



Generative artificial intelligence in dentistry: a bibliometric analysis

Inteligencia artificial generativa en odontología: un análisis bibliométrico

Cristina Valeria Zambrano-Olvera¹
Rosaria Perez-Rizzo¹

¹ Universidad Católica de la Santísima Concepción, Facultad de Medicina. Concepción, Chile.

Submitted: 18-01-2026 | Revised: 27-02-2026 | Accepted: 24-04-2026 | Published: 26-04-2026

Corresponding Author: Rosaria Perez-Rizzo 

ABSTRACT

Introduction: Large language models and generative artificial intelligence have spread rapidly through dentistry, but the volume and structure of the corresponding literature have not been mapped. **Methods:** A bibliometric analysis of the Scopus-indexed literature was performed. **Results:** The query retrieved 1,777 documents published between 2022 and 2026 across 720 sources and 93 countries. Annual production grew from 1 document in 2022 to 900 in 2026 (an annual growth rate of 447.7%), and the corpus accumulated 15,608 citations (mean 8.8; h-index 54), with 42.8% of documents uncited. **Conclusions:** The field emerged in 2022, grew exponentially thereafter, and was led by Turkey, the United States, and China, with concentration in oral-health and dental-education journals.

Keywords: bibliometrics; artificial intelligence; large language models; dentistry; Scopus.

RESUMEN

Introducción: Los modelos de lenguaje de gran tamaño y la inteligencia artificial generativa se han extendido rápidamente por la odontología, pero el volumen y la estructura de la literatura correspondiente no se han cartografiado. **Métodos:** Se realizó un análisis bibliométrico de la literatura indexada en Scopus. **Resultados:** La consulta recuperó 1 777 documentos publicados entre 2022 y 2026 en 720 fuentes y 93 países. La producción anual creció de 1 documento en 2022 a 900 en 2026 (tasa de crecimiento anual del 447,7 %), y el corpus acumuló 15 608 citas (media 8,8; índice h 54), con un 42,8 % de documentos sin citar. **Conclusiones:** El campo surgió en 2022, creció de forma exponencial a partir de entonces y estuvo liderado por Turquía, Estados Unidos y China, con concentración en revistas de salud oral y de educación odontológica.

Palabras clave: bibliometría; inteligencia artificial; modelos de lenguaje de gran tamaño; odontología; Scopus.

INTRODUCTION

Large language models and generative artificial intelligence entered clinical and scientific discourse almost overnight after the public release of ChatGPT in November 2022, and their adoption in health care has since outpaced the evidence needed to regulate it.(1,2) Dentistry has been no exception to this trajectory. Reviews now document applications that range from answering patients' questions and supporting differential diagnosis to generating educational content, drafting manuscripts, and assisting examination preparation, and primary evaluations of model accuracy now number in the dozens each year.(3-6)

The earliest evaluations already reported that ChatGPT and its successors answered dental knowledge questions with variable but improving accuracy, with scores on structured dental examinations and licensure-style questions rising as successive model versions were released.(7-9)

These examination-oriented studies have been especially attractive because dental licensure and specialty examinations provide a ready-made, standardized benchmark against which a general-purpose model can be scored without requiring any dental data or annotation, a methodological convenience that has shaped the early literature and its emphasis on knowledge recall over clinical reasoning, and one that is visible in the prominence of education journals among the top sources of the corpus. Later work extended these comparisons to multimodal models capable of reasoning over radiographs and clinical images, reporting diagnostic performance for oral and maxillofacial differential diagnosis, oral cancer staging, and implant-related questions that approached or, in selected tasks, matched clinician benchmarks.(10-13)

The successive model generations have themselves become moving targets for this literature, because each release, from GPT-3.5 through GPT-4 and its multimodal variants to the newer proprietary and open-weight systems, has been followed by a fresh wave of dental evaluations that re-test the same clinical and educational questions against the latest system.(14,15) Alongside these accuracy studies, a parallel strand has examined the risks and governance of generative tools in dental practice, including hallucination, patient safety, and the detection of machine-generated text in dental manuscripts.(16,17)

The volume of this output is large enough to warrant its own mapping, because the literature has grown faster than the systematic reviews that could synthesize it, and the first comprehensive reviews already flag heterogeneity in methods, tasks, and reporting as barriers to comparison.(18,19)

The gap is bibliometric rather than clinical. Earlier bibliometric studies have mapped artificial intelligence in dentistry as a whole, locating its leading sources, countries, and themes and describing a field that had already expanded through convolutional and diagnostic applications before the generative wave.(20-22)

A handful of studies have begun to map large language models specifically, the broader generative-artificial-intelligence literature, or the use of these tools within particular dental subspecialties and educational contexts.(23-26)

These prior analyses, however, were published before the most recent growth surge, were confined to a single subspecialty or to education, or did not combine the large-language-model and generative-artificial-intelligence terminologies in a single Scopus-defined corpus with a fixed date of extraction. A further reason that an updated mapping is needed is that the earliest bibliometric treatments of this topic were necessarily based on small, early corpora, and their rankings of sources and countries may no longer describe a field that has since expanded by an order of magnitude.

Bibliometric indicators are point-in-time descriptions, and their value decays as a field grows, so a mapping that was accurate in 2024 can be misleading in 2026, which is the situation that motivates the present update. The objective of this study was to characterize the Scopus-indexed literature on large language models and generative artificial intelligence in dentistry between 2018 and 2026 by quantifying its annual production and growth, identifying its leading sources and countries, and describing its citation impact.

METHODS

Design

A bibliometric analysis was performed on a Scopus-only corpus, following the recommendation that bibliometric studies restrict their corpus to a single database to avoid double-counting across databases and to permit reproducible retrieval.(23)The choice of Scopus rather than Web of Science or PubMed was driven by the target journal's scope and by the availability of a structured

export, and the single-database restriction means that the results describe the Scopus-indexed literature specifically rather than the totality of published work on the topic. The study was observational, retrospective, and descriptive, and no human participants or patient records were involved, so no ethical approval was required. The unit of analysis was the individual indexed document, the indicators were computed at the corpus level rather than at the level of individual authors or institutions, and no author-level or institution-level ranking was undertaken, which avoids the interpretive problems associated with attributing documents to the many co-authors of large collaborative evaluations and with whole-counting authorship that systematically favors senior investigators on multi-author papers.

Search

Scopus search expression: TITLE-ABS-KEY(("large language model" OR ChatGPT OR "generative artificial intelligence" OR "generative AI") AND (dental OR dentistry OR oral)) AND PUBYEAR > 2017 AND PUBYEAR < 2027.

The large-language-model and generative-artificial-intelligence terms were joined by OR within one parenthetical clause, the dental, dentistry, and oral terms were joined by OR within a second clause, and the two clauses were combined with AND and applied to the title, abstract, and keyword fields, which is the standard TITLE-ABS-KEY scope for bibliometric retrieval.

The publication-year clause restricted the window to 2018 through 2026, matching the study window. The wildcard in "large language model" captured both the singular and plural forms, and the inclusion of both ChatGPT and the generic generative-artificial-intelligence terms ensured that the corpus was not limited to a single product, because relying on the proprietary name alone would have excluded studies of other models and of the field as a whole.

The choice of "generative artificial intelligence" and its abbreviation "generative AI" as synonyms reflects the terminology that has become standard in the field, and the dental, dentistry, and oral terms were chosen to span the clinical specialty, the discipline, and the broader oral-health domain in which dental applications are published.

The search returned 1,777 unique documents after one duplicate identified by electronic identifier was removed. The deduplication step compared electronic identifiers across records and retained the first occurrence of any identifier that appeared more than once, which is necessary because Scopus records can be duplicated when a document is indexed under multiple affiliations or when a correction is issued.

All document types returned by the query were retained, including articles, reviews, conference papers, and editorials, because the objective was to map the whole literature rather than a filtered subset, and no language restriction was applied, so that non-English dental literature would not be systematically excluded. The broad scope of the query was deliberate, because a more restrictive search limited to "dentistry" alone would have missed the substantial share of relevant work published under "oral health", "oral medicine", and "dental education" headings.

Analysis

The export was read into Python 3.11 and analyzed with pandas. Annual production was tabulated by year of publication, and the compound annual growth rate was computed from the first and last years with nonzero production using the geometric-mean formula over the intervening years, so that the value reflects the average annual multiplicative growth rather than a simple arithmetic average of year-over-year changes.

The year-over-year percentage changes were also computed for each pair of consecutive years and reported where informative, because a single compound rate can obscure a decelerating growth curve, and the two measures together describe both the average and the trajectory of growth.

The leading sources and countries were ranked by document count, with countries extracted from the affiliation field of each record by taking the country token of each listed affiliation and aggregating across the corpus, so that a document with authors from multiple countries contributed to each of those countries. This whole-counting approach, in which every listed affiliation country receives one credit for the document, is the standard convention in bibliometric country analysis and is preferable to fractional counting for a corpus of this size because it is transparent and reproducible, even though it slightly favors countries that participate in large international collaborations.

Citation impact was summarized by the total number of citations, the mean and median citations per document, the proportion of documents with zero citations, and the h-index of the corpus, defined as the largest number h such that at least h documents each received at least h

citations. The median was reported alongside the mean because citation distributions are heavily right-skewed, and the mean alone would overstate the citation count of a typical document.

RESULTS

The study comprised 1,777 documents published between 2022 and 2026, with no documents indexed in 2018, 2019, 2020, or 2021 (Table 1). Annual production was 1 document in 2022, 60 in 2023, 237 in 2024, 579 in 2025, and 900 in 2026, a compound annual growth rate of 447.7% over the four years of observed production (Figure 1). The year-over-year increases were 295.0% from 2023 to 2024, 144.3% from 2024 to 2025, and 55.4% from 2025 to 2026, so that the absolute annual increment grew every year even as the percentage growth rate decelerated from its initial explosive value.

Figure 1. Annual production of documents on generative artificial intelligence in dentistry.

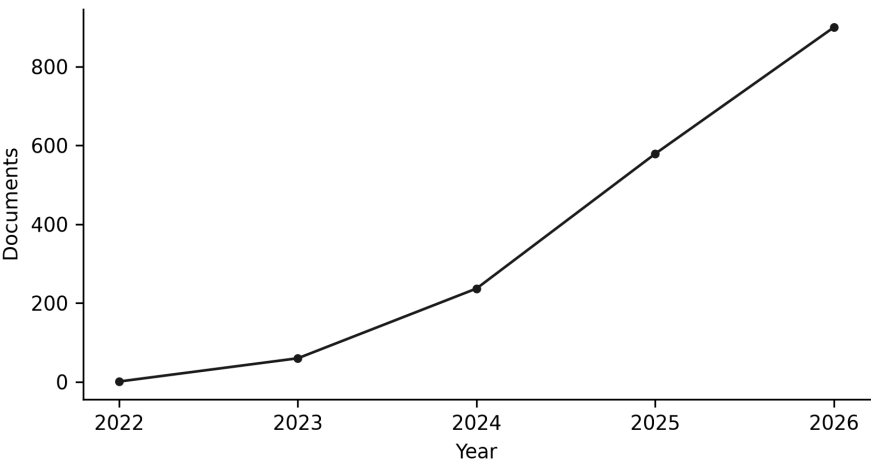


Table 1. Annual production of documents on generative artificial intelligence in dentistry.

Year	Documents
2022	1
2023	60
2024	237
2025	579
2026	900

The production curve rose from a near-zero baseline to nearly one thousand documents in four years, a trajectory amounting to a doubling time of less than one year between 2023 and 2026, and the 2026 output alone exceeded the combined output of the three preceding years. The shape of the curve is that of an early exponential phase with no evidence of the plateau that typically precedes consolidation, which indicates that the field’s literature was still expanding rapidly at the time of the search and that the document count reported here is a lower bound on the eventual size of the 2026 cohort, since documents indexed later in the year would not yet appear in the retrieved export.

The absence of documents before 2022 indicates that the field began with the public release of ChatGPT rather than with earlier transformer-based language models, and that the preceding years of deep-learning research in dental imaging constitute a separate, earlier literature that the present query did not capture.

The single document indexed in 2022 predates the November release of ChatGPT and therefore reflects an early generative-artificial-intelligence application that preceded the public launch, whereas the 60 documents of 2023 represent the first full year of the post-ChatGPT literature and a sixtyfold increase over the baseline. The pattern of one document in the first year followed by rapid multiplication is typical of tool-driven research fronts, in which the enabling technology rather than a scientific question supplies the initial impetus.

The literature was dispersed across 720 distinct sources, and the ten most productive sources were led by BMC Oral Health with 106 documents, followed by the International Dental Journal with 54 and the Lecture Notes in Computer Science series with 45 (Table 2).

Table 2. Top 10 sources by number of documents.

Source	Documents
BMC Oral Health	106
International Dental Journal	54
Lecture Notes in Computer Science Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics	45
European Journal of Dental Education	38
Journal of Dentistry	31
Scientific Reports	27
BMC Medical Education	26
Journal of Prosthetic Dentistry	25
Oral Diseases	25
Journal of Dental Education	23

Note: BMC = BioMed Central.

The three most productive sources together accounted for 205 documents, only 11.5% of the corpus, which confirms that no small group of journals dominated the field, and the distribution across 720 sources means that the average source contributed fewer than three documents.

This fragmentation is characteristic of a field in which publication opportunities are distributed across many specialist and general journals, and it contrasts with more mature bibliometric corpora in which a handful of journals account for a large share of output. The prominence of open-access megajournals and of the BMC series among the top sources reflects the rapid publication model that has accommodated the field's fast growth, and it is consistent with the pattern seen in other fast-moving artificial-intelligence subfields.

Dental education journals accounted for three of the top ten sources, the European Journal of Dental Education with 38 documents, BMC Medical Education with 26, and the Journal of Dental Education with 23, a cluster that reflects the prominence of educational applications and examination-oriented evaluations in the literature.

The remaining leading sources were the Journal of Dentistry with 31, Scientific Reports with 27, the Journal of Prosthetic Dentistry with 25, and Oral Diseases with 25. The presence of a computer-science proceedings series among the top sources indicates that a nontrivial share of the output is published outside the dental literature proper, in venues where machine-learning methods are the primary audience and where dental applications serve as case studies rather than as clinical contributions.

This split between clinical-dental venues and engineering venues is a structural feature of artificial-intelligence research in health care generally, and it has implications for how the field should be reviewed, because a synthesis confined to dental journals would miss the methodological contributions published in the computer-science literature.

The clustering of educational studies in dedicated dental-education journals further suggests that the field's subspecialties are organizing themselves into distinct publication channels rather than converging on a shared venue.

The 93 contributing countries were led by Turkey with 262 documents, the United States with 204, and China with 186, followed by India with 115 and Saudi Arabia with 66 (Table 3). Brazil, Spain, Japan, Germany, and the United Kingdom completed the top ten with 62, 55, 50, 43, and 41 documents, respectively, and the top three countries together accounted for 652 documents, more than one third of the corpus.

The geographical distribution was thus broad at the top but concentrated in the aggregate, with 93 countries contributing at least one document and the top ten countries contributing the large majority, a profile in which a long tail of countries contributed a small number of documents each.

The presence of four middle-income countries among the top ten, Turkey, India, Saudi Arabia, and Brazil, is a distinctive feature of this corpus relative to the imaging-based dental artificial-intelligence literature, and it suggests that the low entry barrier of evaluating a commercial model has broadened participation beyond the countries that historically dominate dental research.

Citation impact was concentrated rather than uniform: the corpus accumulated 15,608 citations, a mean of 8.78 and a median of 1 citations per document, and an h-index of 54, while 42.80% of documents remained uncited. The h-index of 54 means that 54 documents had each been cited at least 54 times, which for a four-year-old field represents a substantial core of highly cited work, and it is driven disproportionately by the early comprehensive reviews and the first high-profile evaluations that established the field's reference points and continue to attract citations as later studies position themselves against those foundations.

The large gap between the mean and the median confirms that a small minority of highly cited documents carried the citation load, whereas nearly half of the corpus had not yet been cited at all, a pattern consistent with a young literature in which a few early and comprehensive contributions attracted the bulk of attention while the rapidly expanding tail of descriptive evaluations accumulated little or no citation.

The distribution was strongly right-skewed: 340 documents had received ten or more citations, 60 had received fifty or more, and 26 had reached one hundred, while the most-cited single document had accumulated 463 citations. This concentration means that the citation metrics of the corpus are driven by a few dozen foundational contributions, and the median of one citation per document indicates that the typical document in this literature has been cited at most once.

Table 3. Top 10 countries by number of documents.

Country	Documents
Turkey	262
United States	204
China	186
India	115
Saudi Arabia	66
Brazil	62
Spain	55
Japan	50
Germany	43
United Kingdom	41

DISCUSSION

The defining finding is the timing and steepness of the growth. No documents were indexed before 2022, and the corpus then expanded from a single document in 2022 to 900 in 2026, a compound annual growth rate of 447.7% that amounts to a doubling time of less than one year.

The field of large language models and generative artificial intelligence in dentistry is therefore almost entirely a post-ChatGPT phenomenon, and its literature is still in an exponential rather than a consolidating phase, with a 55.4% year-over-year increase in the most recent year showing no sign of saturation at the time of retrieval.

The deceleration in the percentage growth rate, from 295.0% to 144.3% to 55.4% across successive years, is the arithmetic signature of a maturing curve, but the absolute annual increment continued to rise, which indicates that the field is still expanding even as its explosive initial phase passes.

The growth is attributable not to a single enabling advance but to a sequence of releases, each of which lowered the barrier to producing an evaluation and thereby invited another cohort of authors into the field, a self-reinforcing dynamic that is characteristic of tool-driven rather than question-driven research.

The comparison with the earlier deep-learning literature in dental imaging is instructive here, because that literature grew out of a methodological and dataset infrastructure that took years to build, whereas the generative literature required only access to a public model and a set of questions, which explains why its growth rate has been an order of magnitude faster. This infrastructure-free entry is both the field's strength and its weakness, because it has democratized participation while lowering the methodological bar, a tension that runs through the limitations and implications discussed below.

The absence of any indexed output before 2022 is itself informative, because it shows that the transformer-based language models of the pre-ChatGPT era, which were available for several years, did not trigger a dental literature until a consumer-facing product made them accessible to investigators without specialized machine-learning expertise.

The growth pattern is consistent with what has been reported for large language models in medicine generally, where the literature also accelerated after 2022 and remains dominated by exploratory evaluations rather than implementation studies.(1,27,28)

The dominance of evaluation over implementation is visible in the present corpus in the repeated testing of model accuracy against licensure-style questions, a design that is inexpensive to execute and easy to publish but that does not by itself establish whether the models change clinical or educational outcomes, a gap that the broader medical large-language-model literature has also identified.(2,28)

In turn, the concentration of output in oral-health journals such as BMC Oral Health and the International Dental Journal matches the distribution reported in earlier bibliometric studies of artificial intelligence in dentistry, which identified the same general-dental and oral-health venues as the field's core.(20,22)

The prominence of dental-education journals is likewise concordant with the finding that educational applications and licensure-examination evaluations constitute a large share of dental large language model research.(9,26,29)

The accuracy-improvement trend documented in temporal meta-analyses, which find rising performance as newer models are released, provides a plausible driver for the sustained growth, because each new model release generates a fresh wave of comparative evaluations that re-test the same clinical questions against the latest system.(15)

The spread of evaluations across education, diagnosis, and patient communication likewise matches the tripartite division that the scoping reviews of dental ChatGPT applications have described, with education and examination performance receiving the earliest and densest coverage.(5,6,30)

The correspondence between the present corpus and the content of the existing qualitative reviews indicates that the query captured the field's mainstream rather than a biased slice of it, notwithstanding the contamination discussed below. The consistency of the country and source distributions with the earlier bibliometric studies of artificial intelligence in dentistry further validates the retrieval, because a query that was substantially off-target would be expected to produce a source and country profile unrelated to the dental literature, whereas the observed profile is recognizably dental and education-focused.(20,22,26)

The country ranking departs from the pattern observed in earlier artificial-intelligence dentistry bibliometrics. Turkey led with 262 documents, ahead of the United States and China, whereas general artificial-intelligence bibliometric studies have usually placed China and the United States first.(20,24)

The discrepancy is plausibly explained by the inclusion of the broad term "oral" in the search, which may have captured oral-health and general-oral publications beyond dentistry strictly defined, and by the comparatively low barrier to publishing descriptive evaluations of commercially available models, which has attracted a large volume of contributions from middle-income settings where access to the models is easy but access to structured evaluation frameworks is less established.

An additional contributing factor is the presence of a substantial Turkish-language dental literature and a large number of Turkish dental journals indexed in Scopus, which, combined with the ease of evaluating freely accessible models, has produced a high volume of descriptive studies from that country.

Whatever the precise cause, the ranking demonstrates that bibliometric country rankings for this topic are sensitive to query construction and should be interpreted with the same caution as any indicator derived from a broad search. The dispersion across 720 sources for 1,777 documents, a ratio of fewer than three documents per source, is a further sign that the corpus is still fragmented and has not yet concentrated in a small set of specialist journals, and the estimated presence of ChatGPT-generated text in dental manuscripts may itself contribute to the volume by lowering the cost of producing a publication.(16)

The strong representation of middle-income and emerging-economy countries in the top ten, including Turkey, India, Saudi Arabia, and Brazil, is consistent with a field whose entry costs are low because the models are freely accessible, and it contrasts with the capital-intensive imaging-based artificial-intelligence literature, whose leading countries reflect access to large annotated datasets and computational infrastructure.(20,21)

The study has limitations that follow from its design, and each suggests a remedy. The query relied on the term “oral”, which is broad and introduced records on oral administration, oral cancer, and other non-dental topics, inflating the corpus and the country counts; the presence among the most-cited documents of a pharmacology paper on a small-molecule inhibitor, retained only because of the term “oral”, confirms this contamination concretely, and a future analysis could restrict the search to dental-specific terms or apply a manual title screen to quantify and remove off-topic records before reporting a corrected corpus.

The export did not include author keywords or cited references, so keyword co-occurrence, thematic clustering, and co-citation structure could not be computed, and these are the dimensions that would reveal the intellectual organization of the field beyond its volume; a dedicated Scopus export with the complete view, which includes those fields, would enable the network analyses that are standard in bibliometric work.

This omission is a consequence of the access tier of the application programming interface used for retrieval, and it should be read as a limitation of the present study's scope rather than as a property of the underlying corpus, which does carry those fields in the full database record.

The citation metrics are biased downward for the most recent documents, which have had less time to accumulate citations, so the h-index and mean citations understate the eventual impact of the 2025-2026 cohort and cannot be compared directly with citation benchmarks from older fields. The study is confined to Scopus, and documents indexed only in other databases were not captured, which may under-represent some regional dental journals and non-English-language literature despite the absence of a language restriction in the query, and a multi-database comparison would be needed to estimate the size of this under-representation. The single-database restriction is a deliberate trade-off that favors reproducibility over completeness, and it means that the absolute document counts should be read as Scopus-indexed counts rather than as estimates of all published work on the topic.

The practical implication is that the field has now reached a scale and growth rate that justify structured synthesis rather than further uncoordinated evaluation. With 900 documents in 2026 alone and a citation distribution in which 42.8% of documents are uncited, the marginal value of additional single-model evaluations is low, and the field would benefit from living systematic reviews, standardized reporting of model and prompt details, and replication of the few high-impact findings that have accumulated most citations.^(14,17) Editors and reviewers have a role in this shift, because the proliferation of redundant evaluations is sustained in part by a publication system that rewards novelty of model over novelty of question, and a shift toward requiring preregistration, standardized reporting checklists, and explicit comparison with the existing literature would raise the threshold that each new evaluation must clear.

The breadth of applications already documented, from caries and cancer staging to implant planning and patient education,^(10,12,13) means that evidence syntheses should be organized by clinical task rather than by model, so that a clinician can locate the state of the evidence for a specific decision rather than for a specific product. Such task-based syntheses would also make the heterogeneity that the existing reviews flag more tractable, because the relevant comparators, outcomes, and reporting standards differ between a diagnostic task and an educational task, and collapsing them into a single review obscures rather than clarifies the evidence.

The concentration of output in education journals further suggests that dental curricula and assessment are the immediate translation frontier, and that evidence syntheses on educational applications should precede those on clinical decision support, whose evidential base is still thinner. The practical consequence of the citation concentration is equally clear: with a median of one citation per document and 42.8% of the corpus uncited, the field is producing more evaluation than it is reading, and the returns to additional single-model evaluations are diminishing even as the absolute number of such studies continues to rise. The data therefore support a rebalancing of effort, away from the production of further descriptive accuracy studies and toward the synthesis, standardization, and implementation research that a field of this size now requires.

CONCLUSIONS

The literature on large language models and generative artificial intelligence in dentistry first emerged in 2022 and grew exponentially to reach a total of 1,777 documents by the year 2026, at an annual growth rate of 447.7% over the period.

The field was dispersed across 720 sources and 93 countries, led by Turkey, the United States, and China, with citation impact concentrated in a minority of highly cited documents. The corpus

has reached a scale that calls for synthesis and standardization rather than additional descriptive evaluations.

REFERENCES

1. Iqbal U, Tanweer A, Rahmanti AR, Greenfield D, Lee LTJ, Li YCJ. Impact of large language model (ChatGPT) in healthcare: an umbrella review and evidence synthesis. *J Biomed Sci*. 2025;32(1):45. doi: 10.1186/s12929-025-01131-z
2. Zhang Z, Momeni Nezhad MJ, Bagher Hosseini SM, Zolnour A, Zonour Z, Hosseini SM, et al. Advancing healthcare with large language models: A scoping review of applications and future directions. *International Journal of Medical Informatics*. 2026;208:106231. doi: 10.1016/j.ijmedinf.2025.106231
3. Akitomo T, Hamada M, Hamaguchi S, Kusaka S, Nomura R. Generative Artificial Intelligence and Large Language Models in Paediatric Dentistry: A Scoping Review. *International Dental Journal*. 2026;76(2):109401. doi: 10.1016/j.identj.2025.109401
4. Alhaidry HM, Fatani B, Alrayes JO, Almana AM, Alfhaed NK. ChatGPT in Dentistry: A Comprehensive Review. *Cureus*. 2023. doi: 10.7759/cureus.38317
5. Hamada M, Kikuchi S, Akitomo T, Kusaka S, Iwamoto Y, Nomura R. Applications and potential of ChatGPT in dentistry: Scoping review of research perspectives. *Journal of Dental Sciences*. 2026;21(1):1-8. doi: 10.1016/j.jds.2025.08.016
6. Puleio F, Lo Giudice G, Bellocchio AM, Boschetti CE, Lo Giudice R. Clinical, Research, and Educational Applications of ChatGPT in Dentistry: A Narrative Review. *Applied Sciences*. 2024;14(23):10802. doi: 10.3390/app142310802
7. Arpaci A, Ozturk AU, Okur I, Sadry S. Evaluation of the accuracy of ChatGPT-4 and Gemini's responses to the World Dental Federation's frequently asked questions on oral health. *BMC Oral Health*. 2025;25(1):1293. doi: 10.1186/s12903-025-06624-9
8. Kandemir M, Sarıbaş EE. Comparative performance of large language models in answering periodontology questions from the Turkish Dental Specialty Examination: a cross-sectional study on accuracy and coverage. *BMC Oral Health*. 2025;25(1):1804. doi: 10.1186/s12903-025-07241-2
9. Kim W, Kim BC, Yeom HG. Performance of Large Language Models on the Korean Dental Licensing Examination: A Comparative Study. *International Dental Journal*. 2025;75(1):176-184. doi: 10.1016/j.identj.2024.09.002
10. de Araujo BMDM, Cavassin ME, Abuabara A, Giacon ML, Fariniuk LF, Kirschneck C, et al. Diagnostic performance of multimodal large language models in detecting pulp stones on panoramic radiographs. *BMC Oral Health*. 2026;26(1):2043. doi: 10.1186/s12903-026-10119-6
11. Shujaat S, Riaz M, Alsubaie JM, Almuammar AA, Pillai A, Alfadley A, et al. Differential Diagnosis Performance of ChatGPT in Oral and Maxillofacial Pathology. *Oral Diseases*. 2026;odi.70501. doi: 10.1111/odi.70501
12. Ai QYH, Kwok HM, Lu MY, Lau TTS, Hung KF, Wong LM, et al. Commercial large language models for oral cavity cancer staging using descriptive pre-treatment MRI reports: ready for standalone use in clinical practice?. *Insights Imaging*. 2026;17(1):233. doi: 10.1186/s13244-026-02409-y
13. Erdal B, Duman FM, Altınkılıç MR, Horoz M. Artificial Intelligence as a Digital Consultant: Evaluating Large Language Models for Patient Inquiries on Dental Caries. *J Esthet Restor Dent*. 2026;jerd.70284. doi: 10.1111/jerd.70284
14. Wu X, Cai G, Guo B, Ma L, Shao S, Yu J, et al. A multi-dimensional performance evaluation of large language models in dental implantology: comparison of ChatGPT, DeepSeek, Grok, Gemini and Qwen across diverse clinical scenarios. *BMC Oral Health*. 2025;25(1):1272. doi: 10.1186/s12903-025-06619-6
15. Doubleday AF, Cheverko CM, Bolgova O, Mavrych V, Mohamed FRR, Westrick J, et al. Temporal trends in large language model (LLM) accuracy: A meta-analysis of multiple-choice question performance in dentistry and dental education. *Journal of Dentistry*. 2026;171:106724. doi: 10.1016/j.jdent.2026.106724
16. Uribe SE, Maldupa I. Estimating the use of ChatGPT in dental research publications. *Journal of Dentistry*. 2024;149:105275. doi: 10.1016/j.jdent.2024.105275
17. Mese I, Gunes ST, Coskun O, Keles A, Banaz T, Kazanbas MF, et al. Reporting Quality of Large Language Model Studies: A Cross-Sectional Audit of High-Ranking Radiology and Medical Imaging Journals. *Korean J Radiol*. 2026;27(9):829. doi: 10.3348/kjr.2026.0505
18. Kong HJ, Kim YL. Large language models in implant dentistry: a scoping review of applications, performance, and limitations. *BMC Oral Health*. 2026;26(1):1492. doi: 10.1186/s12903-026-08747-z
19. KAPLAN TT, AKYOL GE. The role of large language models in dental diagnosis, decision-making, and communication: A systematic review. *Japanese Dental Science Review*. 2026;62:57-67. doi: 10.1016/j.jdsr.2026.01.002
20. Long J, Yang L, Dou J, Zhang L, Tan X. Artificial intelligence in dentistry: a bibliometric analysis. *Br Dent J*. 2025. doi: 10.1038/s41415-025-8885-y
21. Ho YS, Fardi A, Kodonas K. A bibliometric approach to understanding trends and major topics in cone beam computed tomography research. *Journal of Stomatology Oral and Maxillofacial Surgery*. 2026;127(4):102769. doi: 10.1016/j.jormas.2026.102769
22. Kurmanalina M, Ismagulova MB, Tamadon A. The use of AI in the diagnosis of oral diseases: a bibliometric analysis. *J CLIN MED KAZ*. 2025;22(3):8-15. doi: 10.23950/jcmk/16327
23. Lin Y, Choi Y, Park W. Mapping the structure of an emerging field: A scientometric decoding of large language model applications in the dental field. *Japanese Dental Science Review*. 2026;62:225-238. doi: 10.1016/j.jdsr.2026.06.002
24. Li D, Liu Z, Luo B, Hu J, Jiang Q. AI in Fixed Prosthodontics: A Bibliometric and AI-Assisted Qualitative Analysis of Research Trends and Translational Gaps (2011–2026). *J Craniofac Surg*. 2026. doi: 10.1097/scs.00000000000013406

25. Hazrati P, Zibandehkhooy T, Sabri H, Romandini M, Barootchi S. Mapping 60 Years of Discovery: An AI - Driven Bibliometric and Altmetric Analysis of the Journal of Periodontal Research. *J of Periodontal Research*. 2025;60(12):1181-1200. doi: 10.1111/jre.70071
26. Iniesta M, Pérez-Higueras JJ. Global Trends in the Use of Artificial Intelligence in Dental Education: A Bibliometric Analysis. *Eur J Dental Education*. 2026;30(2):427-438. doi: 10.1111/eje.13133
27. Asif S, Hadi F, Qurrat-ul-ain, Yan Y, Wang VY, Xu D. The impact of large language models on medical research and patient care: A systematic review of current trends, challenges, and future innovations. *Computer Science Review*. 2026;59:100847. doi: 10.1016/j.cosrev.2025.100847
28. Mansoor I, Abdullah M, Rizwan MD, Fraz MM. Reasoning with large language models in medicine: a systematic review of techniques, challenges and clinical integration. *Health Inf Sci Syst*. 2025;14(1):6. doi: 10.1007/s13755-025-00403-0
29. Wimalarathna A. Revolutionizing dental education: harnessing the power of ChatGPT for personalized learning in dentistry. *JDSE*. 2024;2(2):40-42. doi: 10.51271/jdse-0031
30. Ribeiro CG, de Oliveira RG, Porto FR, Hespanhol FL, Ferraz Júnior AML. Artificial Intelligence in Dental Examinations: Performance Analysis of ChatGPT , Gemini, and Undergraduate Students. *Eur J Dental Education*. 2026:eje.70319. doi: 10.1111/eje.70319

FUNDING

No financing

CONFLICT OF INTEREST

None

AUTHORSHIP CONTRIBUTIONS

Drafting – original draft: Cristina Valeria Zambrano-Olvera, Rosaria Perez-Rizzo.

Writing–review and editing: Cristina Valeria Zambrano-Olvera, Rosaria Perez-Rizzo.