دانشنامه علمی چاوش طب آذین · chavoshteb.ir · به‌روزرسانی ۱ آذر ۱۴۰۵

گزارش‌دهی، سوگیری و عدالت در هوش مصنوعی تصویربرداری دندانی: چک‌لیست‌ها و راهنماها

کیفیت گزارش مطالعات هوش مصنوعی تصویربرداری دندانی از دقت مدل جداست: چارچوب‌های گزارش‌دهی موجود پایه جامعی فراهم می‌کنند، اما کمتر به کار می‌روند.[۱] در مرور سیستماتیک سوگیری و عدالت، هیچ‌یک از ۱۰ مطالعه تحلیل زیرگروه جمعیتی انجام نداده بود و ۸ مطالعه گزارش جمعیتی معناداری نداشت.[۲]

نکات کلیدی این مقاله

- مطالعات هوش مصنوعی دندانی به‌سرعت افزایش یافته‌اند، اما گزارش آن‌ها پراکنده و ناهمسان است.[۱]
- DECIDE-AI شامل ۱۷ آیتم اختصاصی هوش مصنوعی (۲۸ زیرآیتم) و ۱۰ آیتم عمومی برای ارزیابی بالینی اولیه است.[۳]
- راهنمای تخصصی رادیولوژی دهان و فک برای انتخاب و ترکیب منابع گزارش‌دهی در طراحی‌های پیچیده تدوین شده است.[۴]
- در ۱۰ مطالعه واردشده در مرور عدالت، همه روی تصاویر دوبعدی بودند و هیچ مطالعه CBCT واجد شرایطی یافت نشد.[۲]
- صفر از ۱۰ مطالعه تحلیل زیرگروه داشت؛ ۲ مطالعه گزارش جمعیتی ناقص و ۸ مطالعه هیچ گزارش معناداری نداشتند.[۲]

هوش مصنوعی بیش از پیش در تشخیص، تصویربرداری و برنامه‌ریزی درمان دندانپزشکی وارد شده است، اما پایبندی به چارچوب‌های استاندارد گزارش‌دهی ناهمسان مانده و تکرارپذیری، شفافیت و اعتبار نظارتی را محدود می‌کند.[۱] برای رادیولوژیست دهان و فک که مقاله‌ها را می‌خواند، داوری می‌کند یا ابزاری را برای بخش خود ارزیابی می‌کند، شناخت این چارچوب‌ها و پرسش‌های مربوط به سوگیری و عدالت بخشی از ارزیابی نقادانه است. این مقاله مکمل مقاله دانشنامه درباره خوانش نقادانه مطالعات دقت تشخیصی است.

فهرست مطالب

1. چرا گزارش‌دهی در هوش مصنوعی تصویربرداری دندانی اهمیت دارد

2. راهنماهای اصلی و انتخاب آن‌ها
3. سوگیری، انصاف و عدالت در مدل‌ها
4. نکات کاربردی برای رادیولوژیست
5. جدول جمع‌بندی نکات
6. پرسش‌های متداول
7. جمع‌بندی
8. مطالعه مرتبط در دانشنامه
9. منابع

چرا گزارش‌دهی در هوش مصنوعی تصویربرداری دندانی اهمیت دارد

Khurshid و همکاران راهنماهای گزارش‌دهی پزشکی برای هوش مصنوعی را در حوزه‌های کارآزمایی، پروتکل، مدل‌های پیش‌بینی، ارزیابی سوگیری، سامانه‌های پشتیبان تصمیم، تصویربرداری و مرور سیستماتیک، در کنار چک‌لیست‌های اختصاصی دندانپزشکی و چارچوب‌های اخلاقی، تحلیل کردند. نتیجه آنان این بود که گزارش پژوهش‌های هوش مصنوعی دندانی پراکنده و ناهمسان است.[۱]

> «هماهنگ‌سازی این چارچوب‌ها و ترویج پایبندی فعال از طریق سیاست مجلات، ادغام نظارتی و پایش کمّی پایبندی برای پرکردن فاصله میان عملکرد الگوریتم و پذیرش بالینی قابل اعتماد ضروری است.»
>
> “Harmonising these frameworks and promoting active adherence through journal policies, regulatory integration and quantitative compliance tracking are essential to bridge the gap between algorithmic performance and trustworthy clinical adoption.”
>
> — Khurshid و همکاران، Int Dent J 2026[۱]

راهنماهای اصلی و انتخاب آن‌ها

انتخاب بر اساس طراحی مطالعه

Barioni و همکاران در راهنمای تخصصی رادیولوژی دهان و فک یادآوری می‌کنند که انتخاب منبع مناسب، به‌ویژه در مطالعاتی که رادیومیکس، مدل‌سازی پیش‌بینی، ارزیابی دقت تشخیصی، سگمنتیشن، هوش مصنوعی چندوجهی و اعتبارسنجی بالینی را ترکیب می‌کنند، دشوار است و منابع مختلف نقش مکمل دارند.[۴] چک‌لیست CLAIM برای مطالعات

هوش مصنوعی در تصویربرداری پزشکی در سال ۲۰۲۴ به‌روز شد تا با تحول سریع این حوزه هماهنگ بماند.[۵]

نمونه راهنماهای گزارش‌دهی بر پایه نوع مطالعه

نوع مطالعه یا پرسش	راهنما یا ابزار پرکاربرد	کاربرد در رادیولوژی دهان و فک
مدل هوش مصنوعی روی تصاویر پزشکی	CLAIM (به‌روزرسانی ۲۰۲۴)	تشخیص، سگمنتیشن و طبقه‌بندی روی پانورامیک و CBCT
ارزیابی بالینی اولیه سامانه پشتیبان تصمیم	DECIDE-AI	نخستین استفاده بالینی در مقیاس کوچک، ایمنی و عوامل انسانی
کارآزمایی تصادفی و پروتکل آن	CONSORT-AI و SPIRIT-AI	مقایسه پیامد با و بدون کمک هوش مصنوعی
مدل پیش‌بینی	TRIPOD+AI و PROBAST برای خطر سوگیری	پیش‌بینی خطر یا پیامد از داده تصویری و بالینی
دقت تشخیصی و خطر سوگیری	STARD و QUADAS-2	مقایسه با استاندارد مرجع
مرور سیستماتیک	PRISMA	جمع‌بندی شواهد مدل‌ها

منبع: جمع‌بندی تحریریه بر پایه دسته‌بندی Khurshid و همکاران ۲۰۲۶ و Barioni و همکاران ۲۰۲۶؛ نام راهنماها برای راهنمایی خواننده آمده است.

DECIDE-AI برای ارزیابی بالینی اولیه

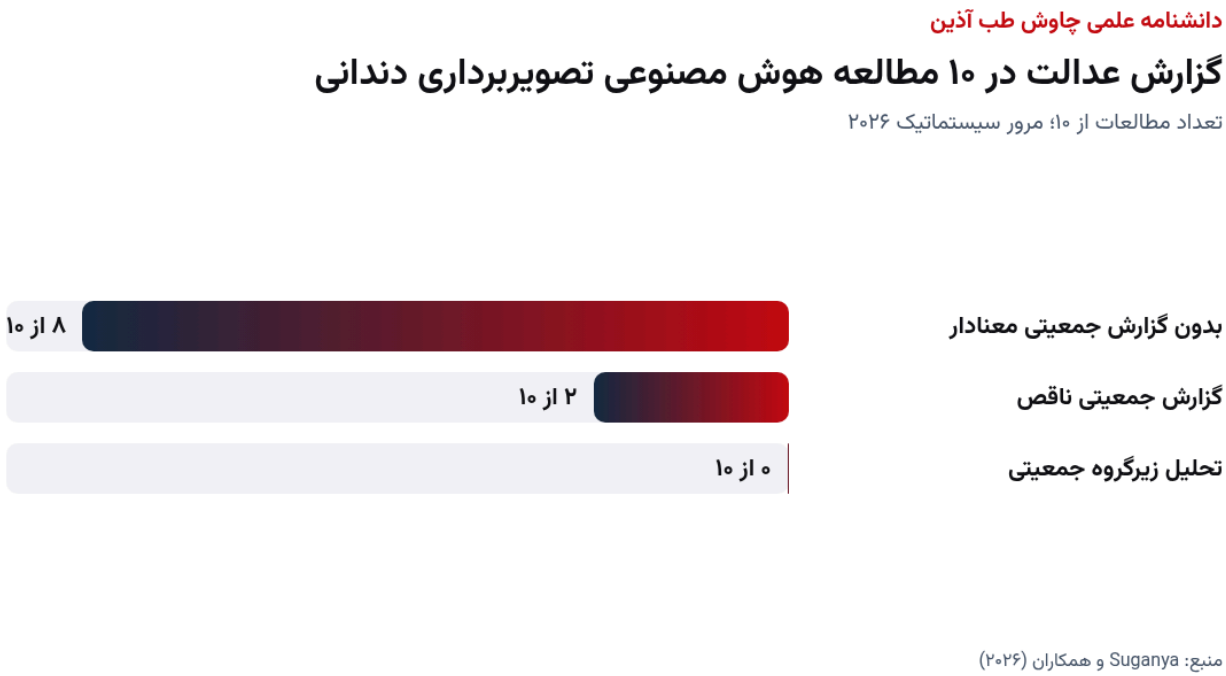
DECIDE-AI با فرایند دلفی اصلاح‌شده دومرحله‌ای و جلسه اجماع تدوین شد؛ ۱۲۳ متخصص در مرحله اول، ۱۳۸ نفر در مرحله دوم و ۱۶ نفر در جلسه اجماع شرکت کردند و متخصصان از ۲۰ گروه ذی‌نفع از پیش تعریف‌شده انتخاب شدند. نسخه نهایی ۱۷ آیتم اختصاصی هوش مصنوعی با ۲۸ زیرآیتم و ۱۰ آیتم عمومی دارد.[۳]

> «شمار فزاینده‌ای از سامانه‌های پشتیبان تصمیم بالینی مبتنی بر هوش مصنوعی در ارزیابی پیش‌بالینی و رایانه‌ای عملکرد امیدوارکننده‌ای نشان می‌دهند، اما تعداد کمی تاکنون فایده واقعی برای مراقبت از بیمار نشان داده‌اند.»

> “A growing number of artificial intelligence (AI)-based clinical decision support systems are showing promising performance in preclinical, in silico, evaluation, but few have yet demonstrated real benefit to patient care.”
>
> — Vasey و همکاران، BMJ 2022[۳]

سوگیری، انصاف و عدالت در مدل‌ها

Suganya و همکاران سوگیری را خطای نظام‌مند ناشی از داده‌های آموزش یا آزمون غیرنماینده یا نامتوازن، انصاف را عملکرد یکسان مدل در گروه‌های جمعیتی و عدالت را ارزش تشخیصی قابل مقایسه در گروه‌هایی با سن، جنس، قومیت، مرحله دندانی یا وضعیت اجتماعی‌اقتصادی متفاوت تعریف کردند.[۲] در ۱۰ مطالعه اصیل واردشده، همه مدل‌ها یادگیری عمیق روی تصاویر دوبعدی مانند پانورامیک، پری‌اپیکال، بایت‌وینگ و عکس داخل دهانی بودند و هیچ مطالعه CBCT واجد شرایطی پیدا نشد.[۲]



وضعیت گزارش عدالت در ۱۰ مطالعه هوش مصنوعی تصویربرداری دندانی؛ نمودار تحریریه بر پایه Suganya و همکاران.

وضعیت گزارش عدالت در مطالعات هوش مصنوعی تصویربرداری دندانی

معیار	نتیجه در ۱۰ مطالعه

معیار	نتیجه در ۱۰ مطالعه
تحلیل عملکرد در زیرگروه‌های جمعیتی	۰ از ۱۰
گزارش جمعیتی ناقص	۲ از ۱۰
بدون گزارش جمعیتی معنادار	۸ از ۱۰
مطالعات CBCT واجد شرایط	هیچ مطالعه‌ای
خطر سوگیری روش‌شناختی	در بیشتر مطالعات پایین

منبع: Suganya و همکاران، Int J Med Inform 2026.

> «سامانه‌های هوش مصنوعی مورد استفاده در تصویربرداری دندانی توانایی فنی قوی نشان می‌دهند اما ارزیابی کافی از سوگیری، انصاف و عدالت ندارند.»
>
> “AI systems used in dental imaging show strong technical capability but lack adequate evaluation of bias, fairness, and equity.”
>
> — Suganya و همکاران، Int J Med Inform 2026[۲]

این مرور پیشنهاد می‌کند اعتبارسنجی منصفانه شامل مجموعه آزمون مستقل از گروه‌های جمعیتی متنوع با گزارش عملکرد به تفکیک زیرگروه باشد و مجموعه داده نماینده ویژگی‌های جمعیت بالینی واقعی باشد.[۲] نکته مهم این است که خطر سوگیری روش‌شناختی پایین در ابزارهایی مانند QUADAS-2 به معنای ارزیابی عدالت نیست.

نکات کاربردی برای رادیولوژیست

- پیش از استفاده از یک ابزار، بپرسید داده آموزش و آزمون از چه جمعیت، چه دستگاه‌هایی و چند مرکز آمده است.[۲]
- در خواندن یا داوری مقاله، راهنمای متناسب با طراحی مطالعه را انتخاب کنید و در صورت ترکیب طراحی‌ها چند منبع را مکمل هم به کار ببرید.[۴]
- عملکرد پیش‌بالینی را با فایده بالینی یکی ندانید؛ ارزیابی بالینی اولیه و عوامل انسانی جداگانه گزارش شوند.[۳]
- برای CBCT، نبود مطالعه واجد شرایط در مرور عدالت یعنی شواهد انصاف عملاً وجود ندارد.[۲]

جدول جمع‌بندی نکات

نکات کلیدی گزارش‌دهی و عدالت در هوش مصنوعی تصویربرداری دندانی

محور	نکته کلیدی
وضعیت گزارش‌دهی	پراکنده و ناهمسان؛ چارچوب‌ها کمتر به کار می‌روند
DECIDE-AI	۱۷ آیتم اختصاصی (۲۸ زیرآیتم) + ۱۰ آیتم عمومی
CLAIM	به‌روزرسانی ۲۰۲۴ برای تصویربرداری پزشکی
تحلیل زیرگروه	۰ از ۱۰ مطالعه
CBCT	بدون مطالعه واجد شرایط در مرور عدالت
اعتبارسنجی منصفانه	مجموعه آزمون مستقل و متنوع با گزارش زیرگروه

منبع: جمع‌بندی تحریریه بر پایه منابع فهرست‌شده در همین مقاله.

پرسش‌های متداول

چرا راهنمای گزارش‌دهی برای مطالعات هوش مصنوعی لازم است؟

زیرا گزارش ناقص تکرارپذیری، شفافیت و اعتبار نظارتی را کاهش می‌دهد و گزارش در هوش مصنوعی دندانی هنوز پراکنده است.[۱]

DECIDE-AI چند آیتم دارد؟

۱۷ آیتم اختصاصی هوش مصنوعی شامل ۲۸ زیرآیتم و ۱۰ آیتم عمومی.[۳]

CLAIM چیست؟

چک‌لیست هوش مصنوعی در تصویربرداری پزشکی که در سال ۲۰۲۴ به‌روزرسانی شد.[۵]

کدام راهنما برای یک مطالعه ترکیبی رادیومیکس و پیش‌بینی مناسب است؟

راهنمای تخصصی رادیولوژی دهان و فک ترکیب مکمل چند منبع را بر اساس اجزای طراحی پیشنهاد می‌کند.[۴]

عدالت در مدل‌های هوش مصنوعی دندانی چگونه سنجیده می‌شود؟

با گزارش ویژگی‌های جمعیتی، نمایندگی مجموعه داده و تحلیل عملکرد در زیرگروه‌ها.[۲]

آیا شواهدی درباره عدالت مدل‌های CBCT وجود دارد؟

در مرور ۲۰۲۶ هیچ مطالعه CBCT واجد شرایطی یافت نشد.[۲]

جمع‌بندی

ارزش یک مدل هوش مصنوعی تصویربرداری دندانی فقط به عدد دقت آن نیست، بلکه به شفافیت گزارش، تناسب راهنمای استفاده‌شده با طراحی مطالعه و ارزیابی عملکرد در گروه‌های مختلف بیماران بستگی دارد. شواهد کنونی نشان می‌دهد چارچوب‌ها موجودند اما کمتر رعایت می‌شوند و تحلیل زیرگروه تقریباً گزارش نمی‌شود. رادیولوژیست دهان و فک با پرسیدن پرسش‌های درست درباره داده، اعتبارسنجی و عدالت، می‌تواند از پذیرش زودهنگام ابزارهای ناکافی جلوگیری کند.

مطالعه مرتبط در دانشنامه

- خوانش نقادانه مطالعات دقت تشخیصی
- هوش مصنوعی در رادیولوژی دندان و فک
- اعتبارسنجی بالینی هوش مصنوعی پوسیدگی
- برنامه‌ریزی ایمپلنت با هوش مصنوعی
- رادیومیکس در رادیولوژی دهان و فک
- مدل‌های زبانی بزرگ در رادیولوژی دهان و فک

منابع

شماره‌های بالانویس در متن به همین فهرست پیوند دارند؛ هر منبع با شناسه DOI یا نشانی رسمی ناشر قابل بررسی است.

1. Khurshid Z, Osathanon T, Shire MA, Schwendicke F, Samaranayake L. Artificial Intelligence in Dentistry: A Concise Review of Reporting Checklists and Guidelines. Int Dent J. 2026;76(1):109322. DOI PubMed PMC

2. Suganya P, Dupada P, Sruthi KG, Panda P, Mohanty JR. Bias, fairness, and equity in artificial intelligence systems used in dental imaging: A systematic review. Int J Med Inform. 2026;214:106433. DOI PubMed

3. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. BMJ. 2022;377:e070904. DOI PubMed PMC

4. Barioni ED, de Oliveira LAP, Santos MMA, Orhan K, de Castro Lopes SLP, Costa ALF. Artificial Intelligence Reporting Guidelines in Dentomaxillofacial Radiology: Specialty-Oriented Practical Guide. Diagnostics (Basel). 2026;16(18):2961. DOI PubMed PMC

5. Tejani AS, Klontzas ME, Gatti AA, Mongan JT, Moy L, Park SH, et al. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. Radiol Artif Intell. 2024;6(4):e240300. DOI PubMed PMC

درباره نویسنده: تیم تحریریه علمی چاوش طب آذین

این مقاله را تیم تحریریه علمی چاوش طب آذین بر پایه منابع داوری‌شده و بیانیه‌های رسمی فهرست‌شده تهیه کرده است؛ ادعای بازبینی بالینی مستقل ندارد و جایگزین قضاوت متخصص نیست. آخرین به‌روزرسانی: ۱ آذر ۱۴۰۵.