South American Publishing

SAP Gamification and Augmented Reality. 2023;2:2
doi: 10.62486/gr20232

ORIGINAL

Device problems and user-related events reported for head-mounted augmented reality navigation systems in United States surveillance data

Problemas de dispositivo y eventos de uso notificados para navegación con realidad aumentada de visor cefálico en Estados Unidos

Maria Carla Hernández Martínez[1]

[1] Independent Researcher. Buffalo, New York

Submitted: 16-02-2023 | Revised: 23-04-2023 | Accepted: 23-07-2023 | Published: 26-07-2023

Corresponding Author: Maria Carla Hernández Martínez

ABSTRACT

Introduction: Augmented reality head-mounted displays are entering surgical practice, but their post-market safety profile is undocumented. The objective was to describe the device problems and user-related events reported for these systems in the United States device surveillance database. Methods: Retrospective descriptive study of spontaneous device reports received between 2010 and 2023 and retrieved from the openFDA device event endpoint. Reports were screened for augmented or mixed reality systems with a head-mounted display and analysed by dimension: reporting volume, device component, event type and outcome, user-related events, and relations. Results: Of 490 distinct reports, 44 (9.0%) met the criteria, grouped in 25 events; 41 concerned one spinal navigation system. No report preceded 2020, and none concerned mixed reality headsets, training or wearer symptoms. Injury events accounted for 77.3% of reports (95% CI 63.0-87.2). A user-related cause was stated in 9 events (36.0%) and reference movement in 11. User-related causes were associated with reference movement (odds ratio 8.3; 95% CI 1.4-50.0). Conclusions: Reported problems concerned the registration chain and handling, not the headset display or the wearer. The record is narrow, manufacturer-led and cannot yield event rates.

Keywords: augmented reality; medical device adverse events; surgical navigation; head-mounted display; use error.

RESUMEN

Introducción: Los visores de realidad aumentada montados en la cabeza se incorporan a la práctica quirúrgica, pero su perfil de seguridad posterior a la comercialización no está documentado. El objetivo fue describir los problemas de dispositivo y los eventos relacionados con el usuario notificados para estos sistemas en la vigilancia estadounidense. Métodos: Estudio descriptivo retrospectivo de notificaciones espontáneas recibidas entre 2010 y 2023 y recuperadas de openFDA. Se cribaron los sistemas de realidad aumentada o mixta con visor en la cabeza y se analizaron por dimensión: volumen, componente, tipo de evento y desenlace, eventos de uso y relaciones. Resultados: De 490 notificaciones distintas, 44 (9,0%) cumplieron los criterios y se agruparon en 25 eventos; 41 correspondieron a un sistema de navegación espinal. Ninguna precedió a 2020 y ninguna correspondió a visores de realidad mixta, formación o síntomas del portador. Los eventos con lesión representaron 77,3% de las notificaciones. Se declaró causa relacionada con el usuario en 9 eventos (36,0%) y movimiento de la referencia en 11. Las causas de uso se asociaron con el movimiento de la referencia (razón de momios 8,3; IC 95 % 1,4-50,0). Conclusiones: Los problemas notificados afectaron a la cadena de registro y a la manipulación, no a la pantalla del visor ni al portador. El registro es estrecho, dependiente de los fabricantes y no permite estimar tasas.

Palabras clave: realidad aumentada; eventos adversos de dispositivos médicos; navegación quirúrgica; visor montado en la cabeza; error de uso.

© 2023; The Authors. This Open access article is distributed under CC BY 4.0 licence.

INTRODUCTION

Augmented reality overlays computer-generated content on the user's direct view of the real world, keeps the overlay registered to that world in three dimensions and responds in real time.(1) When the display is worn on the head, the surgeon sees the navigation image in the line of sight instead of turning to a monitor. Systematic reviews describe head-mounted augmented and mixed reality as the most common intraoperative interface in spine surgery, alongside microscope and projector interfaces.(2,3) Their pooled clinical outcomes for pedicle screw guidance resemble those reported for robot-assisted systems.(4)

Clinical series with a commercial headset system report screw accuracy between 97% and 100% on the Gertzbein-Robbins scale.(5-9)Those series are small, are reported by early adopters and focus on accuracy, so they say little about the problems that arise when many surgeons use the systems in routine operating rooms. One of them recorded two procedures not completed because of the intraoperative scanner,(5)and another recorded four screws revised intraoperatively out of 222.(8)Neither was designed to classify failures by their cause.

The headset adds problems that conventional navigation does not have. Usability reviews, in-theatre performance tests and ergonomic analyses describe sensitivity to lighting, registration time, display latency and head-worn weight.(10-13)Cybersickness is documented for virtual and mixed reality headsets in general.(14,15) Mixed reality headsets are also tested in surgical education and simulated training,(16-18)where the unit is not necessarily a regulated patient device. Whether any of these problems reaches the post-market record is not known.

Spontaneous adverse event reports are the main post-market source for medical devices in the United States. Analyses of the Manufacturer and User Facility Device Experience database (MAUDE) have mapped failures of surgical robots, image-guided sinus navigation, robotic arthroplasty systems and skull clamps.(19-22)The same literature documents the weaknesses of the source, which include free-text narratives, voluntary reporting and the absence of denominators.(23-25)The literature search for this study did not identify a MAUDE analysis restricted to augmented or mixed reality head-mounted displays. The objective was to describe the device problems and user-related events reported for augmented and mixed reality head-mounted displays in the MAUDE database between 2010 and 2023, organised by five dimensions: reporting volume and source, device component, event type and outcome, user-related events and attributed causes, and the relations among them.

METHODS

Design and data source

The study was a retrospective descriptive analysis of spontaneous medical device reports. Reporting followed the STROBE statement and its extension for routinely collected data.(26,27) The source was the open device adverse event endpoint of openFDA, which publishes the medical device reports received under the federal medical device reporting regulation.(28,29) Reports received between 2010 and 2023 were eligible. The endpoint is public, requires no login and was queried programmatically.

Augmented reality was defined as a system that combines real and virtual content, is interactive in real time and registers the virtual content in three dimensions.(1)A report was eligible when its device was a component of such a system whose virtual content is shown on a head-mounted display, either a see-through headset or smart glasses. Mixed reality headsets were eligible under the same registration criterion. The search combined 34 queries on the narrative text, brand name and generic name fields. The terms covered augmented, mixed and virtual reality, head mounted, smart glasses, headset, glasses, heads-up display, holographic, exoscope and the abbreviation HMD, plus the brand names of the two systems that the first queries identified. Each query was paged at 100 reports with a date filter on the receipt date, and the pages were stored unchanged so that screening and analysis could be rerun on the same files. The endpoint returns narratives in upper case, and the scripts lowercased them before applying text rules. A further 33 count probes searched the three fields for six consumer and enterprise headset brand names and five generic headset phrases.

Screening used the device name, generic name and narrative. Reports were excluded when the match was lexical or incidental, for example an ophthalmic Latin term, a femoral head mounted on an implant, an audio headset or a catalog prefix. Virtual reality devices and video magnifiers

without a see-through overlay were excluded because they are not augmented or mixed reality. The reasons for exclusion were counted by rule in the analysis script.

Reporting volume and source

The variables were the year of receipt, the year of the event, the report source code, the reporter occupation, the device operator and the remedial action. Manufacturers often file one report per device component for a single event, so a distinct event was defined as a set of reports sharing the first 200 characters of the narrative. Counts and proportions are presented per report and per distinct event, with Wilson 95% confidence intervals.(30)

Device component

The component was classified from the model field of each report, because the brand name was common to all reports of a system. Four classes were used: headset, computer and software, tool or pin adapter, and other system component. Reports for the same event were checked for mixed component classes. Counts were cross-tabulated against the system and the event type.

Event type and outcome

The event type was the value coded in the report (injury, malfunction or death). Under the reporting regulation, a serious injury includes an injury that requires medical or surgical intervention to preclude permanent impairment, and a malfunction is reportable when it would be likely to cause or contribute to a serious injury if it recurred.(29)The FDA product problem codes and patient problem codes were tallied once per report. The outcome of each distinct event was coded from the narrative into five exclusive classes: neurological symptoms or deficit, implant placement corrected intraoperatively without harm, reference or registration problem corrected without implant revision, component detached or broken without harm, and bone fracture.

User-related events and attributed causes

The cause named or suspected in the narrative was coded for each distinct event into five non-exclusive categories: reference fixation, marker or camera movement; tool, screw interface, calibration or skiving; patient factors such as body mass, bone quality, anatomy and prior screw tracks; hardware component failure; and root cause not concluded. Reference fixation covered the clamp, pin, marker and camera that anchor the patient coordinate system. Tool, screw interface, calibration or skiving covered a loose or bent instrument interface, a calibration that did not follow the manual and the deviation of a tool at the entry point. A separate flag recorded a user-related cause, defined as user error, misuse or non-adherence to the instructions for use stated or concluded in the manufacturer narrative; a hedged mention of a possible user error was not coded. A text rule on the narrative terms checked the flag, and the number of events flagged by the rule but not coded is reported. Keyword searches counted narratives that described wearer symptoms, headset or display faults, and training or simulation settings.

Relations among dimensions

Two relations were specified before the data were examined: user-related cause by reference fixation, marker or camera movement, and neurological symptoms by patient factors. Each was tested with the two-sided Fisher exact test on distinct events, with the Holm adjustment over the two tests. Odds ratios used the Haldane-Anscombe correction and a Woolf interval. The analysis is exploratory and has no power calculation. Cross-tabulations of the five causes by user-related cause, implant correction and neurological symptoms complete the dimension.

RESULTS

Reporting volume and source

In the dimension reporting volume and source, the queries returned 490 distinct reports, and 44 of them (9.0%) described an augmented reality system with a head-mounted display. The included reports comprised 41 for the xVision Spine System and 3 for the NextAR system, grouped in 22 and 3 distinct events. The 18 probes on six headset brand names returned 0 reports (Table 1).

Reports presented a short history. The first was received in 2020, and none was received in

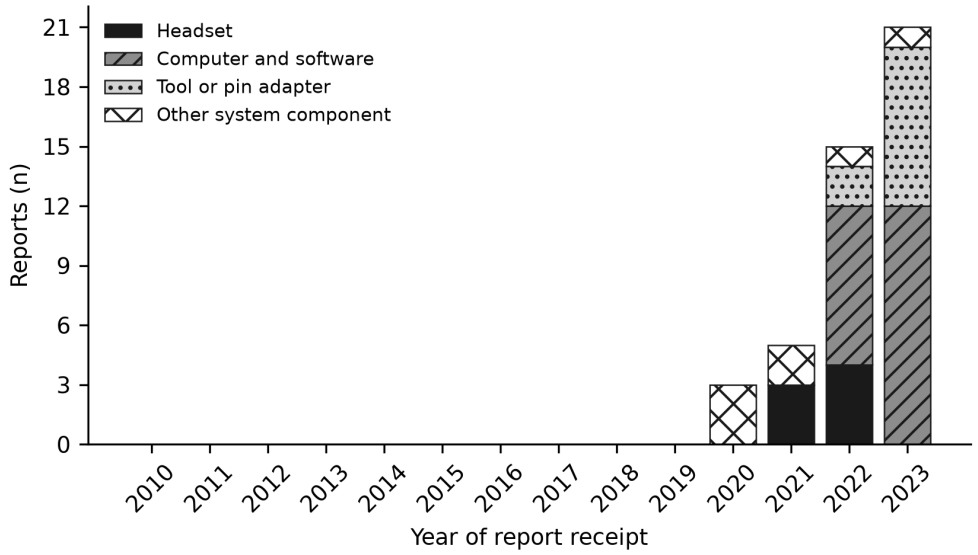
the 10 earlier years of the window. The year 2023 obtained 21 reports against 3 in 2020, and the last two years concentrated 81.8% of the total (95% CI 68.0-90.5) (Figure 1).

Table 1. Screening of retrieved reports and source of the included reports.

Characteristic	Reports, n (%)
Distinct reports retrieved	490
Included (augmented reality systems with a head-mounted display)	44 (9.0)
Excluded: Other lexical match (acronyms, catalog prefixes, other terms)	359 (73.3)
Excluded: Ophthalmic Latin term (oculus)	47 (9.6)
Excluded: Femoral or screw head mounted on an implant	17 (3.5)
Excluded: Augmented, mixed or virtual reality mentioned in the narrative of a device without a head-mounted display	9 (1.8)
Excluded: Headset without a display (audio, hearing protection, skull clamp system)	8 (1.6)
Excluded: Virtual reality or video magnifier head-mounted device, not augmented or mixed reality	3 (0.6)
Excluded: Consumer head-worn display near an implanted neurostimulator	3 (0.6)
Included reports by source	
Manufacturer report	43 (97.7)
Distributor report	1 (2.3)
Reporter occupation, physician	38 (86.4)
Reporter occupation, other or coded numerically	6 (13.6)
Device operator, health professional	43 (97.7)

Note: n (%) of the retrieved reports for screening rows and of the included reports for source rows. Percentages use one decimal.

Figure 1. Reports received per year by device component.



Manufacturers filed 43 reports (97.7%) and a distributor filed one. Physicians were the reporter occupation in 38 reports. In relation to the other dimensions, 19 of the 25 events were filed twice with the same narrative, and no event mixed component classes (0 events). No narrative described training or simulation (0 matches), and all 25 described surgical procedures.

Device component

For this variable, the model field named a computer and software unit in 20 reports (45.5%) and a tool or pin adapter in 10. A headset was named in 7 reports, and another system component in 7 (Table 2).

Table 2. Device component named in the included reports by system and event type.

Component named in the report	xVision, n	NextAR, n	Total, n (%)	Injury, n	Malfunction, n
Headset	7	0	7 (15.9)	7	0
Computer and software	20	0	20 (45.5)	20	0
Tool or pin adapter	10	0	10 (22.7)	2	8
Other system component	4	3	7 (15.9)	5	2
All components	41	3	44 (100.0)	34	10

Note: Components were classified from the model field of each report. Percentages refer to 44 reports.

Headsets obtained 15.9% of the reports (95% CI 7.9-29.4), and all 7 headset reports described injury events. Adapters concentrated the malfunctions: 8 of their 10 reports were malfunctions. The 11 distinct model names belonged to 41 xVision and 3 NextAR reports.

In relation to the event dimension, the narratives of the headset reports described screw misplacement after reference movement or tool calibration, and none described a display fault, a discomfort or a visual symptom (0 events matched the symptom terms). Headset or smart glasses terms appeared in 1 narrative, where the unit was reviewed without anomaly.

Event type and outcome

Table 3. Event type, outcome and FDA problem codes of the included reports.

Characteristic	Reports (n = 44), n (%)	Distinct events (n = 25), n (%)
Event type, injury	34 (77.3)	19 (76.0)
Event type, malfunction	10 (22.7)	6 (24.0)
Outcome: Neurological symptoms or deficit reported	NA	3 (12.0)
Outcome: Implant placement corrected intraoperatively, no harm reported	NA	14 (56.0)
Outcome: Reference or registration problem corrected, no implant revised	NA	2 (8.0)
Outcome: Component detached or broken, no harm reported	NA	5 (20.0)
Outcome: Bone fracture	NA	1 (4.0)
FDA problem code: Use of Device Problem	28 (63.6)	NA
FDA problem code: Insufficient Device Problem Information	8 (18.2)	NA
FDA problem code: Improper or Incorrect Procedure or Method	3 (6.8)	NA
FDA problem code: Computer Software Problem	1 (2.3)	NA
FDA problem code: Unintended Movement	1 (2.3)	NA
FDA problem code: Difficult or Delayed Positioning	1 (2.3)	NA
FDA problem code: Mechanics Altered	1 (2.3)	NA
FDA problem code: Retraction Problem	1 (2.3)	NA
FDA problem code: Adverse Event Without Identified Device or Use Problem	1 (2.3)	NA
FDA problem code: Fracture	1 (2.3)	NA
FDA problem code not assigned	0 (0.0)	NA
FDA patient problem code: No Clinical Signs, Symptoms or Conditions	39 (88.6)	NA
FDA patient problem code: Spinal Column Injury	2 (4.5)	NA
FDA patient problem code: Insufficient Information	1 (2.3)	NA
FDA patient problem code: No Known Impact Or Consequence To Patient	1 (2.3)	NA
FDA patient problem code: Bone Fracture(s)	1 (2.3)	NA
Recall as remedial action	5 (11.4)	3 (12.0)

Note: NA = not applicable.

Within this component, injury events accounted for 34 reports (77.3%) and malfunctions for 10; no death was reported (Table 3). The problem code Use of Device Problem was assigned to 28 reports (63.6%), and the patient code No Clinical Signs, Symptoms or Conditions to 39 (88.6%).

At the level of distinct events, injury obtained 76.0% (95% CI 56.6-88.5). Implant placement was corrected intraoperatively without harm in 14 events (56.0%; 95% CI 37.1-73.3), and neurological symptoms were reported in 3 (12.0%; 95% CI 4.2-30.0). A component detached or broke in 5 events, and one event ended in a bone fracture.

In relation to the component dimension, hardware failure codes carried 5 of the 6 malfunction events, against 1 injury event. Recall was recorded as remedial action in 3 events, all of them adapter malfunctions.

User-related events and attributed causes

In the dimension user-related events and attributed causes, reference fixation, marker or camera movement was the most frequent attributed cause, with 11 events (44.0%; 95% CI 26.7-62.9). Patient factors followed with 9 events and tool, screw interface, calibration or skiving with 7. Hardware failure was attributed in 6 events, and 4 narratives stated that the root cause was not concluded (Table 4).

Table 4. Causes attributed in the narratives of distinct events.

Attributed cause	Events, n (%)	95% CI	Reports, n
Reference fixation, marker or camera movement	11 (44.0)	26.7-62.9	19
Tool, screw interface, calibration or skiving	7 (28.0)	14.3-47.6	13
Patient factors (body mass, bone quality, anatomy)	9 (36.0)	20.2-55.5	15
Hardware component failure	6 (24.0)	11.5-43.4	11
Root cause not concluded in the narrative	4 (16.0)	6.4-34.7	6
User-related cause stated by the manufacturer	9 (36.0)	20.2-55.5	16
No system malfunction found, stated explicitly	4 (16.0)	NA	8
Wearer symptoms described	0 (0.0)	NA	NA

Note: CI = confidence interval (Wilson). Categories are not mutually exclusive. NA = not applicable.

A user-related cause was stated in 9 events (36.0%; 95% CI 20.2-55.5), and all of them concerned the xVision system (9 events). The narratives stated explicitly that no system malfunction was found in 4 events. More than one cause was coded in 9 events, and the text rule flagged 1 hedged mention of user error that was not coded.

Within this component, no narrative described a wearer symptom, a display failure or a latency problem. The causes named the registration chain (clamp, pin, marker or camera), the instrument interface and the patient, and not the headset optics.

Relations among dimensions

In the dimension relations, user-related causes were stated in 7 of the 11 events with reference fixation or marker movement (63.6%) and in 2 of the 14 remaining events (14.3%) (Table 5).

Table 5. Attributed cause by user-related cause, implant correction and neurological symptoms.

Attributed cause	Events, n	User-related cause, n (%)	Implant corrected, n (%)	Neurological symptoms, n (%)
Reference fixation, marker or camera movement	11	7 (63.6)	9 (81.8)	1 (9.1)
Tool, screw interface, calibration or skiving	7	2 (28.6)	7 (100.0)	1 (14.3)
Patient factors (body mass, bone quality, anatomy)	9	2 (22.2)	8 (88.9)	3 (33.3)
Hardware component failure	6	1 (16.7)	1 (16.7)	0 (0.0)
Root cause not concluded in the narrative	4	1 (25.0)	4 (100.0)	2 (50.0)
All events	25	9 (36.0)	17 (68.0)	3 (12.0)

Note: Implant corrected = implant placement corrected intraoperatively, including the three events with neurological symptoms. Fisher exact tests are reported in the text.

The first prespecified test gave an odds ratio of 8.3 (95% CI 1.4-50.0; Fisher p = 0.017; Holm-adjusted p = 0.034). Neurological symptoms were reported in 3 of the 9 events with patient factors (33.3%) and in 0 of the 16 events without them (odds ratio 17.8; 95% CI 0.8-394.1; Fisher p = 0.037; Holm-adjusted p = 0.037).

Among the cross-tabulated causes, the hardware category obtained the lowest implant correction, with 1 of 6 events, and the category root cause not concluded the highest neurological share, with 2 of 4 events.

DISCUSSION

This study set out to describe the device problems and user-related events reported for augmented and mixed reality head-mounted displays. The record that answers the question is narrow. It describes one spinal navigation system and, to a smaller extent, one knee, shoulder and spine system, and the problems it documents lie in the registration chain and in handling, not in the headset itself. The discussion follows the dimensions of the Methods and ends with a joint reading.

Reporting volume and source

The absence of reports before 2020 and the rise in the last two years follow the commercial life of the systems more than any change in safety. Surgical series with a commercial headset system started to appear in 2021,(7,31) and adoption reports run to 2023.(8,9) The same pattern, with peaks that follow adoption and learning, was found for the da Vinci system in MAUDE, where the proportion of console and cart malfunctions changed with surgeon experience and technology generation.(32)Counting reports against time is therefore not a measure of risk, because no denominator of procedures is available.

The dominance of manufacturer reports shapes what is visible. Manufacturers file when their own investigation finds a reportable serious injury or malfunction, and the narratives in this record contain the manufacturer's investigation, including the statement that a redirected screw is a serious injury under the regulation. Clinical series reach the opposite impression. A prospective cohort recorded no intraoperative adverse events with the same class of system while revising four of 222 screws,(8)and an accuracy series reported 98.5% of screws within the clinically acceptable zone.(6)The two sources do not conflict, because a revised screw is a correction in a clinical series and a reportable injury in the regulation. Studies of MAUDE share this property, and reviews of the database find that most studies built on it are case series of poor quality, chiefly because of its structural limits as passive surveillance with free-text reports.(23-25)

The window of the search also shapes the record. Research applications of general-purpose headsets, such as those in cadaver and visualisation studies,(33,34) are not marketed medical systems, and no manufacturer carries a reporting duty for them. A search of narrative text, brand and generic names finds a device only when its report carries one of the terms, and the probes for headset brand names returned none. The absence of reports before 2020 is therefore the absence of a regulated, marketed product with a surveillance obligation, and it should not be read as ten years of problem-free use.

The finding that no report described training or simulation is consistent with the regulatory scope. Mixed reality headsets are tested in training and surgical education,(16-18)but a unit used to teach is not a patient-contact device and has no event to report. The absence of reports from training therefore says nothing about the safety of training with these devices, which is a reading the data do not allow.

Device component

Headsets were named in a minority of reports, and no narrative attributed the problem to its optics or display. This does not mean the headset is free of problems. Experimental work documents registration time that is longer in the operating room than in the laboratory and a strong influence of direct illumination on performance,(11)worse motor performance at higher display latency,(12)and ergonomic requirements on weight and balance.(13)Wireless headsets that remove head turning can still lengthen task completion, as in a randomised arthroscopy simulation.(35)The wearer's side of the device has also been associated with benefits, such as lower surgeon fatigue with smart glasses in scoliosis surgery.(36)A surveillance system built on device failure and patient injury is not designed to capture these effects, and a surgeon who finds the display tiring has no obvious route to report it.

The FDA problem codes show a second gap, which is granularity. The code Use of Device Problem was assigned to most reports, including events whose narratives state that the system displayed the tracked tool correctly and that the fault lay in the clamp, the calibration or the entry point, and the code Insufficient Device Problem Information was assigned to adapter malfunctions whose narratives describe the detached part precisely. Codes are chosen by the reporter from a fixed list, and the narrative carries the information that the code lacks. Analyses that rely on codes alone would classify these events as generic use problems, which is why the narratives were coded here. A framework that separates device problems from patient injuries, developed for orthodontic aligners, shows how a structured reading of the same fields can be organised.(37)

Studies that measure the wearer's experience directly give a different picture from the surveillance record. A pilot of an augmented reality knee navigation system viewed through smart glasses assessed motion sickness with a questionnaire immediately after surgery,(38) and virtual reality sickness is documented across headset types and mitigated by specific strategies.(14,15) None of this reached the reports, which suggests that wearer symptoms are measured in studies and not reported as device events.

The probes for consumer and enterprise headset brands returned no report. Studies compare two mixed reality headsets for visualising radiological content and evaluate a successor model for orthopaedic oncological interventions,(34,39) and a cadaver study used a commercial mixed reality headset for pedicle screw navigation.(33)These are research uses or uses of general-purpose hardware under a specific application, and a manufacturer of the hardware has no reporting duty for them. The consequence is that the systems most discussed in the literature are the least visible in the surveillance record.

Event type and outcome

The share of injury events differs from other MAUDE analyses. In robot-assisted surgery, about three-quarters of reports were malfunctions and 13.1% were injuries,(19)and in image-guided sinus navigation 97.1% were malfunctions and 2.9% involved patient adverse events.(20)Skull clamps showed 73.2% malfunctions,(22)and an analysis of angiographic equipment classified two-thirds of events as near misses.(40)In the present record, injuries predominated because the reports were driven by the manufacturer's rule that a screw redirected after a confirmation scan is a serious injury, and because most events were screw misplacement detected and corrected during the same operation. The harm reported was therefore mostly a correction, with no clinical signs recorded in the large majority of reports. Robotic arthroplasty offers a comparable picture, with surgical delay and conversion to manual surgery as the usual consequences and only a few reinterventions among the injuries.(21)

The classification of an event as an injury deserves a caution. In most injury reports the patient code was No Clinical Signs, Symptoms or Conditions, so the injury category reflects the regulatory definition and not a clinical consequence. A reader who compares the injury share of this record with that of a record dominated by device malfunctions, such as the sinus navigation or skull clamp analyses,(20,22) would conclude that augmented reality navigation is more harmful, and the data do not support that conclusion. The reverse comparison is equally unsafe: the low injury share of robotic surgery per procedure was computed with a denominator of procedures that this record lacks.(19)

Three events with neurological symptoms deserve separate attention, because they show that the correction of a misplaced screw does not always remove its consequences. One event ended in foot drop and revision surgery. In the other two, symptoms were reported as improving. The same range of consequences is described for robot-assisted placement, where unsatisfactory screws occur even in series with high accuracy.(41-43)Malfunction events, in contrast, were adapter failures in which a spring, ball or trigger detached, with no harm reported. The comparison with robotic arthroplasty, where a retained registration checkpoint was the most frequent injury for hip procedures,(21)suggests that fragments of navigation hardware left in the field are a recurring failure class across navigation technologies.

User-related events and attributed causes

Reference movement is the cause attributed most often, and it is also the best documented in the navigation literature. Navigation error grows with the distance from the dynamic reference frame and with vertebral motion from manipulation and respiration, with a mean error of 2 mm or more at two or more levels from the frame.(44) A drift phenomenon occurred in 16.7% of 341 deformity patients navigated with an intraoperative scanner, and the distance from the reference

frame was an independent risk factor.(45) Perforation rates also rose with the distance from the frame in scoliosis surgery.(46,47) The narratives in the present record describe the same mechanism in clinical terms: a clamp not fully closed, a marker touching the skin, a bump on the pin and respiration during registration. This agreement suggests that the headset inherits the error structure of optical navigation, which the headset does not remove.

The failure of the clamp has a counterpart in other surgical fields. Slippage accounted for nearly half of the skull clamp malfunction reports in a ten-year MAUDE analysis, next to mechanical failure,(22)and the unexpected movement of the robotic arm was the most frequent adverse event of robotic knee arthroplasty.(21)In each case the element that anchors the patient to the navigation frame, or the element that moves relative to it, is the point of failure. The headset, which only displays the result, sits downstream of that chain, and an accurate display of a misregistered position would be reported as a use or registration problem, not as a display fault.

The patient factors in the narratives, namely high body mass, soft bone and prior screw tracks, coincide with some but not all of the literature. Obesity was an independent risk factor for unsatisfactory robot-assisted placement in one series,(41)whereas in a multicentre study obesity and other patient variables contributed little to the variation in accuracy.(48)The discrepancy may come from the outcome measure, which was unsatisfactory placement in one study and offset from the planned screw in the other, and from the system. The present record cannot arbitrate between them, because it has no unaffected cases for comparison.

Manufacturer attribution of the cause to the user needs caution. The manufacturer investigates its own device, and the narratives reach the conclusion of user error with different degrees of certainty. Other safety analyses find the same ambiguity. In machine learning devices reported to MAUDE, problems of use were 7% of events but four times more likely to cause harm,(49) and in da Vinci malfunction reports 18% of injuries were attributed to misuse by surgeon or staff.(50) Use errors tied to interface design are documented for electrosurgical units,(51) and a plea for manufacturers to assess rather than dismiss user complaints argues that device-use reports hide design signals.(52) The recommendation to mine MAUDE and recall data for use issues (53) applies here: the repeated clamp and calibration errors in the narratives are design opportunities, such as a closure indicator for the clamp and a calibration that detects a loose screw interface.

Relations among dimensions

The association between reference movement and user-related cause was the more precise of the two tests, and it survived the adjustment. It is partly built into the coding, because movement of a clamp that the surgeon left loose is a user-related cause by definition, and the manufacturer's narrative is the only source. The association with neurological symptoms was imprecise, with an interval that includes values from no association to very large ones, and all three events with symptoms had patient factors attributed. With 25 events, the tests are signals to be followed, in line with the use of spontaneous reports for signal detection and not for confirmation.(54)

The coded categories are not independent of each other. A narrative that attributes the event to a loose clamp also tends to describe the surgeon's handling, and a narrative that attributes it to soft bone also tends to mention the pin. The cross-tabulation shows which categories coincide with user-related causes, implant correction and neurological symptoms, and it does not separate their effects. The hardware category stands apart, because it coincides with malfunctions without patient harm and not with screw correction, which supports the separation between failures of components and failures of the workflow.

Joint reading, strengths and limitations

Read together, the dimensions describe a safety problem of the navigation workflow carried by a headset, not a headset problem. The headset appears in the reports as a labelled component, and the events arise from clamps, pins, markers, instrument interfaces and patients. The inference holds only for the systems found.

The strengths are the transparent screening, the public source, the deduplication of events and the script that reproduces every figure. The limitations are specific. First, passive reporting and the absence of denominators prevent rates, and a procedure registry linked to the reports would allow them. Second, the narratives are written by manufacturers, so independent adjudication of a sample by clinicians would test the attribution. Third, 41 of 44 reports concern one system, and the findings should not be generalised to other vendors or to mixed reality headsets; surveillance of those would need searches by manufacturer and product code beyond the text terms used here. Fourth, a text search in a free-text field can miss reports filed under generic names, and the probes

reduce but do not remove that risk. Fifth, the number of events limits the relations to exploratory status, and the Holm adjustment covers only the two prespecified tests. A related framework for structuring device problems and patient injuries in MAUDE, developed for another device family, offers a template for that repetition.(37)

The practical implications come from the corrective actions stated in the narratives. Surgeons were instructed to close the clamp fully, to verify registration on a bone landmark and not on the skin or the tool, to re-scan after any bump or suspected marker movement and to follow the calibration method for swivelled instruments. Hospitals should add adapter inspection to the sterile processing routine, since several adapters at one site had damaged parts. For teaching, the record has a different message. The studies of headset-based training report feasibility and acceptance in small samples,(16-18)and the surveillance record contains no training event. Programmes that adopt headsets for teaching should therefore define their own incident log, because the regulatory system will not capture it, and should include display latency, lighting and wearer symptoms among the items logged, since the experimental literature links the first two to performance loss and documents the third in headset use.(11,12,14)

For Latin America, where hospitals that adopt these systems depend on local vigilance systems, the transferability of the record is limited: the content of local reports may differ from the federal reports analysed, and no regional evidence was analysed. Local registries of use problems that include the navigation workflow, and not only device failure, would give the regional view that this record lacks.

CONCLUSIONS

Reports for augmented reality head-mounted display systems appeared only from 2020, came from manufacturers and described surgical use, with no report of training use.

The device component named was most often the computer and software unit or a tool adapter, and no report concerned a headset display fault or a wearer symptom.

Injury events predominated, they consisted of screw placement corrected during the operation, and malfunctions were detached adapter parts without harm.

User-related and attributed causes centred on the fixation and movement of the reference, followed by patient factors, the instrument interface and hardware failure.

Reference movement was associated with user-related causes and patient factors with neurological symptoms, and both associations are exploratory.

REFERENCES

1. Azuma RT. A Survey of Augmented Reality. Presence: Teleoperators & Virtual Environments. 1997;6(4):355-385. doi: 10.1162/pres.1997.6.4.355
2. McCloskey K, Turlip R, Ahmad HS, Ghenbot YG, Chauhan D, Yoon JW. Virtual and Augmented Reality in Spine Surgery: A Systematic Review. World Neurosurgery. 2023;173:96-107. doi: 10.1016/j.wneu.2023.02.068
3. Ghaednia H, Fourman MS, Lans A, Detels K, Dijkstra H, Lloyd S, et al. Augmented and virtual reality in spine surgery, current applications and future potentials. The Spine Journal. 2021;21(10):1617-1625. doi: 10.1016/j.spinee.2021.03.018
4. 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. JCM. 2023;12(11):3788. doi: 10.3390/jcm12113788
5. Harel R, Anekstein Y, Raichel M, Molina CA, Ruiz-Cardozo MA, Orrú E, et al. The XVS System During Open Spinal Fixation Procedures in Patients Requiring Pedicle Screw Placement in the Lumbosacral Spine. World Neurosurgery. 2022;164:e1226-e1232. doi: 10.1016/j.wneu.2022.05.134
6. Liu A, Jin Y, Cottrill E, Khan M, Westbroek E, Ehresman J, et al. Clinical accuracy and initial experience with augmented reality–assisted pedicle screw placement: the first 205 screws. Journal of Neurosurgery: Spine. 2022;36(3):351-357. doi: 10.3171/2021.2.spine202097
7. Yahanda AT, Moore E, Ray WZ, Pennicooke B, Jennings JW, Molina CA. First in-human report of the clinical accuracy of thoracolumbar percutaneous pedicle screw placement using augmented reality guidance. Neurosurgical Focus. 2021;51(2):E10. doi: 10.3171/2021.5.focus21217
8. Bhatt FR, Orosz LD, Tewari A, Boyd D, Roy R, Good CR, et al. Augmented Reality-Assisted Spine Surgery: An Early Experience Demonstrating Safety and Accuracy with 218 Screws. Global Spine Journal. 2023;13(7):2047-2052. doi: 10.1177/21925682211069321
9. Butler AJ, Colman MW, Lynch J, Phillips FM. Augmented reality in minimally invasive spine surgery: early efficiency and complications of percutaneous pedicle screw instrumentation. The Spine Journal. 2023;23(1):27-33. doi: 10.1016/j.spinee.2022.09.008
10. Brown EJ, Fujimoto K, Blumenkopf B, Kim AS, Kontson KL, Benz HL. Usability Assessments for Augmented Reality Head-Mounted Displays in Open Surgery and Interventional Procedures: A Systematic Review. MTI. 2023;7(5):49. doi: 10.3390/mti7050049
11. Suter D, Hodel S, Liebmann F, Fürnstahl P, Farshad M. Factors affecting augmented reality head-mounted device performance in real OR. Eur Spine J. 2023;32(10):3425-3433. doi: 10.1007/s00586-023-07826-x
12. Khan T, Zhu TS, Downes T, Cheng L, Kass NM, Andrews EG, et al. Understanding Effects of Visual Feedback Delay in AR on Fine Motor Surgical Tasks. IEEE Trans Visual Comput Graphics. 2023;29(11):4697-4707. doi: 10.1109/tvcg.2023.3320214
13. D'Amato R, Cutolo F, Badiali G, Carbone M, Lu H, Hogenbirk H, et al. Key Ergonomics Requirements and Possible Mechanical Solutions for Augmented Reality Head-Mounted Displays in Surgery. MTI. 2022;6(2):15. doi: 10.3390/mti6020015
14. Ang S, Quarles J. Reduction of cybersickness in head mounted displays use: A systematic review and taxonomy of current strategies. Front Virtual Real. 2023;4:1027552. doi: 10.3389/frvir.2023.1027552
15. Kirollos R, Merchant W. Comparing cybersickness in virtual reality and mixed reality head-mounted displays. Front Virtual Real. 2023;4:1130864. doi: 10.3389/frvir.2023.1130864
16. Peng Y, Xie Z, Chen S, Wu Y, Dong J, Li J, et al. Application effect of head-mounted mixed reality device combined with 3D printing model in neurosurgery ventricular and hematoma puncture training. BMC Med Educ. 2023;23(1):670. doi: 10.1186/s12909-023-04659-6
17. Guha P, Lawson J, Minty I, Kinross J, Martin G. Can mixed reality technologies teach surgical skills better than traditional methods? A prospective randomised feasibility study. BMC Med Educ. 2023;23(1):144. doi: 10.1186/s12909-023-04122-6
18. Jain S, Timofeev I, Kirollos RW, Helmy A. Use of Mixed Reality in Neurosurgery Training: A Single Centre Experience. World Neurosurgery. 2023;176:e68-e76. doi: 10.1016/j.wneu.2023.04.107
19. Alemzadeh H, Raman J, Leveson N, Kalbarczyk Z, Iyer RK. Adverse Events in Robotic Surgery: A Retrospective Study of 14 Years of FDA Data. PLoS ONE. 2016;11(4):e0151470. doi: 10.1371/journal.pone.0151470
20. Hassanin SW, Kshirsagar RS, Eide JG, Chang J, Liang J, Palmer JN, et al. Image-Guided Surgical Device Failures in Functional Endoscopic Sinus Surgery: A MAUDE Analysis. The Laryngoscope. 2023;133(6):1310-1314. doi: 10.1002/lary.30296
21. Pagani NR, Menendez ME, Moverman MA, Puzzitiello RN, Gordon MR. Adverse Events Associated With Robotic-Assisted Joint Arthroplasty: An Analysis of the US Food and Drug Administration MAUDE Database. The Journal of Arthroplasty. 2022;37(8):1526-1533. doi: 10.1016/j.arth.2022.03.060
22. Shah HA, Mehta NH, Mehta SH, Ward M, McBriar JD, D'Amico RS. Characterizing complications associated with skull clamps: A review of the manufacturer and user facility device experience database. Clinical Neurology and Neurosurgery. 2023;235:108043. doi: 10.1016/j.clineuro.2023.108043
23. Davies C, Nieri CA, Sheyn A, Rangarajan S, Yawn RJ. The Use and Utility of Food and Drug Administration Adverse Event Data from the Manufacturer and User Facility Device Experience Database in Otology: A Systematic Review. Otol Neurotol. 2023;44(6):534-541. doi: 10.1097/mao.0000000000003899
24. Ziapour B, Zaepfel C, Iafrati MD, Suarez LB, Salehi P. A systematic review of the quality of cardiovascular surgery studies that extracted data from the MAUDE database. Journal of Vascular Surgery. 2021;74(5):1708-1720.e5. doi: 10.1016/j.jvs.2021.01.050

25. Sawaya J, Champlain A, Cohen J, Avram M. Barriers to Reporting: Limitations of the Maude Database. Dermatol Surg. 2021;47(3):424-425. doi: 10.1097/dss.0000000000002832
26. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Journal of Clinical Epidemiology. 2008;61(4):344-349. doi: 10.1016/j.jclinepi.2007.11.008
27. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLoS Med. 2015;12(10):e1001885. doi: 10.1371/journal.pmed.1001885
28. U.S. Food and Drug Administration. openFDA device adverse event API (Manufacturer and User Facility Device Experience, MAUDE) [Internet]. 2023. Available from: https://open.fda.gov/apis/device/event/
29. U.S. Food and Drug Administration. Code of Federal Regulations Title 21, Part 803: Medical device reporting [Internet]. 2023. Available from: https://www.ecfr.gov/current/title-21/chapter-I/subchapter-H/part-803
30. Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. Journal of the American Statistical Association. 1927;22(158):209-212. doi: 10.1080/01621459.1927.10502953
31. Yanni DS, Ozgur BM, Louis RG, Shekhtman Y, Iyer RR, Boddapati V, et al. 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. 2021;51(2):E11. doi: 10.3171/2021.5.focus21209
32. Lucas SM, Pattison EA, Sundaram CP. Global robotic experience and the type of surgical system impact the types of robotic malfunctions and their clinical consequences: an FDA MAUDE review. BJU International. 2012;109(8):1222-1227. doi: 10.1111/j.1464-410x.2011.10692.x
33. Felix B, Kalatar SB, Moatz B, Hofstetter C, Karsy M, Parr R, et al. Augmented Reality Spine Surgery Navigation. Spine. 2022;47(12):865-872. doi: 10.1097/brs.0000000000004338
34. Zari G, Condino S, Cutolo F, Ferrari V. Magic Leap 1 versus Microsoft HoloLens 2 for the Visualization of 3D Content Obtained from Radiological Images. Sensors. 2023;23(6):3040. doi: 10.3390/s23063040
35. FANG C, MO P, CHAN H, CHEUNG J, WONG JSH, WONG TM, et al. Can a Wireless Full-HD Head Mounted Display System Improve Knee Arthroscopy Performance? - A Randomized Study Using a Knee Simulator. Surg Innov. 2023;30(4):477-485. doi: 10.1177/15533506221142960
36. Sakai D, Schol J, Kawachi A, Sako K, Hiyama A, Katoh H, et al. Adolescent Idiopathic Scoliotic Deformity Correction Surgery Assisted by Smart Glasses Can Enhance Correction Outcomes and Accuracy and Also Improve Surgeon Fatigue. World Neurosurgery. 2023;178:e96-e103. doi: 10.1016/j.wneu.2023.06.144
37. Escalante Freitez MF. Device problems and patient injuries reported for clear aligners: a framework from the FDA MAUDE database. SAP Dentistry. 2023;1:60. doi: 10.62486/dent202360
38. Iacono V, Farinelli L, Natali S, Piovan G, Screpis D, Gigante A, et al. The use of augmented reality for limb and component alignment in total knee arthroplasty: systematic review of the literature and clinical pilot study. J exp orthop. 2021;8(1):52. doi: 10.1186/s40634-021-00374-7
39. Pose-Díez-de-la-Lastra A, Moreta-Martinez R, García-Sevilla M, García-Mato D, Calvo-Haro JA, Mediavilla-Santos L, et al. HoloLens 1 vs. HoloLens 2: Improvements in the New Model for Orthopedic Oncological Interventions. Sensors. 2022;22(13):4915. doi: 10.3390/s22134915
40. Healy GM, Ahrari A, Alkhalifah F, Oreopoulos G, Tan KT, Jaberi A, et al. Typology, Severity, and Outcomes of Adverse Events Related to Angiographic Equipment—A Ten-Year Analysis of the FDA MAUDE Database. Can Assoc Radiol J. 2023;74(4):737-744. doi: 10.1177/08465371231167990
41. Zhang JN, Fan Y, Hao DJ. Risk factors for robot-assisted spinal pedicle screw malposition. Sci Rep. 2019;9(1):3025. doi: 10.1038/s41598-019-40057-z
42. Matur AV, Palmisciano P, Duah HO, Chilakapati SS, Cheng JS, Adogwa O. Robotic and navigated pedicle screws are safer and more accurate than fluoroscopic freehand screws: a systematic review and meta-analysis. The Spine Journal. 2023;23(2):197-208. doi: 10.1016/j.spinee.2022.10.006
43. Luengo-Matos S, Sánchez-Gómez LM, Hijas-Gómez AI, García-Carpintero EE, Ballesteros-Massó R, Polo-deSantos M. Efficacy and safety of robotic spine surgery: systematic review and meta-analysis. J Orthop Traumatol. 2022;23(1):49. doi: 10.1186/s10195-022-00669-0
44. Guha D, Jakubovic R, Gupta S, Fehlings MG, Mainprize TG, Yee A, et al. Intraoperative Error Propagation in 3-Dimensional Spinal Navigation From Nonsegmental Registration: A Prospective Cadaveric and Clinical Study. Global Spine Journal. 2019;9(5):512-520. doi: 10.1177/2192568218804556
45. Zhao J, Jiang H, Zhuge Y, Gao R, Wang C, Ma J, et al. Risk Factors for the Drift Phenomenon in O-arm Navigation-Assisted Pedicle Screw Placement during Spinal Deformity Surgery. Orthopaedic Surgery. 2023;15(1):118-123. doi: 10.1111/os.13557
46. Uehara M, Takahashi J, Ikegami S, Kuraishi S, Shimizu M, Futatsugi T, et al. Are pedicle screw perforation rates influenced by distance from the reference frame in multilevel registration using a computed tomography-based navigation system in the setting of scoliosis?. The Spine Journal. 2017;17(4):499-504. doi: 10.1016/j.spinee.2016.10.019
47. Oba H, Ikegami S, Kuraishi S, Uehara M, Takizawa T, Munakata R, et al. Perforation Rate of Pedicle Screws Using Hybrid Operating Room Combined With Intraoperative Computed Tomography Navigation for Adolescent Idiopathic Scoliosis. Spine. 2020;45(20):E1357-E1364. doi: 10.1097/brs.0000000000003673
48. Toossi N, Vardiman AB, Benech CA, Kanaly CW, Maltenfort MG, Backes DM, et al. Factors Affecting the Accuracy of Pedicle Screw Placement in Robot-Assisted Surgery. Spine. 2022;47(23):1613-1619. doi: 10.1097/brs.0000000000004473

49. Lyell D, Wang Y, Coiera E, Magrabi F. More than algorithms: an analysis of safety events involving ML-enabled medical devices reported to the FDA. Journal of the American Medical Informatics Association. 2023;30(7):1227-1236. doi: 10.1093/jamia/ocad065
50. Correa J, Aribo C, Stuparich M, Nahas S, Cheung C, Behbehani S. Malfunction Events in the US FDA MAUDE Database: How Does Robotic Gynecologic Surgery Compare with Other Specialties?. Journal of Minimally Invasive Gynecology. 2022;29(2):300-307.e1. doi: 10.1016/j.jmig.2021.08.021
51. Estock JL, Li A, Casey MC, Taylor MA, Boudreaux-Kelly MY, Choma MD, et al. Assessing Use Errors Related to the Interface Design of Electrosurgical Units. AORN Journal. 2018;107(1):72-82. doi: 10.1002/aorn.12006
52. Krouwer JS. Manufacturers Need to Assess Rather Than Ignore User Complaints About Their Diabetes Devices. J Diabetes Sci Technol. 2023;17(2):592-592. doi: 10.1177/19322968211049804
53. Liebel TC, Daugherty T, Kirsch A, Omar SA, Feuerstein T. Analysis : Using the FDA MAUDE and Medical Device Recall Databases to Design Better Devices. Biomedical Instrumentation & Technology. 2020;54(3):178-188. doi: 10.2345/0899-8205-54.3.178
54. Chung G, Etter K, Yoo A. Medical device active surveillance of spontaneous reports: A literature review of signal detection methods. Pharmacoepidemiology and Drug. 2020;29(4):369-379. doi: 10.1002/pds.4980

FUNDING

No financing

CONFLICT OF INTEREST

None

AUTHORSHIP CONTRIBUTIONS

Drafting – original draft: Maria Carla Hernández Martínez.
Writing–review and editing: Maria Carla Hernández Martínez.