, SSQ, adapted SUS, Tool preference, Task time, Task error, Task attempts, General feedback
(Alsofy et al. 2021)	Ortho	VR	HMD	Pl	10(NR)	N/A	Potential influence of VR on past cases
(Amini and Kersten-Oertel 2019)	Surg/IR Onc	AR	HMD	Pl	13 (5)	N/A	Task time, SUS
(Andersen et al. 2017)	General	AR	DT	Pr	20 (10)	Phantom	UX, Attention shift, Task error, Task time
(Andersen et al. 2019)	Trauma	AR	DT, HHD	Pr	20 (NR)	Phantom	Idle time, Recall error of procedural steps, Task performance rating
(Andersen et al. 2021)	ENT	VR	DT+glasses	Pl	9(2)	N/A	UX
(Arpaia et al. 2022)	General	AR	HMD	Pr	NR	Phantom, Live human	Display accuracy, System latency, adapted SUS
(Asgar-Deen et al. 2020)	Surg/IR Onc	AR	Mirror	Pr	10(NR)	Phantom	Task error, Task time, UX
(Azar et al. 2004)	Neuro	VR	HMD, DT+glasses	Pr	28 (NR)	Phantom	Task error, Task time, UX
(Ballocca et al. 2018)	Cardio	VR	DT+glasses	Pl	1 (NR)	N/A	Agreement of measurements, Task time, Interobserver variability
(Bartella et al. 2019)	O/M	VR	HMD	Pl	4 (NR)	N/A	UX
(Bartlett et al. 2020)	Endovasc	AR	HMD	Pr	20(NR)	Non-medical	Head rotation, Hand activity, Body posture, Task time, Time spent viewing displays
(Benmahdjoub et al.)	Neuro, O/M	AR	HMD, HUD	Pr	10 (4)	Phantom	Tool location error, Calibration time and error, Overlay registration error, Task error
(Bichlmeier et al. 2009)	Ortho	AR	HMD, DT	Pr	5 (NR) 12 (NR)	Phantom	UX, Task error, Frame rate
(Block, Yablok, and McDonald 1995)	Anesthes	AR	HUD	Pr	11 (NR)	Live human	UX
(Boedecker et al. 2021)	General	VR	HMD	Pl	5(NR)	N/A	SUS
(Bong, Kim, and Park 2017)	Ortho	AR	DT	Pl, Pr	5 (1)	Phantom	System error, Task error
(Bruckheimer et al. 2016)	Cardio	AR	Hologram	Pl	4 (NR)	N/A	Image quality rating, UX, Identification of landmarks
(Brun et al. 2018)	Cardio	AR	HMD	Pl	36 (12)	N/A	Identification of landmarks, Diagnosis, Suggested surgical plan, UX
(Cardin et al. 2007)	Surg/IR Onc	VR	DT+glasses	Pr	28 (14)	Virtual	Task error, Task time
(Cartucho et al. 2020)	Neuro	AR	HMD	Pl	9 (NR)	N/A	UX
(Cattari et al. 2021)	Surg/IR Onc	AR	HMD	Pr	10 (NR)	Phantom	Overlay registration error, Task success rate, Task error, Visual discomfort, Perceived overlay error
(Chen et al. 2014)	Ortho	AR	DT	Pr	8 (NR)	Phantom	Time for re-calibration, Number of x-ray image acquisitions needed
(Coelho et al. 2020)	Neuro, O/M	AR	HHD	Pl	38 (NR)	N/A	UX
(Condino et al. 2021)	Neuro	AR	HMD	Pr	11(~6)	Phantom, Live human	Registration error, Navigation accuracy, Qualitative registration assessment, Template adjustments
(Cutolo, Parchi, and Ferrari 2014)	Ortho	AR	HMD	Pr	10 (NR)	Phantom	Registration error, Refresh rate
(Cutolo et al. 2016)	Ortho	AR	HMD	Pl, Pr	1 (NR)	Phantom	Task error
(Cutolo et al. 2017)	Neuro	AR	HMD	Pl, Pr	12 (NR)	Phantom	User rating, Choice of visualization modality, Task success, Frame rate
(Davrieux et al. 2020)	Surg/IR Onc	AR	HMD	Pr	23 (NR)	Phantom	Task time, Task attempts, Task error

(Deib et al. 2018)	Ortho	AR	HMD	Pr	1 (NR)	Phantom	Visibility of landmarks and materials, Fluoroscopy time, Task time, UX
(Desender et al. 2013)	Endovasc	VR	DT	Pl	24 (NR)	N/A	User rating, Change of plan due to rehearsal, Task time, Fluoroscopy time, Dexterity/Efficiency metrics, Number of angiographs taken
(El-Hariri et al. 2018)	Ortho	AR	HMD	Pr	NR	Phantom	Registration error
(El-Khameesy and El-Wishy 2012)	Surg/IR Onc	VR	various	Pl	20 (NR)	N/A	UX, Planning time
(Enislidis et al. 1997)	O/M	AR	HUD	Pl, Pr	NR	Phantom, Live human	Feasibility
(Erat et al. 2013)	Ortho	AR	DT	Pr	3 (NR)	Animal cadaver	User rating
(Fick et al. 2021)	Neuro	AR	HMD	Pr	NR	Phantom	Registration error
(Fischer et al. 2016)	Ortho	AR	DT	Pr	7 (NR)	Phantom	Task time, Number of x-ray image acquisitions, Task error, Surg-TLX
(Fotouhi et al. 2018)	Ortho	AR	HMD	Pr	4 (NR)	Phantom	Calibration error, Task error
(Freschi et al. 2009)	Surg/IR Onc	VR	DT	Pl, Pr	3 (NR)	Phantom	Task error, Task success rate
(Furtado et al. 2010)	Endovasc	VR	DT	Pr	17 (NR)	Phantom	Task time, Task error, UX
(Galati et al. 2020)	Surg/IR Onc	AR	HMD	Pr	6 (NR)	Live human	UX
(Gao et al. 2021)	Surg/IR Onc	AR	Proj	Pr	10(NR)	Phantom, Animal cadaver	Projection error, Task success rate, UX
(Gavaghan et al. 2012)	Ortho, O/M, Surg/IR Onc	AR	Proj	Pl, Pr	12 (NR)	Live human, Human cadaver, Phantom	UX, Frame rate
(Gerard et al. 2018)	Neuro	AR	DT	Pr	2 (NR)	Live human	UX, Registration error
(Glas et al. 2021)	O/M	AR	HMD	Pr	12 (0)	Phantom	Task error, Task speed, UX
(Gökyar, Bahadır, and Çokluk 2020)	Neuro	AR	Proj	Pr	NR	Animal cadaver	Task success rate
(Gsaxner et al. 2021)	ENT	AR	HMD	Pr	11 (NR)	Phantom	SUS, UX, Training time, Task time
(Gu et al. 2022)	Ortho	AR	HMD	Pr	12 (NR)	Phantom	Detection of registration error
(Guiraudon et al. 2010)	Cardio	VR	DT	Pr	NR	Live animal, Animal cadaver	Feasibility, Task error
(Guo, Cai, and Gao 2018)	Endovasc	VR	DT	Pr	NR	Virtual	Refresh rate, UX
(Hans et al. 1999)	ENT	VR	DT	Pl	17 (NR)	N/A	Image quality rating, Interobserver agreement, User visualization preferences
(Hansen et al. 2010)	General	AR	DT	Pr	6 (NR)	Live human (video), Animal cadaver	UX, Depth perception
(Harake et al. 2020)	Cardio	VR	DT+glasses	Pl	22(NR)	N/A	UX, User preference (Engagement with images based on manipulation and viewing time)
(Harel et al. 2022)	Ortho	AR	HMD	Pr	6 (NR)	Live human	Task accuracy, UEQ, UX
(Heining et al. 2006)	Ortho	AR	DT	Pr	NR	Human cadaver, Phantom	Feasibility, Fluoroscopy time
(Heinrich et al. 2020)	Surg/IR Onc	AR	HMD, DT, Proj	Pr	21(13)	Phantom	Mental workload (Reaction time for secondary task), Task time, Task error, Task success, UX
(Hendel et al. 2021)	Derm/Plastic	VR	HHD	Pl	3(NR)	N/A	Accuracy of generated models; Inter- and intra-investigator reliability of generated models
(Hombeck et al. 2022)	General	VR	DT, HMD	Pl	56 (21)	N/A	User confidence, Task error, Display preferences, Task time
(Huang et al. 2019)	Surg/IR Onc	AR	HMD	Pr	NR	Phantom	Feasibility of shared visualization; Error in location of shared virtual content across HMD users; Task error
(Huettl et al. 2021)	Surg/IR Onc	VR	DT HMD	Pl	30 (15)	N/A	Task success. Task time, Titmus test for stereoacuity, Mental Rotation Test, SUS, SSQ, UX
(Incekara et al. 2018)	Neuro	AR	HMD	Pl, Pr	2 (NR)	Live human	Planning time, Registration error
(Jeon, Choi, and Kim 2014)	Anesthes	AR	Proj	Pr	20 (NR)	Phantom	Task time, Task efficiency

(Jinnin et al. 2011)	Derm/ Plastic	Stereo	Image	Pl	26 (8)	N/A	UX
(Juhnke et al. 2013)	Ortho	VR	DT	Pl	17 (15)	N/A	Task success rate, Task accuracy, Task time, User preferences
(Katic et al. 2015)	O/M	AR	HMD	Pr	NR	Animal cadaver	Calibration time, Calibration error, Task error, UX
(Kendoff et al. 2009)	Ortho	AR	DT	Pr	9 (NR)	Live human	Setup time, UX, Change of intraoperative decisions, Image quality rating, System failure rate, Fluoroscopy time
(Kersten-Oertel, Chen, and Collins 2014)	Neuro	VR	DT+glasses	Pl	35 (10)	N/A	Depth perception, User rating, Refresh rate
(Kim et al. 2019)	Cardio	AR	HHD	Pr	NR	Phantom	Task error
(Kim et al. 2020)	Cardio	VR	Proj, HMD	Pl	22(19)	N/A	User's previous medical and VR experience, Discussion time by case and display type, Accuracy of diagnosis, User preferences, UX
(Kim et al. 2022)	Cardio	VR	HMD	Pl	5 (NR)	N/A	Task time, Outcomes
(Koike et al. 2021)	Neuro	AR	DT	Pr	NR	Live human	Registration error, Task time, Task efficiency
(Kozlowski, Urheim, and Samset 2017)	Cardio	Stereo	DT	Pr	5 (NR)	Phantom	Frame rate, System lag, Task time, UX
(Kubben and Sinlae 2019)	Neuro	AR	HMD	Pr	2(NR)	Operating room	UX, Visualization error, System performance (Recognition of hand gestures)
(Kumar et al. 2020)	Cardio	AR	HMD	Pl	62(7)	N/A	UX, Task success
(Léger et al. 2018)	Neuro	AR	HHD	Pr	17(9)	Phantom	Registration error, UX
(Léger et al. 2020)	Neuro	AR	HHD	Pr	10 (3)	Phantom	Task error, Task time, Registration error, Latency, NASA-TLX
(Lo et al. 2009)	Cardio	VR	DT	Pr	8 (NR)	Phantom	Task error, Task time, NASA TLX
(Londei et al. 2015)	Ortho	AR	DT	Pr	4 (NR)	Phantom	Task error, Task success rate, Number of x-ray acquisitions, Task time
(Ma et al. 2017)	Ortho	AR	Mirror	Pl, Pr	NR	Animal cadaver, Phantom	Calibration error, Registration error, Task error
(Marschollek et al. 2006)	Ortho	VR	DT	Pl	3 (NR)	N/A	UX
(Molina et al. 2019)	Ortho	AR	HMD	Pr	4 (NR)	Human cadaver	Task error, UX
(Moosburner et al. 2018)	General	AR	HMD	Pr	17 (NR)	Virtual	SUS
(Moreta-Martinez et al. 2021)	Surg/IR Onc	AR	HHD	Pl, Pr	10(NR) 2(NR)	Phantom	Registration error, UX
(Napa, Moore, and Bardyn 2019)	Cardio	VR	HMD	Pl	8 (3)	N/A	SUS, NASA TLX, UX
(Nguyen et al. 2020)	Ortho	AR	HMD	Pr	9(NR)	Live human	Display opacity, Color preferences, UX
(Novotny et al. 2020)	Ob	VR	Proj+glasses	Pl	8 (NR)	N/A	Task error, Task time, Image resolution
(Ogawa et al. 2018)	Ortho	AR	HHD	Pr	2 (NR)	Live human	Task error
(Pahuta et al. 2012)	Ortho	VR	DT	Pl	48 (NR)	N/A	Mental Rotations Test, Task error
(Pandey et al. 2021)	Ortho	Stereo	DT	Pr	21 (NR)	Virtual	Task time, Task error, UX,
(Pellegrino et al. 2019)	O/M	AR	HMD	Pr	1 (NR)	Live human	Task error, Task success
(Pepe et al. 2019)	O/M	AR	HMD	Pr	6 (NR)	Phantom	Task error, ISO-9241/110-based questionnaire
(Perkins et al. 2017)	Surg/IR Onc	AR	HMD	Pl	8 (NR)	N/A	Visualization error
(Pfeiffer et al. 2018)	Surg/IR Onc, Cardio	VR	HMD	Pl	77 (NR)	N/A	UX, Feasibility
(Ponce et al. 2016)	Neuro, Ortho	AR	HHD	Pr	31 (NR)	Live human	UX
(Pushparajah et al. 2021)	Cardio	VR	HMD	Pl	5(NR)	N/A	User confidence in decision-making, Revisions to decision-making
(Qian et al. 2017)	Ortho	AR	HMD	Pr	23(4)	Phantom	Text readability, Contrast perception, NASA-TLX, Frame rate, System lag
(Qian et al. 2019)	General	AR	HMD	Pr	20 (5)	Phantom	Task time, Task variability, UX
(Reitinger et al. 2006)	Surg/IR Onc	VR	DT+glasses	Pl	19 (NR)	N/A	Task time, Task success, Task error, Quality of task performance, UX
(Rojas-Munoz et al. 2019)	General	AR	HMD, DT	Pr	20 (6)	Phantom	Task error, Task time, Focus shifts, UX
(Rojas-Muñoz et al. 2020)	Trauma	AR	HMD	Pr	20 (3)	Phantom	Task time, Emergency Cricothyroidotomy Performance scores, Global Rating Scale, Evaluator's Overall Rating, Critical Criteria

(Rojas-Munoz et al. 2020)a	Trauma	AR	HMD	Pr	20 (6)	Human cadaver	Number of task errors, Task time, Global Rating Score, Evaluator Overall Rating, Confidence in independent task recall/performance
(Rolls et al. 2013)	Neuro	AR	DT	Pr	21 (NR)	Phantom	Task and tool efficiency, Task time, Fluoroscopy time, Task error, Task performance rating
(Sadri et al. 2019)	Endovasc	AR	HMD	Pr	>40 (NR)	Virtual, Phantom	Task error, Task time, NASA-TLX, Mental rotation test, Stereo Optical Co. Inc. Stereo Fly Test, Pseudo-Isochromatic Plate Color vision test, UX
(Sampogna et al. 2017)	General	VR	HMD	Pl, Pr	20 (NR)	Phantom	UX
(Saylany et al. 2020)	Ortho	AR	HUD	Pr	1 (NR)	Live human	Task success, UX
(Scherl et al. 2020)	ENT	AR	HMD	Pr	NR	Phantom, Live human	Registration error, UX, Effect of training on registration
(Scherl et al. 2021)	ENT	AR	HMD	Pr	NR	Live human	Registration error, UX, Preparation time, Task time, Complications
(Seitel et al. 2009)	Cardio	VR	DT	Pl	3 (NR)	N/A	Registration error, Visualization error, Intra- and inter-operator variability of task performance, Task time
(Si et al. 2018)	Surg/IR Onc	AR	HMD	Pr	10 (NR)	Phantom	Task error, Planning and setup time, Task time, UX
(Southworth et al. 2020)	Cardio	AR	HMD	Pr	10(NR)	Live human (observation)	Geometric distortion (bench and human testing), Projection error, Similarity of 2D and 3D images, Object positioning accuracy, Color accuracy, Contrast accuracy, Latency, Frame rate, Battery runtime
(Stadie and Kockro 2013)	Neuro	VR	DT+glasses	Pl	20 (NR)	N/A	UX, Refresh rate
(Stamm et al. 2002)	O/M	VR	DT	Pl	1 (NR)	N/A	Feasibility
(Suenaga et al. 2013)	O/M	AR	Mirror	Pr	NR	Live human	Task error
(Suenaga et al. 2015)	O/M	AR	Proj	Pr	NR	Live human	Registration error, Frame rate
(Tabrizi and Mahvash 2015)	Neuro	AR	Proj	Pr	NR	Live human, Phantom	Registration success, Registration time, Visualization error
(Tai et al. 2020)	Neuro	VR	HMD	Pl	1	N/A	Feasibility, Differences in surgical approach
(Testi et al. 2006)	Ortho	VR	DT+glasses	Pl	5 (NR)	N/A	System error, Task time, Task error
(Thabit et al. 2022)	Neuro	AR	HMD	Pl	12 (NR)	Phantom	Task error, UX
(Tian et al. 2015)	Surg/IR Onc	VR	DT	Pl, Pr	1 surgical team	Live human	Visualization error
(Timonen et al. 2020)	ENT	VR	HMD	Pl	5(NR)	N/A	Measurement errors; Bland-Altman agreement of measurement, Face validity, Content validity, UX
(Traub et al. 2006)	Ortho	AR	HMD	Pr	3 (NR)	Human cadaver	Feasibility
(Tsukada et al. 2019)	Ortho	AR	HHD	Pr	1 (NR)	Phantom	Visualization error
(Ungar et al. 2019)	ENT	VR	DT	Pl	1 (NR)	N/A	Feasibility
(Urakov 2020)	Ortho	AR	HMD	Pr	1(NR)	Phantom	Task time, Task completion; Task error, Complications
(Vaghela, Lee, and Akhtar 2018)	Ortho	VR	DT	Pr	25 (NR)	Phantom	Construct validity
(van den Berg et al. 2016)	Derm/Plastic	VR	DT	Pr	NR	Live human	Visualization error
(Vertemati et al. 2019)	General	VR	HMD	Pl	33 (NR)	N/A	UX
(Viehofer et al. 2020)	Ortho	AR	HMD	Pr	2(NR)	Phantom	Task error, Task performance by level of experience
(Vos et al. 2015)	Surg/IR Onc	VR	Proj+ glasses	Pl	2 (NR)	N/A	Interobserver reliability of surgical evaluation, Patient outcome assessment, Surgeon outcome assessment
(Wacker et al. 2005)	Surg/IR Onc	AR	HMD	Pr	NR	Live animal, Phantom	Task error, Task time, UX
(Wang et al. 2012)	Endovasc	AR	DT	Pl, Pr	16 (NR)	Phantom	UX
(Wang, Wang, et al. 2016)	Ortho	AR	HMD	Pr	1 (NR)	Human cadaver	Task error
(Wang, Habert, et al. 2016)	Ortho	AR	DT	Pr	24 (NR)	Virtual	Perception of relative depth, Visualization accuracy, Preferred visualization method, UX

(Wellens et al. 2019)	Surg/IR Onc	VR	HMD	Pl	7 (NR)	N/A	UX, Comparison of virtual and 3D-printed model
(Werkgartner et al. 2004)	Surg/IR Onc	AR	HMD	Pl	NR	N/A	Feasibility
(Willaert et al. 2010)	Neuro	VR	DT	Pl, Pr	7 (NR)	Virtual, Live human	Face validity, UX, Task time, Fluoroscopy time, Tool and material choices
(Willaert, Aggarwal, Daruwalla, et al. 2012)	Neuro	VR	DT	Pl	20 (NR)	Virtual	Fluoroscopy time, Procedure time, Dexterity/Efficiency metrics, Objective Structured Assessment of Technical Skills (OSATS)-derived generic endovascular (Global Rating Scale, GRS) and procedure-specific rating scales (PSRS) for carotid artery stenting procedures (CAS)
(Willaert, Aggarwal, Van Herzeele, et al. 2012)	Neuro	VR	DT	Pl, Pr	Pl: 3 (NR), Pr =<15 surgical teams	Virtual, Live human	Tool choice, Dexterity/Efficiency metrics, Procedure time, Qualitative task performance, Objective Structured Assessment of Technical Skills (OSATS)-derived generic endovascular (Global Rating Scale, GRS) and procedure-specific rating scales (PSRS) for carotid artery stenting procedures (CAS), Non-technical skills (NOTECHS) for surgeons, Mayo High Performance Teamwork Scale (MHPTS), Face validity, Content validity
(Wu et al. 2022)	General	AR	DT	Pr	NR	Live human	Task time, Blood loss, Surgical outcomes
(Yanni et al. 2021)	Ortho	AR	HMD	Pr	11 (NR)	Phantom	Task performance rating, UX
(Ye et al. 2021)	Cardio	AR	HMD	Pl	Team	N/A	Task time, 3D image reconstruction time, Accuracy of diagnosis, Similarity of plan to procedure
(Yoon et al. 2017)	Ortho	AR	HUD	Pr	2(0)	Live human	Task time, UX
(Zawy Alsofy, Sakellaropoulou, and Stroop 2020)	Neuro	VR	DT	Pl, Pr	9 (NR)	Virtual, Live human	Task success, Post-op patient symptoms, Potential influence of VR on past cases
(Zuo et al. 2020)	Endovasc	AR	HMD	Pr	15 (9)	Phantom in OR	Task error, Frame rate, Muscle fatigue, System performance during user movement, NASA TLX, Visualization quality, UX

Ortho = Orthopedic or Spinal Surgery; Neuro = Neurosurgery or Neuro-interventional; Surg/IR Onc = Surgical or Interventional Oncology; General = General Surgery; Cardio = Interventional Cardiology or Cardiac Surgery; O/M = Oral and Maxillofacial Surgery; Endovasc = Endovascular Surgery; Derm/Plastic = Dermatology or Plastic Surgery; Anesth = Anesthesiology; ENT = Otolaryngology; Ob = Obstetrics; Trauma = Emergency Medicine or Trauma Surgery; VR = Virtual Reality; Stereo = Stereoscopic 3D Visualization; HMD = Head-Mounted Display; HHD = Handheld Display; HUD = Head-Up Display; Glasses = Glasses for stereoscopic 3D-visualization; DT = Desktop Display or Monitor; Pl = Planning, Pr = Procedure; UX = User Experience, subjective (e.g. survey, questionnaire, interview or general feedback); NR = Not reported; N/A = Not Applicable