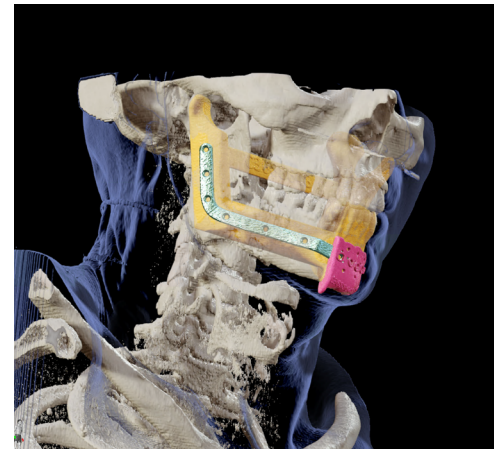


Editorial

Inteligencia artificial en cirugía oral y maxilofacial: la herramienta que hemos tomado prestada

Artificial intelligence in oral and maxillofacial surgery: the tool we have borrowed

Hace unos meses operamos a un paciente con un carcinoma escamoso de mandíbula. La indicación no tenía nada de excepcional, pero el caso ilustra bien por dónde estamos hoy en nuestra especialidad. Sobre el TC, con ayuda de un sistema de segmentación automática en el programa de planificación virtual de nuestro navegador, delimitamos el tumor y los márgenes adecuados de resección; realizamos la resección virtual del tumor sobre la mandíbula y determinamos y analizamos el defecto a reconstruir. Sobre ese modelo virtual planificamos la reconstrucción con un colgajo libre de peroné, y diseñamos unas guías de corte de la resección mandibular que contenía además la posición exacta de los tornillos de la placa de reconstrucción, una placa de osteosíntesis customizada, hecha para ese defecto y para ese paciente. Diseñamos las guías de corte del peroné para la reconstrucción específica del defecto del paciente. La residente que preparó la sesión usó un asistente conversacional para localizar el último consenso sobre márgenes y criterios radiológicos de invasión ósea. Cuando revisamos el caso en comité, el sistema nos sugirió tres diagnósticos diferenciales que ya habíamos descartado, pero que nos obligaron a justificar por escrito por qué los descartábamos. Con las gafas de realidad aumentada pudimos compartir y discutir en plan de resección y reconstrucción entre colegas y residentes. En el quirófano pudimos comprobar con la navegación quirúrgica los márgenes de resección tumoral planificados, la ubicación de las perforantes cutáneas del colgajo de peroné. La placa se adaptó con precisión. El colgajo reconstruyó con exactitud el defecto mandibular e intraoral.



Cuento esto sin entusiasmo tecnológico ni nostalgia, porque ninguna de esas dos posturas ayuda a operar mejor. Lo cuento porque ilustra lo que ha cambiado nuestra especialidad en los últimos años. La inteligencia artificial ya no aparece solo en el quirófano. Aparece también en nuestra sesión del comité de cáncer de cabeza y cuello de los miércoles, en la búsqueda bibliográfica de la residente, en las sesiones bibliográficas del servicio, en los resúmenes que se mandan a los congresos, en las presentaciones de trabajos en los congresos, en el primer borrador del manuscrito que se manda a una revista. Y en todos esos sitios reaparece la misma pregunta: ¿quién decide?

En el ámbito clínico la evidencia ya es considerable. La segmentación automática de la mandíbula sobre TC alcanza coeficientes Dice del 97 %, con errores de superficie por debajo del milímetro, y procesa cada estudio en menos de ocho segundos¹. La planificación virtual con implantes específicos del paciente para reconstrucciones con peroné², cirugía ortognática, prótesis de ATM o mallas de reconstrucción orbitaria, mejoran la precisión respecto a las placas premoldeadas y de stock. La detección automática de puntos cefalométricos opera con errores radiales medios menores de un milímetro^{3,4}. En oncología, los modelos de aprendizaje profundo predicen la supervivencia libre de enfermedad en el carcinoma epidermoide de cavidad oral con índices c-Harrell por

*Autor para correspondencia:

Correo electrónico: adeanferrer@yahoo.es (Alicia Dean Ferrer).

<http://dx.doi.org/10.20986/recom.2026.1772/2026>

encima de los modelos clásicos de Cox^{5,6}. Las redes convolucionales clasifican muestras histopatológicas con áreas bajo la curva superiores a 0,96, y los patólogos que las consultan antes de emitir su informe mejoran su propio rendimiento⁷. Los modelos de inteligencia artificial aplicados a la imagen también empiezan a aportar información útil en el diagnóstico diferencial de tumores parotídeos, aunque la evidencia se limita a series pequeñas⁸. Pero conviene aportar un matiz: en un estudio multicéntrico reciente con casos reales de traumatología maxilofacial, los residentes superaron de forma significativa a Chat GPT-4 y Google Bard en exactitud, claridad y completitud diagnóstica⁹. La utilidad real de la herramienta diagnóstica, hoy, está más en obligarnos a justificar el descarte de hipótesis que en sustituir el juicio clínico. No es poco, pero tampoco es lo que se promete en algunos foros.

La inteligencia artificial está cambiando la forma en que aprenden nuestros residentes, y lo está haciendo sin pedimos permiso. La consultan para preparar sesiones, para preparar resúmenes para congresos, para repasar antes de un examen, para resolver una duda en una guardia. Pelearse con eso me parece una batalla perdida. Chat GPT-4 ha demostrado superar el examen del USMLE, el examen de licencia médica estadounidense¹⁰ y rendir de manera competente en preguntas sobre cirugía oral y maxilofacial¹¹. Pero el mismo trabajo que valida ese rendimiento muestra, sin contradicción, que las referencias bibliográficas que el modelo proporciona contienen errores graves: en cirugía maxilofacial, hasta una de cada cuatro citas que aporta es inventada¹². El residente que copia sin verificar está construyendo conocimiento sobre arena movediza.

Eso nos da un trabajo nuevo como tutores y mentores. Enseñarles a usar la herramienta sin que la herramienta los use a ellos. Verificar la fuente, leer el original, distinguir un resumen útil de una alucinación elegante. Las principales sociedades han recomendado ya integrar formalmente la alfabetización en inteligencia artificial en los programas de residencia¹³, y a mí me parece sensato, importante y necesario. Se parece bastante a la radioprotección: una competencia que se evalúa antes de dejar que alguien use el aparato.

En la redacción científica los modelos de lenguaje se han instalado con la misma rapidez con la que en su día lo hizo el procesador de textos. Asisten en la búsqueda bibliográfica, sugieren estructuras, traducen, pulen el inglés y resumen artículos extensos. Bien usados, democratizan el acceso para quienes escribimos en una lengua que no es la nuestra. Mal usados, generan referencias que no existen, parafrasean sin atribuir y producen una prosa indistinguible que las revistas ya empiezan a detectar. La línea entre asistencia legítima y autoría delegada se está dibujando ahora mismo, y conviene posicionarse. Las grandes revistas internacionales lo han hecho de forma razonable: la inteligencia artificial puede ser herramienta, pero no autora; debe declararse su uso; y quien firma asume íntegramente la responsabilidad del texto.

En investigación, las posibilidades son mayores y los riesgos también. Los modelos de aprendizaje automático permiten integrar variables clínicas, anatómicas y moleculares para predecir complicaciones quirúrgicas, supervivencia, riesgo de osteorradionecrosis o evolución de lesiones potencialmente malignas^{5,7,14-16}. El aprendizaje federado permite que varios hospitales entrenen juntos un mismo modelo sin compartir físicamente los datos de sus pacientes, lo que resuelve buena parte del problema clásico entre disponer de muestras grandes y proteger la confidencialidad¹⁷. Y sin embargo, una revisión reciente de 99 estudios de inteligencia artificial en cirugía oral y maxilofacial publicados entre 2015 y 2025 concluye que predominan los trabajos exploratorios, retrospectivos y realizados en un solo centro. La validación externa multicéntrica es infrecuente, y no se ha identificado un solo ensayo clínico aleatorizado de gran tamaño en nuestra especialidad^{17,18}. La evidencia que tenemos es preliminar, que no es lo mismo que decir que sea débil. Y conviene tenerlo presente antes de adoptar como rutina algo que aún se está validando.

Mirado el conjunto, la frontera se formula con claridad. La inteligencia artificial propone, segmenta, predice, redacta y resume. La decisión, la firma y la responsabilidad siguen siendo nuestras. Cuando llegue una complicación quirúrgica, el comité de morbilidad no convocará al algoritmo. Cuando un manuscrito contenga una referencia inventada, la retracción del artículo no llevará la firma del modelo de lenguaje. Cuando un residente apruebe un examen sin entender lo que ha respondido, la responsabilidad de su formación seguirá siendo nuestra. La literatura jurídica reciente coincide en lo esencial: en sistemas de apoyo a la decisión, la responsabilidad profesional sigue siendo del clínico; en funciones autónomas entran en juego también las obligaciones de seguridad del fabricante; pero en ningún caso se diluye^{17,19}. La opacidad de los modelos, los sesgos de los datos de entrenamiento, la deriva del rendimiento una vez en uso clínico y el riesgo de alucinaciones son problemas reales que exigen auditoría continua. Conviene recordar a Watson for Oncology, de IBM, retirado del mercado tras fracasar por la falta de datos clínicos sólidos y por interrumpir el trabajo asistencial de los oncólogos¹⁹. Es un precedente que conviene tener presente.

¿Qué nos toca entonces a quienes nos tomamos en serio nuestro trabajo? Lo de siempre, en realidad. **Formación**, para que residentes y especialistas sepan leer una métrica de calibración, reconocer un sesgo de muestra, verificar una referencia y explicarle a un paciente, en un lenguaje que entienda qué papel ha tenido un sistema automático en su planificación. **Gobernanza interna**, para que cualquier herramienta que integremos en el flujo clínico, docente o investigador pase por validación local, auditoría periódica del rendimiento, declaración explícita en publicaciones y consentimiento informado proporcional al riesgo^{17,20,21}. Y **participación activa en el desarrollo**, porque las herramientas que se construyen sin nosotros acaban no sirviéndonos. La historia de Watson lo demostró bien.

La inteligencia artificial no nos va a sustituir, ni en el quirófano, ni en las sesiones clínicas, ni en los congresos ni en las publicaciones. Pero los servicios que la usen con cabeza probablemente acaben sustituyendo a quienes la rechacen, por orgullo o por falta de visión, y a quienes la abracen sin criterio, por puro deslumbramiento. Es una herramienta prestada, y como toda herramienta prestada nos pone a prueba. Lo demás es ir aprendiendo a distinguir entre la herramienta y la mano que la sostiene.

B I B L I O G R A F Í A

1. Xu J, Liu J, Zhang D, Zhou Z, Zhang C, Chen X. A 3D segmentation network of mandible from CT scan with combination of multiple convolutional modules and edge supervision in mandibular reconstruction. *Comput Biol Med.* 2021;138:104925. DOI: 10.1016/j.combiomed.2021.104925.
2. Ettinger KS, Mohamed AK, Nathan JM, Vierkant RA, Morris JM, Sears VA, et al. Patient-specific implants improve volumetric surgical accuracy compared to stock reconstruction plates in modern paradigm virtual surgical planning of fibular free flaps for head and neck reconstruction. *J Oral Maxillofac Surg.* 2024;82(10):1311-28. DOI: 10.1016/j.joms.2024.06.166.
3. Tang H, Liu S, Shi Y, Wei J, Peng J, Feng H. Automatic segmentation and landmark detection of 3D CBCT images using semi-supervised learning for assisting orthognathic surgery planning. *Sci Rep.* 2025;15:8814. DOI: 10.1038/s41598-025-93317-6.
4. Abdali HM, Ayoub AY, Omer RA, Alanazi RA, Alsaedi BS, Almutairi LB, et al. Applications of artificial intelligence in oral and maxillofacial cosmetic surgery: a systematic review of diagnostic, planning, and outcome assessment tools. *Cureus.* 2025;17(9):e92185. DOI: 10.7759/cureus.92185.
5. Kim DW, Lee S, Kwon S, Nam W, Cha IH, Kim HJ. Deep learning-based survival prediction of oral cancer patients. *Sci Rep.* 2019;9(1):6994. DOI: 10.1038/s41598-019-43372-7.
6. Abdul NS, Shivakumar GC, Sangappa SB, Di Basio M, Crimi S, Cicciú M, et al. Applications of artificial intelligence in the field of oral and maxillofacial pathology: a systematic review and meta-analysis. *BMC Oral Health.* 2024;24:122. DOI: 10.1186/s12903-023-03533-7.
7. Sukegawa S, Ono S, Tanaka F, Inoue Y, Hara T, Yoshii K, et al. Effectiveness of deep learning classifiers in histopathological diagnosis of oral squamous cell carcinoma by pathologists. *Sci Rep.* 2023;13(1):11676. DOI: 10.1038/s41598-023-38343-y.
8. Rao Y, Ma Y, Wang J, Xiao W, Wu J, Shi L, et al. Performance of radiomics in the differential diagnosis of parotid tumors: a systematic review. *Front Oncol.* 2024;14:1383323. DOI: 10.3389/fonc.2024.1383323.
9. Benedetti S, Frosolini A, Catarzi L, Vaira LA, Consorti G, Paglianiti M, et al. Large language models and surgical decision-making: evaluation of generative unimodal AI in facial traumatology practice. *J Maxillofac Oral Surg.* 2026;25(1):69-77. DOI: 10.1007/s12663-025-02556-7
10. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health.* 2023;2(2):e0000198. DOI: 10.1371/journal.pdig.0000198.
11. Quah B, Yong CW, Lai CWM, Islam I. Performance of large language models in oral and maxillofacial surgery examinations. *Int J Oral Maxillofac Surg.* 2024;53:881-6. DOI: 10.1016/j.ijom.2024.06.003.
12. Frosolini A, Franz L, Benedetti S, Benedetti S, Vaira LA, Filippis C, et al. Assessing the accuracy of ChatGPT references in head and neck and ENT disciplines. *Eur Arch Otorhinolaryngol.* 2023;280(11):5129-33. DOI: 10.1007/s00405-023-08205-4.
13. Miragall MF, Knoedler S, Kauke-Navarro M, Saadoun R, Grabenhorst A, Grill FD, et al. Face the Future-Artificial intelligence in oral and maxillofacial surgery. *J Clin Med.* 2023;12(21):6843. DOI: 10.3390/jcm12216843.
14. Nathan JM, Shim J, Arce K, Ettinger KS. Artificial intelligence for the oral and maxillofacial surgeon. *J Oral Maxillofac Surg.* 2026;84:608-20. DOI: 10.1016/j.joms.2025.12.013.
15. Mäkitie AA, Alabi RO, Ng SP, Takes RP, Robbins KT, Ronen O, et al. Artificial intelligence in head and neck cancer: a systematic review of systematic reviews. *Adv Ther.* 2023;40:3360-80. DOI: 10.1007/s12325-023-02527-9.
16. Dong F, Yan J, Zhang X, Zhang Y, Liu D, Pan X, et al. Artificial intelligence-based predictive model for guidance on treatment strategy selection in oral and maxillofacial surgery. *Heliyon.* 2024;10:e35742. DOI: 10.1016/j.heliyon.2024.e35742.
17. Consorti G, Catarzi L, Frosolini A, Vaira LA, Committeri U, Cirignaco G. Artificial intelligence in oral and maxillofacial surgery: a scoping review of clinical applications, ethical challenges, and legal considerations. *Int J Oral Maxillofac Surg.* 2026;S0901-5027(26)00079-2. DOI: 10.1016/j.ijom.2026.02.020.
18. Huang Y, Wang Y, Hou C, Song F, Jiang Y, Chen K, et al. Mapping artificial intelligence research in oral and maxillofacial surgery: a bibliometric analysis. *Int Dent J.* 2026;76:109456. DOI: 10.1016/j.identj.2026.109456.
19. Maliha G, Gerke S, Cohen IG, Parikh RB. Artificial intelligence and liability in medicine: balancing safety and innovation. *Milbank Q.* 2021;99:629-47. DOI: 10.1111/1468-0009.12504.
20. Collins GS, Moons KGM, Dhiman P, Diley RD, Beam AL, Calster BV, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ.* 2024;385:e078378. DOI: 10.1136/bmj-2023-078378.
21. Rana M, Sakkas A, Zimmermann M, Kostyuk M, Schwarz G. Artificial intelligence in oral and maxillofacial surgery: integrating clinical innovation and workflow optimization. *J Clin Med.* 2026;15:427. Doi: 10.3390/jcm15020427.