

Beyond Transcription

Why dental AI scribes must produce structured SOAP notes

DentScribe.AI | Version 1.0

White paper

The patient has left. The next patient is waiting. Your AI note reads well. But did it keep the distal-buccal caries separate from the lingual fracture? Did it remember that the patient deferred the crown? Did it record the second anesthetic carpule as requested, or as administered?

Those are documentation questions a dentist should be able to answer at a glance. A fluent paragraph can conceal the wrong answer.

DentScribe.AI publishes Dental AI SOAP Note Benchmark 1.0 to make note quality visible. It combines ten focused challenges with two source-based, scattered-conversation challenges, expected clinical content, and a straightforward comparison sheet.

Our position

We believe DentScribe.AI is the best technology for creating dental SOAP notes. We are putting that conviction to a practical test: ask DentScribe and any alternative to document the same encounters, then judge the final notes yourself.

The quality of the clinical record is the standard we want customers to demand. A good scribe preserves what was found, where it was found, what the clinician concluded, what actually happened, and what remains to be done.

Three outputs that are easy to confuse

Output	What it gives you	What remains to be checked
Transcript	The spoken conversation in text.	Clinical relevance, attribution, corrections, organization, and decisions.
Narrative summary	A shorter, smoother account of the visit.	Whether compression removed teeth, surfaces, measurements, exceptions, or conditions.
Structured dental SOAP note	A clinical record organized for review and follow-through.	Whether every important statement is supported, correctly linked, and consistent.

S, O, A, and P stand for Subjective, Objective, Assessment, and Plan. These represent different kinds of information. Adding those headings to generic prose does not, by itself, preserve the distinctions.

The ADA describes SOAP as a way to organize patient-specific documentation and encourages tooth-by-tooth status and procedure detail. Benchmark 1.0 turns those practical expectations into visible tests of AI-generated notes. [1]

Clinical structure keeps the relationships intact

The difficult part is often the relationship between facts. A tooth can have an existing restoration, new decay on one surface, a fracture on another, a preferred treatment, and an alternative that depends on cost. A useful note keeps those facts together without blending them.

A dense restorative example

Encounter detail. #31 has distal-buccal caries and a lingual fracture. A crown is recommended; an MLB restoration is considered if the patient cannot proceed with the crown. Six other teeth have surface-specific monitoring instructions. No restorative care occurred today.

A readable but inadequate summary. Multiple teeth require monitoring. Decay and fractures were discussed, with options for a crown or filling.

Structured clinical content. One acceptable rendering is:

- #31: DB caries; L fracture. Crown recommended; MLB restoration considered as an alternative if the patient cannot proceed with the crown.
- #30D, #8MD, #9M, #13D, #4D, #3M: Monitor.
- No restorative treatment performed today.

The difference is observable: six monitoring targets remain identifiable, the two findings on #31 keep their surfaces, and the alternative remains conditional. Separate bullets can do the same job; grouping is useful only when it preserves the meaning.

A periodontal example

A patient can have generalized bone loss and recession, localized pocketing, a healing ulcer between #13 and #14, Grade I furcation involvement at #30, and concentrated occlusal loading at #31. The note must preserve both the mouth-wide findings and the local exceptions.

Required distinction	What a vague summary can lose
At least 4 mm bone loss versus 4–5 mm pockets	The type of measurement and the minimum or range.
#30 furcation and possible grafting	The specific tooth and the difference between consultation and a definite procedure.
#31 loading and planned bite adjustment	A separate problem with a separate action.
Limited bleeding and inflammation	The actual finding, rather than an invented absence or greater severity.
Brushing-related contributors and advice	The clinician’s stated reasoning and the education already delivered.

“Gum disease; deep cleaning discussed” cannot substitute for this record. The full restorative and periodontal challenges appear in Cases 1 and 2 of Benchmark 1.0.

Dentists do not speak in SOAP order

An examination is a conversation. A patient mentions a symptom, the dentist checks a tooth, someone asks about medication, the discussion shifts to travel, and the clinician returns to the original tooth with a different plan. Findings arrive in fragments, not finished sections.

The source transcripts make this visible. We reviewed complete encounters and selected two for chronological walkthroughs in Benchmark 1.0. Their public excerpts retain interruptions, partial statements, late additions, and changes in procedure count while omitting identifying material.

One consultation revisits the same problems repeatedly

The opening request concerns the back of #14. Root-surface decay is discussed later. The patient's temporary dry-mouth complaint appears in another passage. Much later, rapid charting adds #14 distal and #19–#20 distal decay. The clinician then returns to how #14 might be restored, discusses fillings and a crown, and adds another restorative area near the end.

The note has to join the tooth, surface, finding, and conditional plan across those passages. It must also keep patient-reported dryness separate from the clinician's assessment of acid-related wear. Where an excerpt does not establish the tooth for a crown replacement, guessing would defeat the point.

A procedure continues while the conversation wanders

In the extraction encounter, the explanation and numbing steps occur early. “One down” appears during family small talk. A story about another patient's drained abscess interrupts the procedure narrative. Later, the discussion changes from two extractions to four, and a subsequent statement confirms four teeth were removed. Gauze instructions and advice about soft foods arrive near the end.

A useful note assembles the completed extractions and after-care while keeping the other patient's abscess out. It does not invent the extracted tooth letters or an anesthetic dose simply because those would be convenient fields to fill.

The test that a tidy example cannot replace

Give the scribe the conversation in its original order. Do not sort the findings into examination, diagnosis, treatment, and plan before testing. Then ask whether a late correction updated the right earlier fact, whether repeated mentions became one coherent entry, and whether each decision retained its patient and tooth.

Cases F1 and F2 provide public excerpt tests. A practice should also run complete authorized encounters without pre-editing. Short cases expose individual weaknesses; long, interrupted conversations reveal whether the scribe can construct the whole note.

Complexity extends across the dental visit

The benchmark goes beyond restorative and periodontal summaries. These additional examples test whether the note preserves clinical meaning as it organizes a conversation. Every example below is synthetic; none is a measured vendor output.

Situation	What the structured note must preserve	A misleading shortcut
Site measurements	#30MB: corrected probing depth 5 mm, recession 2 mm, bleeding. #31MB: probing 6 mm with bleeding and supuration.	Both teeth have 6 mm pockets.
Mixed dentition	#AMO and #JDO caries; restorations planned. #30 composite already exists. Parent requests scheduling after a trip.	Fillings completed on the upper molars.
Endodontic assessment	#12 cold pain lasts 30 seconds on testing; percussion non-tender. Possible irreversible pulpitis, pending evaluation; treatment deferred.	Irreversible pulpitis; root canal planned.
Treatment and allergy	Penicillin allergy with hives. One carpule of 2% lidocaine administered; second requested but not given. #14MO composite completed, A2 shade.	Two carpules administered; filling completed.
Implant procedure	Site 20 implant 4.3 × 10 mm; corrected torque 35 Ncm. Site 29 discussed for next month. Midazolam prepared but not administered.	Implants placed at 20 and 29 under sedation.
Orthodontic refinement	Scan completed; #2–#3 space corrected to 0.3 mm and #3–#4 space 0.1 mm. No IPR. Reported wear 16 hours; advised wear 22 hours.	IPR completed; patient wears aligners 22 hours daily.
Soft tissue examination	Right lateral tongue ulcer 6 × 4 mm; no induration. Trauma is possible, unconfirmed. Biopsy considered if persistent at review.	Traumatic ulcer diagnosed; biopsy scheduled.
Unresolved information	Unnumbered chipped crown remains unnumbered. Conflicting #12D margin entries are flagged; unrelated #30 discussion stays excluded.	Crown at #12 needs replacement; #30 abscess.

The task is to document the clinician's interpretation faithfully. A scribe should not turn a possibility into a diagnosis, silence into a negative examination, or a recommendation into performed treatment.

These distinctions affect review, handoffs, and follow-up. The ADA emphasizes accurate records of findings, diagnoses, procedures, plans, and patient communications, with dentist oversight of chart entries. [2]

Make the next demonstration a real comparison

Ask to see the final note

An accurate transcript or a strong examination subsection is not enough. The details must survive in the final note that the dentist reviews. A missing allergy, a lost surface, or a contradiction between treatment and plan is still a failure, even when an earlier screen looked correct.

DentScribe is built around natural chairside conversation, dental SOAP documentation, and practice follow-through. Its wider platform includes CoPilot, GPS morning-huddle support, after-care summaries, specialist reports, and treatment coordinator notes. [3] The benchmark focuses on the clinical note, where those subsequent workflows obtain important context.

For a practice owner, the practical question is how much checking and repair remains. A crown discussion can support follow-up only if the note keeps the correct tooth, the recommendation, and the patient's decision. No documentation score alone establishes a revenue increase.

Use Benchmark 1.0 in four steps

1. Give each product the same ten focused cases and two scattered-conversation cases under equal input conditions. Record its version and template.
2. Keep the first final note. Compare it with the expected clinical content before making edits.
3. Record focused passes out of ten and scattered-conversation passes out of two, plus clinical errors, organization-only issues, and missing outputs. Keep error examples beside the totals.
4. Then test authorized encounters from your own practice separately. Include routine visits and your difficult cases, and measure review and correction time.

A pass requires complete, supported clinical content and usable organization. Equivalent wording and layouts receive the same credit. Public-case results apply to those cases; they do not establish performance on every patient or situation.

Our invitation to dentists

Bring the difficult case to your next DentScribe demonstration: the crowded restorative chart, the periodontal findings with local exceptions, the patient who defers treatment, or the procedure with corrected measurements. Ask to see what survives in the final SOAP note.

We believe the comparison will make DentScribe's dental focus clear. Use the benchmark, inspect the notes, and choose the technology that earns your confidence.

[Explore DentScribe and request a demonstration](#)

Sources and publication note

[1] American Dental Association — [Templates, Smart Phrases and SOAP](#)

[2] American Dental Association — [Writing in the Dental Record](#)

[3] DentScribe — [Product and workflow overview](#)

Published by DentScribe.AI, a provider evaluated by the same rules. No comparative product results are presented here. The benchmark is a customer evaluation tool, not an independent certification or ADA endorsement. Proprietary prompts and engineering methods are not disclosed.