Journal of Digital Health and Advanced Biomaterials
e-ISSN 3086-5972 Diamond Open Access journal.jdhab.org

Digital outcome measures in orofacial myofunctional therapy

Toward auditable, telepractice-ready, and clinically anchored assessment

Aline Cazante, MA[1]

[1] State Secretariat of Education of Paraná, Curitiba, Paraná, Brazil.

cazante.a@gmail.com • ORCID 0000-0002-4535-5863

Published 14 April 2026 • pp. 33–46 • DOI 10.67463/w9vkca18

ABSTRACT

Objective: To examine how digital outcome measures may strengthen assessment in orofacial myofunctional therapy (OMT) and to propose a practical framework for moving the field beyond observer-dependent reporting toward auditable, clinically anchored endpoints that remain usable in telepractice. **Data sources:** Recent direct and adjacent literature was prioritized, with emphasis on studies and reviews addressing OMT, telepractice, web-based structured assessment, facial video analysis, clinical acoustic markers, instrumented swallowing technologies, digital swallowing care, and governance issues in automated audiovisual data. **Eligibility criteria:** Evidence was retained when it clarified at least one of three questions: which OMT domains are most amenable to digital capture; which adjacent methods are plausibly transferable to OMT assessment; and which methodological conditions are needed if digital endpoints are to improve reproducibility rather than merely add technical novelty. **Methods of synthesis:** This critical narrative review was updated through March 2026. Direct OMT studies and adjacent rehabilitation literature were appraised separately and synthesized through a three-tier evidence model linked to an endpoint-construction pathway: clinical target, elicited task, acquisition conditions, variable extraction, clinical anchoring, and reproducibility testing. **Main findings:** Direct evidence supports structured telepractice workflows, web-enabled assessment scaffolds, and longitudinal patient-reported monitoring; however, it does not yet establish a validated core set of digital OMT endpoints. Adjacent literature provides firmer methodological support for disciplined video-derived facial variables, constrained signal interpretation, clinically benchmarked instrumentation, and explicit governance procedures. The most defensible near-term targets are lip competence and oral resting posture, facial symmetry and movement coordination, task-evoked perioral dynamics, swallowing-adjacent behavior, and longitudinal patient-reported monitoring. **Conclusion:** Progress will depend less on expanding digital feature inventories than on prospectively testing small, condition-sensitive endpoint sets under controlled acquisition, transparent analytical pipelines, explicit clinical anchors, and reproducibility checks across raters, sessions, devices, and sites.

KEYWORDS orofacial myofunctional therapy; digital outcome measures; telepractice; facial movement analysis; swallowing assessment; auditability

EXPLORE Article text References Online record

Introduction

Interest in orofacial myofunctional therapy has expanded across speech-language pathology, sleep-disordered breathing, temporomandibular disorder care, and broader orofacial rehabilitation. Yet the central barrier to cumulative progress is no longer the absence of therapeutic rationale; it is the instability of the outcome architecture used to support therapeutic claims. Recent reviews continue to show marked heterogeneity in protocols, endpoint definitions, and overall study quality[1,2].

The problem is methodological, not cosmetic. Studies still describe improvement through broad formulations such as better function, improved oral habits, enhanced coordination, or greater stability, while leaving the underlying variables, acquisition conditions, and interpretive thresholds only partially specified. Under such conditions, positive findings may be clinically plausible and remain difficult to audit, compare, or reproduce across teams and settings[1,2].

Telepractice-oriented studies have made that weakness harder to ignore. A structured telemedicine-supported model for adults with obstructive sleep apnea and primary snoring showed that remote care can be organized around patient-reported outcomes and longitudinal monitoring, while a 2026 prospective controlled study suggested that telemedicine-delivered myofunctional therapy may coexist with measurable upper-airway remodeling[3,4]. These studies matter methodologically because they shift outcome capture into environments in which acquisition conditions, task design, and follow-up architecture can no longer be treated as incidental.

Adjacent speech-language pathology literature offers a clearer picture of what digitally explicit assessment may look like in practice. The web-based OMES study showed that a structured clinical scaffold can be transferred to a digital environment without loss of clinical logic, with full heuristic satisfaction in 90% of evaluations, a mean overall Computer System Usability Questionnaire score of 1.31, and shorter completion times after the first task[5,6]. At the instrumentation end of the spectrum, work on hybrid multimodal wearable sensing clarifies why future digital endpoints are unlikely to rest on isolated measurements alone: credible systems depend on deliberate signal selection, coherent integration architecture, flexible acquisition platforms, and analytical workflows that remain interpretable to clinicians[17]. Video-based facial movement analysis and methodological work on acoustic markers point in the same direction. Digital measures become useful when acquisition, preprocessing, and interpretation are explicit[9,10].

Swallowing research reinforces the same principle. Telehealth guidance, digital rehabilitation reviews, and instrumented swallowing studies suggest that remote-compatible and sensor-based methods become clinically meaningful only when they are tied to clear tasks, defensible reference standards, and transparent analytical rules[11–15]. In this setting, digital health is best understood less as a technological add-on than as measurement infrastructure. Its value lies in making clinically relevant change more traceable, more comparable, and less dependent on impressionistic reporting.

Review objective

This review is driven by a methodological problem rather than by generalized enthusiasm for technology. The central question is not whether digital tools should be adopted in OMT, but which outcome architecture could make the literature more reproducible, more comparable, and more clinically interpretable.

Three questions organize the discussion: which functional domains in OMT are most amenable to reproducible digital capture; which methods from OMT and adjacent rehabilitation fields are plausibly transferable; and which methodological conditions are required if digital outcome measures are to mature into clinically anchored endpoints rather than merely add technical novelty.

The aim is deliberately practical. Rather than cataloguing tools, this review proposes a restrained set of design principles intended to move the field from locally persuasive reporting toward auditable, clinically usable outcome assessment.

Review method and interpretive strategy

This manuscript was developed as a critical narrative review focused on digital outcome measurement in orofacial myofunctional therapy. It did not pursue PRISMA-style exhaustive retrieval. Instead, it used a targeted search logic updated through March 2026, centered on two evidence families: direct OMT literature with structured or explicitly digital outcome capture, and adjacent digital-rehabilitation literature capable of informing endpoint design. The purpose was not to estimate pooled effects, but to clarify how digital capture, digital outcome measures, and digital endpoints can be designed, interpreted, and judged for near-term use in OMT. The evidence tiers are summarized in Table 1.

Table 1. Evidence grading and immediate research use for digital outcome development in orofacial myofunctional therapy.

Evidence tier	Typical source	What it can support	What it cannot support alone	Immediate use in OMT research
Tier 1: direct OMT evidence	Telepractice OMT studies, web-based OMT assessment, digitally mediated follow-up within OMT populations	Field-specific feasibility, implementation constraints, clinically relevant task candidates	Validated consensus endpoint sets or strong metric generalizability by themselves	Suitable for structured prospective use, but with explicit validation caveats
Tier 2: adjacent transferable evidence	Facial video analysis, acoustic markers, digital dysphagia assessment, instrumented rehabilitation workflows	Acquisition logic, variable design, reference-standard thinking, workflow discipline	Automatic clinical transfer to OMT populations or indications	Suitable for provisional translation when clinically justified and clearly labeled as adjacent evidence
Tier 3: evidence-informed design inference	Framework proposals derived from comparative reading across direct and adjacent literature	Protocol construction, reporting standards, endpoint architecture, research prioritization	Claims of effectiveness, validity, or readiness without prospective testing	Suitable for exploratory design and protocol planning only

Source: Author's synthesis based on references 1–17.

Table note: The table distinguishes what direct OMT evidence, adjacent transferable evidence, and evidence-informed design inference can legitimately support at the current stage of the literature.

Source identification was guided by combinations of terms related to orofacial myofunctional therapy, telepractice, telemedicine, OMES or AMIOFE, facial movement analysis, acoustic markers, sensor-based swallowing assessment, dysphagia telehealth, digital swallowing rehabilitation, auditability, and ethics of automated digital data. Literature published primarily from 2020 onward was prioritized, whereas earlier sources were retained only when they anchored validated clinical frameworks or still-relevant methodological principles[5–7,9–16].

To avoid rhetorical overreach, the literature was interpreted through three tiers. Tier 1 comprised direct OMT studies with structured or digitally mediated outcome capture. Tier 2 comprised adjacent speech-language and rehabilitation literature that contributed transferable logic for acquisition, signal interpretation, telepractice workflow, instrumentation, or governance. Tier 3 comprised evidence-informed design recommendations derived from comparative synthesis across Tiers 1 and 2. Throughout the review, claims are framed according to the highest tier that can reasonably support them, and recommendations are presented as provisional, evidence-informed guidance rather than as field-wide standards.

Priority was assigned to papers that made endpoint architecture explicit: which functional variable was targeted, how the behavior was elicited, under which acquisition conditions it was captured, how the resulting material was processed, and which clinical anchor justified the interpretation. Reports limited to broad claims of improvement without operationally defined endpoints were treated as hypothesis-generating rather than as strong methodological evidence. This review therefore functions less as a catalogue of tools than as a framework paper for constructing and testing auditable endpoints.

Operational framework for endpoint design and evidence grading

For a digital endpoint to be useful in study design or editorial appraisal, it should be built through a fixed sequence: define the phenotype and clinical target, specify the elicited task, standardize acquisition, extract a limited variable set, anchor those variables to recognizable clinical constructs, and test reproducibility across raters or sessions. The software layer matters, but less than that sequence, because the sequence determines whether the result can be interpreted, audited, and compared. The aim is not to declare a validated consensus endpoint set for OMT, but to provide a disciplined framework for designing one.

The evidence base likewise benefits from explicit grading. Direct OMT studies can support claims about feasibility within the field, but they rarely establish robust endpoint validation. Adjacent digital-rehabilitation literature can clarify acquisition logic, workflow discipline, and signal interpretation, but it does not automatically transfer clinically to OMT populations. Recommendations derived from both bodies of literature are useful for protocol construction, yet they remain inferential until prospectively tested within OMT.

Accordingly, this review evaluates endpoint proposals along two axes: proximity of the evidence to OMT itself and immediate readiness for prospective research. That framing helps distinguish what can reasonably be incorporated now from what should still be treated as translational development.

Thematic synthesis of the literature

Why current evidence remains difficult to compare

The principal obstacle to cumulative science in this area is not simply the small number of studies. More fundamentally, it is the instability of the measurement architecture underlying those studies. Even when therapeutic aims appear similar, comparisons remain weak because the literature often collapses distinct target phenomena into a generic language of improvement. Oral resting posture, swallowing-related coordination, speech-adjacent performance, chewing efficiency, symptom burden, and adherence may all be clinically relevant, but they are not interchangeable endpoint classes[1,2].

Task design is a second source of instability. Outcomes are often discussed as though they were stable traits, when in practice many of them are task-dependent behaviors. Lip competence at rest is not equivalent to lip performance during speech initiation, saliva swallow, resisted closure, or repeated oral opening and closing. Likewise, symmetry in a static image is not equivalent to coordination across repeated dynamic tasks. Unless the eliciting task is specified, cross-study comparison remains vulnerable to false equivalence[9,10].

A third weakness lies in the incomplete bridge between structured clinical scales and digital variables. The OMES or AMIOFE family remains valuable because it organizes the examiner's reasoning across appearance, posture, mobility, and function. Yet these instruments were not designed as self-sufficient digital endpoints. At present, their strongest role is to function as clinically recognizable phenotype maps from which narrower digital targets can be selected.

Operationally, this means that a broad clinical item such as lip incompetence should not migrate unchanged into digital research as a vague binary label. It should be decomposed into a task-defined observable and a limited variable set—for example, closure duration at rest, time to re-establish seal after standardized opening, or side-to-side asymmetry during task execution—each interpreted against the original clinical construct. That bridge is what converts structured examination from a documentation scaffold into auditable measurement architecture.

Telepractice introduces an additional comparability risk. Remote care can expand access and support longitudinal monitoring, but it also magnifies variability related to framing, lighting, resolution, network stability, participant positioning, instruction clarity, and the presence or absence of an assistant. For that reason, telepractice conditions should be treated as part of endpoint architecture rather than as minor operational detail[7,8,11].

A further recurring weakness is the failure to separate technical feasibility from clinical validity. Video capture, automated tracking, and wearable sensing can be technically successful while remaining clinically underdetermined if the selected variables do not map onto recognizable functional change. In editorial terms, a technically sophisticated workflow with weakly anchored outcomes remains methodologically immature, whereas a more modest system with explicit tasks, stable acquisition, and interpretable variables may represent the stronger scientific contribution.

Candidate domains for digital outcome capture

A digitally mature framework should begin with functional domains rather than with technology for its own sake. Five domains appear especially relevant in the current literature: lip competence and oral resting posture; facial symmetry and movement coordination; task-evoked perioral dynamics; swallowing-adjacent behavior; and longitudinal patient-reported monitoring. Together, they provide a coherent map of where digital outcome development is most likely to yield clinically useful progress[1–5].

Lip competence and oral resting posture are especially attractive because they are both clinically central and visually accessible. In routine reporting, lip seal is often described categorically as adequate or inadequate. A more reproducible digital approach would instead treat it as a structured functional signal, potentially incorporating closure duration, intertrial variability, compensatory recruitment, and time to re-establish seal after a standardized opening task.

Facial symmetry and movement coordination form a second major domain. Symmetry is not merely an aesthetic observation; it may function as a proxy for coordinated recruitment, temporal consistency, and task stability across repeated trials. Video-based movement analysis is especially relevant because it allows symmetry to be reframed as a measurable property of performance rather than as a static impression[9].

A third domain is perioral dynamics during functional tasks. Many clinically meaningful changes emerge under activity rather than at rest: speech initiation, saliva swallow, resisted lip closure, repeated oral opening and closing, or short motor sequences that combine timing with postural stability. From a digital-outcome perspective, task-evoked behavior may be more informative than static imagery because it reveals variability, fatigue, compensation, and consistency across repetitions.

Swallowing-adjacent behavior represents a fourth domain. OMT is not reducible to dysphagia care, but oral containment, pre-swallow posture, labial seal, and coordination of oral structures remain relevant in several patient groups. Reviews of digital swallowing rehabilitation, AI-supported screening, and multimodal wearable assessment show that clinically meaningful instrumentation requires disciplined task structure, reference benchmarking, and transparent interpretation[12–15].

Finally, patient-reported outcomes and home monitoring should not be treated as peripheral add-ons. Digitally administered symptom scales, adherence logs, perceived effort, sleep-related reports when relevant, and repeated home recordings may capture dimensions of function that remain invisible in a single clinic encounter. Their value depends on being embedded within explicit longitudinal designs and linked to clearly defined clinical targets[3,4].

Telepractice, acquisition standardization, and methodological reproducibility

Telepractice is no longer merely an access solution in speech-language pathology. It has become a methodological condition that forces greater clarity in how functional outcomes are defined, elicited, recorded, and interpreted. In OMT, that shift is especially important because several clinically relevant targets—including lip seal, facial symmetry, perioral coordination, and task execution—are visually accessible and therefore compatible with digitally mediated assessment[3,7,8,11].

A digitally credible study cannot rely on generic statements that video was collected remotely or that follow-up occurred online. It must specify the acquisition architecture of the outcome itself. At minimum, that includes camera position, framing, head orientation, lighting conditions, recording distance, the elicited task set, number of repetitions, task duration, and the exact instructions given to the participant. Without that level of detail, digital endpoints risk reproducing the same heterogeneity that has long constrained observer-based literature in the field[1,2,7,8].

The methodological advantage of digital assessment does not lie in the mere presence of software or video. It lies in the combination of standardized acquisition, interpretable variables, and transparent analytical rules. This principle is consistent with work on clinical acoustic markers and with the swallowing literature, where meaningful digital measures depend on explicit processing steps, reference benchmarking, and careful feature interpretation rather than on uncontrolled data extraction[10,12–15].

Telepractice also creates an opportunity for repeated, ecologically relevant measurement. In rehabilitation, the most informative outcomes are often not captured in a single clinic snapshot, but across serial observations in which task stability, adherence, fatigue, and generalization can be examined. When acquisition is standardized, telepractice-compatible follow-up may allow OMT research to move beyond isolated visual documentation toward structured longitudinal monitoring. Candidate outcome domains are summarized in Table 2.

Table 2. Candidate outcome domains for digitally explicit assessment in orofacial myofunctional therapy.

Domain	Candidate digital variables	Clinical meaning	Tele-practice feasibility	Evidence maturity	Clinical reference standard
Lip competence and oral resting posture	Lip closure duration; time to re-establish seal; resting lip aperture; intertrial variability	Stability of lip seal; oral posture control; perioral efficiency	High	Direct clinical relevance; direct metric validation still limited	OMES-E or AMIOFE, complemented by stopwatch-based observation from 2 blinded raters during a standardized resting-posture and re-sealing task
Facial symmetry and movement coordination	Side-to-side excursion difference; trajectory consistency; asymmetry indices; timing coordination	Recruitment symmetry; dynamic coordination; task stability	Moderate to high	Adjacent validation stronger than direct OMT validation	Structured clinician scoring of facial symmetry and mobility by blinded raters, with frame-based comparison during a standardized movement task
Task-evoked perioral dynamics	Movement amplitude; temporal stability; repetition-related variability; task-dependent recruitment pattern	Motor performance under functional load	Moderate	Conceptually strong; task standardization remains incomplete	Standardized clinical observation of motor execution and task stability, with blinded expert consensus where appropriate
Swallowing-adjacent behavior	Pre-swallow posture; oral containment; timing-related variables; task completion consistency	Oral-phase efficiency; clinically relevant coordination in selected groups	Moderate	Indirect translational support; condition-specific validation needed	Structured oral-phase clinical assessment, complemented by synchronized blinded observation and, where justified, an instrumental reference standard
Longitudinal patient-reported monitoring	Symptom trajectories; adherence logs; perceived effort; repeated home recordings	Change over time; ecological relevance; treatment engagement	High	Feasible now; interpretation depends on tighter linkage to clinical targets	Pre-defined PROMs, standardized digital diaries, and serial clinical assessment over time

Source: Author's synthesis based on references 1–16.

Table note: These domains are proposed as a translational scaffold for future research rather than as a fixed consensus outcome set; the added clinical reference-standard column is intended to convert plausible digital domains into operationally testable endpoint families.

Adjacent digital approaches and translational relevance

From a translational perspective, the most relevant technologies are not necessarily the most sophisticated. They are the ones that align visible orofacial function with repeatable acquisition and clinically interpretable output. Standardized telepractice workflows provide the operational backbone for repeated assessment, adherence monitoring, and remote follow-up. Web-based structured scoring platforms improve examiner consistency, centralize documentation, and create a bridge between clinical scoring and richer digital phenotyping[3,5,7,8].

Video-based facial movement analysis is arguably the most immediately translatable measurement approach in this context. Lip seal stability, side-to-side asymmetry, movement range, timing, and consistency across repetitions can all, in principle, be derived from structured recordings. The bulbar assessment literature shows that clinically relevant inferences can be drawn from video-derived features when acquisition and analysis are disciplined[9]. OMT does not yet have an equally mature evidence base, but it does target behaviors that are highly compatible with this analytical logic. Relevant approaches are summarized in Table 3.

Table 3. Translational relevance of adjacent digital approaches for OMT assessment.

Approach	Current use in adjacent fields	Potential relevance for OMT	Main strength	Main limitation or evidence boundary
Structured telepractice workflows	Remote dysphagia care; longitudinal follow-up; digitally supported rehabilitation	Standardized remote assessment; repeated task capture; monitoring over time	Scalable and clinically transferable	Acquisition variability; feasibility evidence exceeds endpoint validation
Web-based structured assessment tools	Digital clinical scoring; centralized records; stored visual documentation	More consistent examiner logic; easier comparison across visits; structured data capture	Improves traceability and organization	Improves documentation, but does not by itself validate the endpoint
Video-based facial movement analysis	Bulbar assessment; facial kinematics; clinical staging support	Lip competence; asymmetry; perioral coordination; dynamic task analysis	Noninvasive and highly compatible with visible targets	Adjacent validation stronger than field-specific validation
Clinical acoustic markers	Speech science; digital voice and speech analysis	Speech-adjacent timing and coordination variables	Explicit signal-processing logic when well designed	Useful only for narrow questions; high overreach risk
Sensor-based swallowing instrumentation	AI-supported screening; multimodal wearable monitoring; VFSS-benchmarked classification	Methodological model for instrumented capture and criterion benchmarking	Demonstrates how instrumentation can be clinically anchored	Primarily swallowing-focused; transfer to OMT must remain selective

Source: Author's synthesis based on references 3, 5, 7–17.

Table note: The strongest immediate candidates for translation are disciplined telepractice workflows, web-based structured assessment, and standardized facial video analysis; current evidence is stronger for feasibility and measurement logic than for validated endpoint performance within OMT itself.

Signal-based and sensor-based approaches play a complementary role. Clinical acoustic markers are not direct substitutes for myofunctional examination, but they illustrate how digital variables should be selected, constrained, and interpreted. In telepractice-compatible workflows, however, acoustic comparability depends on more than simply obtaining a recording. Reverberation, background noise, microphone placement, automatic gain control, device-specific frequency response, and compression can all alter formant structure, jitter-like measures, and other timing- or spectrum-dependent outputs. For that reason, studies using speech-adjacent or swallowing-adjacent audio should report the recording environment, reverberation control or room conditions, noise-control strategy, gain-normalization logic, and the type and position of microphone used, while avoiding clinically interpretive claims when capture relies on uncalibrated consumer devices with unknown response characteristics. Methodological guidance from voice research further suggests that valid acoustic analysis depends on sufficiently linear microphone response across the frequency range of interest; in practical terms, audio-derived variables should therefore be treated as restricted, context-sensitive candidate measures unless minimum signal conditions are explicitly declared and reasonably stable across sessions and devices[10,12,13,17,19].

Digital swallowing rehabilitation platforms add a further layer of relevance. OMT is not identical to dysphagia care, yet both areas depend on visible behavior, task-based performance, repeated monitoring, and the need to distinguish technical readout from clinical meaning. The strongest near-term strategy for OMT is therefore not maximal automation, but careful integration of feasible tools that improve consistency, auditability, and longitudinal comparison without obscuring clinical interpretation[14–16].

Minimum reporting items for digitally explicit OMT studies

Future studies will not become methodologically stronger merely by adding remote appointments, video capture, or app-based follow-up to otherwise conventional designs. The substantive advance is architectural: a clinically defined target must be translated into a task, the task into controlled acquisition, the acquisition into a restricted variable set, and those variables back into clinically interpretable evidence. Framed this way, digitally explicit OMT research is best understood as a problem in auditable measurement design rather than as a generic expansion of technology use. That distinction is central to this review's contribution.

At the clinical level, authors should define the target phenotype, therapeutic indication, and primary outcome domain with enough precision to distinguish, for example, lip competence, oral resting posture, facial symmetry, perioral task performance, swallowing-adjacent behavior, and patient-reported change. At the intervention level, they should report treatment logic, session frequency, total duration, supervision model, and whether care was delivered in person, remotely, or in hybrid form. At the task level, they should specify elicited activities, number of repetitions, prompts, and the exact moment at which outcome capture occurred[1,2,3,7,8].

Digital acquisition should be treated as part of study design rather than as a technical footnote. Reports should include camera position, framing, distance, lighting, head orientation, environmental constraints, device type, platform, software version, and relevant settings whenever analysis depends on digital tools. When more than one modality is used, sensor placement, temporal synchronization, fusion logic, and modality-specific quality thresholds should likewise be stated explicitly. When preprocessing or annotation is involved, the workflow should be described in sufficient detail, including landmark-selection rules, filtering, segmentation, manual correction steps, and quality-control thresholds[9,10,12,13,17].

A further requirement is a clear distinction between technical validity and clinical relevance. Digitally explicit studies should predefine a restricted set of primary variables, justify their physiological meaning, and state which clinical anchor or reference standard is being used. When human judgment remains part of the workflow, intra-rater and inter-rater reliability should be reported. Unusable recordings, missing data, telepractice constraints, and governance procedures should likewise be described as part of the methodological logic rather than left implicit[10,12–16].

Terminological discipline also matters. In this review, digital capture refers to raw audiovisual or sensor acquisition; a digital outcome measure refers to the operational metric or measurement procedure derived from that capture; and a digital endpoint refers to the final clinically interpretable variable or summary used for inference, longitudinal comparison, or validation against a clinical anchor. Keeping these levels distinct helps avoid ambiguity about whether the object under evaluation is the recording workflow, the extracted metric, or the clinically meaningful result. More broadly, this hierarchy helps prevent a common category error in the literature: treating technically extractable signals as though they were already validated clinical endpoints.

Minimum technical acquisition requirements and hardware-agnostic implementation

For a digital endpoint to reflect patient function rather than device-specific artifact, acquisition must satisfy minimum technical requirements that are explicit, reproducible, and reportable. In telepractice-compatible OMT studies, a defensible provisional minimum is 1280×720 pixels at 30 frames per second, with 1080p preferred when bandwidth and device capacity permit and 60 fps preferable when the task includes rapid perioral or swallowing-adjacent events. The frame-rate requirement should be tied to task speed rather than treated as a generic specification: under Nyquist logic, sampling must remain sufficiently high to represent the fastest clinically relevant movement components without substantial temporal distortion. The camera should be positioned at eye level, in approximate 90° alignment with the frontal facial plane, and at a pre-specified distance of 50 cm to 1.0 m in order to reduce geometric variation across sessions. Framing should include, at minimum, the segment from the top of the head to the base of the neck so that stable anatomical references remain available for longitudinal comparison and, where relevant, facial landmark tracking[7,8].

Illumination should be frontal, homogeneous, and diffuse, avoiding backlighting, strong unilateral shadowing, and localized glare that may alter perceived asymmetry, labial contour, or landmark visibility. This point is especially important on moist reflective surfaces such as the lips and visible oral mucosa, where specular highlights may saturate pixels and destabilize landmark detection; when feasible, protocols should therefore control not only light direction and intensity but also white balance and reflective artifacts. In remote protocols, ultrawide lenses, beauty filters, aggressive digital stabilization, and excessive zoom should be avoided or explicitly documented. File format should likewise be treated as part of acquisition quality rather than as a neutral storage choice: lossy compression, aggressive bitrate reduction, and spatial smoothing may introduce block artifacts, blur edges, and shift sub-pixel landmark coordinates relative to the underlying anatomy. Whenever feasible, studies should report device brand and model, operating system, capture application, effective resolution, frame rate, camera orientation, acquisition distance, and the compression pathway or codec used for analytical files. In this context, hardware-agnosticism does not mean ignoring the device; it means controlling minimum acquisition conditions so that the endpoint remains interpretable across iPhone, Android, or tablet platforms, provided that the device satisfies pre-defined technical thresholds. Together, these requirements define the minimum technical conditions under which digitally derived facial variables can begin to claim clinical interpretability rather than merely technical extractability[7,8,11].

A practical methodological precedent for acquisition standardization is a session-oriented photographic governance framework that integrates structured capture, quality-control checkpoints, traceable session metadata, and reproducible follow-up documentation. Although not developed as a validation study for OMT endpoints themselves, this model is relevant because it operationalizes capture discipline, longitudinal comparability, and record governance at the level of the acquisition workflow[18].

Adverse-data handling, signal-quality targets, and exclusion rules

Auditability depends not only on describing the analytical pipeline when data are usable, but also on declaring what happens when they are not. Digitally explicit OMT studies should therefore predefine signal-quality targets and operational decision rules for retaining, reacquiring, segmenting, or excluding audiovisual material. In facial video analysis, four failure classes are especially relevant: clinically meaningful occlusion of the region of interest, excessive head rotation, lighting instability, and inconsistent landmark tracking[9,10,12,13,16].

As an initial operational standard, recordings should be excluded or reacquired when more than 20% of relevant facial landmarks are persistently occluded, when sustained head rotation exceeds 15°, when facial framing is incomplete, when illumination precludes stable landmark detection, or when the task is performed outside the standardized instruction set. The protocol should further distinguish between errors that are frame-level and recoverable, allowing limited exclusion of affected frames, and errors that invalidate the entire trial and require reacquisition. In repeated tasks, a minimum of 3 valid repetitions is recommended, with the number of discarded attempts and the reason for exclusion reported explicitly.

When multimodal capture is used, including audio or additional sensors, signal quality should also include verification of temporal synchronization, file integrity, timestamp stability, and gross signal plausibility. The goal is not to maximize retention at any cost, but to ensure that clinical inference is based only on material that satisfies pre-declared quality thresholds. Future studies should report exclusion rates, operational reasons for unusable recordings, and the distribution of failure across device, environment, participant, and operator. Those losses are not peripheral; they define the real boundary conditions of validity.

Clinical anchoring, reference standard, and cross-validation strategy

The notion of clinical anchoring should be operationalized more explicitly. For each proposed digital variable, the manuscript should state which clinical reference standard is being used and how correspondence will be tested. Rather than claiming in general terms that a digital endpoint maps onto recognizable function, the protocol should declare whether validation is being performed against OMES-E, AMIOFE, structured clinician observation, blinded stopwatch-based human timing, expert consensus, or a combination of these approaches[5,6].

In near-term applications, the most defensible strategy is digital-versus-human cross-validation, in which the digital variable is compared against a prespecified clinical reference under a standardized task, with human raters blinded to digital output whenever possible. This makes it possible to separate two distinct levels of evidence: first, the technical stability and reproducibility of the extracted variable; second, its clinical interpretability relative to a recognized reference construct. A digital endpoint becomes methodologically stronger only when both levels are specified and tested.

Audiovisual data governance, de-identification, and privacy protection

In studies based on facial video, image-derived features, or other audiovisual material, data governance should be treated as part of methodological validity rather than as an administrative afterthought. The protocol should explicitly state whether the workflow involves storage of raw video, retention only of derived numerical features, or a hybrid architecture in which raw material is temporarily retained for quality control and auditability before scheduled deletion. This distinction matters because facial images may function as directly identifying material, while landmarks, vectors, and derived templates may still retain sensitive inferential value depending on the processing pathway[16].

Whenever possible, studies should adopt a data-minimization architecture in which only the smallest amount of identifiable material necessary to answer the clinical question is retained. This includes clear separation between identifying files and analytical derivatives, access control, retention rules, audit logs, and defined destruction timelines. If landmarks, vectors, or derived metrics constitute the primary analytic object, the manuscript should say so explicitly and should also declare the policy governing raw video. In telepractice designs, the protocol should further specify whether processing occurs locally, on institutional infrastructure, or through a third-party environment, and should describe how those arrangements are aligned with applicable privacy and data-protection requirements. A condensed operational checklist and a model cross-validation matrix are provided in Appendices A and B[16,18].

Knowledge gaps and research priorities

Despite growing interest, the literature still lacks an outcome architecture stable enough for high-confidence comparison across studies. The first major gap is the absence of a shared logic for how a digital OMT endpoint should be built, graded, and interpreted. What the field needs is not merely a list of candidate metrics, but explicit rules linking phenotype, task, acquisition, variable, anchor, and reliability.

A second gap concerns validation within OMT itself. The adjacent literature is now sufficiently mature to show that telepractice workflows, video-based facial analysis, AI-supported screening, multimodal wearables, and digitally supported swallowing care can generate structured and clinically interpretable assessment logic. Yet those methods have not been translated into enough standardized endpoint studies within OMT to justify routine methodological confidence[9,12–15].

A third gap is the limited distinction between technical feasibility and clinical validity. The most useful candidate metrics will not necessarily be the most computationally complex. They will be the measures that remain stable across sessions, interpretable to clinicians, sensitive to therapeutic change, and workable in both in-person and telepractice settings[10,16].

A fourth gap is the scarcity of multicenter reproducibility work. Many promising digital approaches remain tied to single teams, single workflows, or highly local expertise. For the field to develop a cumulative digital-outcome literature, future studies will need to test whether the same acquisition protocol, task set, and analytical rules yield comparable results across clinicians, devices, and institutions.

These gaps carry direct translational implications. Telepractice should be understood not only as a mode of service delivery, but also as a stress test for measurement rigor. Web-based structured tools should be developed not merely for documentation efficiency, but as bridges between clinical scoring and richer digital phenotyping. Instrumented approaches should be incorporated only when they add a clinically interpretable layer of evidence rather than a technically impressive but weakly anchored one.

A clinically useful next step would be the development of condition-sensitive digital outcome pathways rather than a disease-agnostic template. OMT in obstructive sleep apnea, temporomandibular disorders, oral habit modification, and broader orofacial rehabilitation may share common methodological infrastructure, but they do not necessarily require the same primary endpoint. The field will advance more coherently if it adopts a shared measurement logic while allowing condition-specific endpoint profiles.

Another priority is the integration of clinician-rated outcomes with patient-generated data. One of the more promising features of digital assessment in OMT is the possibility of linking structured clinical examination to repeated home-based observations, adherence information, and symptom trajectories. That promise will remain methodologically weak, however, if home-generated data are collected without explicit quality rules or without a clear plan for how they will inform interpretation.

Adverse-case documentation is also underdeveloped. Digital outcome studies tend to emphasize successful remote encounters and usable recordings, but a mature literature must also report where workflows fail: poor camera positioning, inadequate lighting, low adherence, unreliable task execution, difficulty visualizing intraoral targets, mismatch between platform prompts and patient capability, or inability to obtain acceptable data quality. Those negative operational findings define the boundary conditions of validity.

What auditability means in practice

The concept of auditability warrants explicit definition because it sits at the center of the transition proposed here. In digital assessment for OMT, auditability is not simply a matter of storing files or archiving outputs. It refers to whether another clinician, reviewer, or research team can reconstruct how an endpoint was generated, under which acquisition conditions it was collected, which preprocessing or scoring steps were applied, and why the final interpretation was considered clinically meaningful.

In practice, an auditable study in this domain would document the elicited task, acquisition setup, device and platform, operator instructions, quality-control thresholds, annotation or preprocessing workflow, primary variables, reference standard, and rules for excluding unusable recordings. If the endpoint is multimodal, it should also declare how signals were synchronized, whether features were fused early or late, and which modality was expected to carry the decisive clinical information. It should further specify which elements of interpretation remained judgment-dependent and how the reliability of those steps was evaluated. When feasible, representative raw or semi-processed outputs should be preserved so that independent readers can judge whether the final claim is proportionate to the observed data.

Auditability also serves a translational purpose. Clinicians are more likely to trust digital outcomes when they can see how those outcomes were produced and how they map onto familiar functional reasoning. In that sense, auditability is not a bureaucratic burden imposed on innovation; it is part of clinical credibility.

Limitations

This review is narrative and interpretive rather than systematic. It does not claim exhaustive retrieval, pooled effect estimation, or formal risk-of-bias synthesis. Another limitation is that the strongest methodological guidance still comes largely from adjacent fields rather than from OMT trials themselves. The argument therefore rests on converging methodological evidence, not on a mature field-specific digital literature[1,2,9,12–15].

A further limitation is temporal. Digital tools, web platforms, and remote-care architectures are evolving quickly, and the methodological landscape may shift faster than the conventional rehabilitation literature. For that reason, this review should be read less as a closed taxonomy of technologies than as a framework for judging which approaches are likely to be scientifically useful: those that improve traceability, reproducibility, and clinical interpretability rather than merely increasing technical novelty.

Conclusion

The present evidence base supports a precise, rather than expansive, conclusion. Direct OMT literature already justifies structured remote follow-up, digitizable assessment scaffolds, and patient-centered longitudinal monitoring. It does not yet justify a validated consensus set of digital endpoints with demonstrated cross-site reproducibility. At this stage, the more robust methodological guidance still comes mainly from adjacent work in facial video analysis, acoustic and multimodal sensing, instrumented swallowing assessment, telehealth, and data governance[1–17].

Observer-dependent documentation is no longer sufficient on its own, but it should not be discarded. Clinical judgment remains indispensable.

What must change is the structure around that judgment: predefined tasks, controlled acquisition, clinically anchored variables, explicit analytical rules, and transparent handling of missing or unusable data.

The field is ready for structured digital endpoint design, but not yet for consensus endpoint adoption. It does not need an unlimited expansion of digital features. It needs smaller, condition-sensitive endpoint families that can be prospectively validated against clinical anchors and tested for reproducibility across raters, sessions, devices, and sites. Only under those conditions will digital assessment move from technical feasibility to reproducible evidence.

Required statements

Ethics approval: Not applicable. This study is a narrative review of previously published literature and did not involve human participants, animals, or identifiable personal data.

Consent to participate: Not applicable.

Consent for publication: Not applicable.

Competing interests: The author declares no conflicts of interest related to this work.

Editorial role disclosure (added 18 October 2026): Aline Cazante is listed as Copyediting and Language Editor on the journal's masthead. OJS account records show that role assigned from 1 to 3 April 2026, overlapping the start of this manuscript's evaluation. Reviewer assignments and acceptance on 11 April 2026 are recorded under Editor-in-Chief Lisandro Gonçalves; no such actions are attributed to the author in those records. This disclosure supplements the competing-interests statement and does not change the scientific content or the original publication date.

Data availability: No new data were generated or analyzed in the preparation of this review.

Funding: This work received no external funding.

Code availability: Not applicable.

AI use disclosure: An AI-assisted image-generation tool was used exclusively to prepare Figure 1, which is presented as a conceptual schematic under direct human supervision. No AI tool was used for data generation, reference selection, scientific interpretation, or development of the manuscript conclusions. The author critically reviewed all content and assumes full responsibility for the final manuscript.

Author contributions: Aline Cazante was responsible for conceptualization, literature review, evidence synthesis, interpretation of findings, original draft preparation, manuscript review and editing, and approval of the final version.

Acknowledgements: The author has no acknowledgements to declare.

Figure 1 — Conceptual framework for auditable digital outcome assessment in orofacial myofunctional therapy



Source: Conceptual schematic developed by the author from the evidence synthesized in this review.
Panel A outlines the operational pathway linking clinical target definition to task specification, acquisition standardization, restricted variable extraction, clinical anchoring, and reproducibility testing. Panel B summarizes candidate digital outcome domains with their translational readiness. Note: Digital measures require explicit task definition, acquisition control, clinical anchoring, and reproducibility to be effective.

References

1. Shortland HA, Hewat S, Vertigan A, Webb G. Orofacial myofunctional therapy and myofunctional devices used in speech pathology treatment: a systematic quantitative review of the literature. Am J Speech Lang Pathol. 2021;30(1):301-17. https://doi.org/10.1044/2020_AJSLP-20-00245.
2. Stefani CM, de Almeida de Lima A, Stefani FM, Kung JY, Flores-Mir C, Compton SM. Effectiveness of orofacial myofunctional therapy in improving orofacial function and oral habits: a scoping review. Can J Dent Hyg. 2025;59(1):59-72.
3. Capacho EER, Bossa CPD, Campos MDC, Rincon-Yanez D, Rangel-Navia H, Bianchini EMG. Telemedicine-supported structured orofacial myofunctional therapy model for obstructive sleep apnea: patients' report outcomes measurements. Respir Med. 2025;249:108460. https://doi.org/10.1016/j.rmed.2025.108460.

4. Rodríguez-Alcalá C, Rodríguez-Alcalá L, Ignacio-García JM, Plaza G, Gozal D, Baptista P, et al. Telemedicine-delivered myofunctional therapy remodels upper airway anatomy in obstructive sleep apnea: a prospective controlled study. J Clin Sleep Med. 2026;22(1):12. https://doi.org/10.1007/s44470-025-00008-0.
5. Ataide MCG, Bernardi FA, Marques PMA, de Felício CM. Web version of the protocol of the orofacial myofunctional evaluation with scores: usability and learning. CoDAS. 2023;35(2):e20220026. https://doi.org/10.1590/2317-1782/20232022026.
6. de Felício CM, Ferreira CLP. Protocol of orofacial myofunctional evaluation with scores. Int J Pediatr Otorhinolaryngol. 2008;72(3):367-75. https://doi.org/10.1016/j.ijporl.2007.11.012.
7. Boshart CA, James S, Fisher PK. Guidelines for successful online service delivery in orofacial myology. Int J Orofac Myol Myofunct Ther. 2020;46(1):59-76. https://doi.org/10.52010/ijom.2020.46.1.6.
8. Amaral MS, Furlan RMMM, Almeida-Leite CM, Motta AR. Orofacial myofunctional speech-language-hearing teletherapy program focusing on chewing and swallowing in temporomandibular disorders. Rev CEFAC. 2025;27(3):e1524. https://doi.org/10.1590/1982-0216/20252731524.
9. Guarin DL, Taati B, Abrahao A, Zinman L, Yunusova Y. Video-based facial movement analysis in the assessment of bulbar amyotrophic lateral sclerosis: clinical validation. J Speech Lang Hear Res. 2022;65(12):4667-78. https://doi.org/10.1044/2022_JSLHR-22-00072.
10. Schultz BG, Vogel AP. A tutorial review on clinical acoustic markers in speech science. J Speech Lang Hear Res. 2022;65(9):3239-63. https://doi.org/10.1044/2022_JSLHR-21-00647.
11. Malandraki GA, Hahn Arkenberg R, Mitchell SS, Bauer Malandraki JL. Telehealth for dysphagia across the life span: using contemporary evidence and expertise to guide clinical practice during and after COVID-19. Am J Speech Lang Pathol. 2021;30(2):532-50. https://doi.org/10.1044/2020_AJSLP-20-00252.
12. Wong DWC, Wang J, Cheung SMY, Lai DKH, Chiu ATS, Pu D, et al. Current technological advances in dysphagia screening: systematic scoping review. J Med Internet Res. 2025;27:e65551. https://doi.org/10.2196/65551.
13. Shin B, Lee SH, Kwon K, Lee YJ, Crispe N, Ahn SY, et al. Automatic clinical assessment of swallowing behavior and diagnosis of silent aspiration using wireless multimodal wearable electronics. Adv Sci (Weinh). 2024;11(34):e2404211. https://doi.org/10.1002/advs.202404211.
14. Hwang NK, Yoon TH, Chang MY, Park JS. Dysphagia rehabilitation using digital technology: a scoping review. J Evid Based Med. 2025;18(1):e70009. https://doi.org/10.1111/jebm.70009.
15. Alter IL, Dias C, Briano J, Rameau A. Digital health technologies in swallowing care from screening to rehabilitation: a narrative review. Auris Nasus Larynx. 2025;52(4):319-26. https://doi.org/10.1016/j.anl.2025.05.002.
16. Elbaum B, Perry LK, Sarangoulis CM, Goodman KW, Messinger DS, Cejas I. The use of automated digital data in speech, language, and hearing research: confronting a new ethical landscape. J Speech Lang Hear Res. 2025;68(8):4087-93. https://doi.org/10.1044/2025_JSLHR-24-00819.
17. Mahato K, Saha T, Ding S, Sandhu SS, Chang AY, Wang J. Hybrid multimodal wearable sensors for comprehensive health monitoring. Nat Electron. 2024;7:735-50. https://doi.org/10.1038/s41928-024-01247-4.
18. Gonçalves L, Silva D, Galdino KS. Mapping botulinum toxin injection sites and automated unit totalization in orofacial harmonization: a session-based web module for standardized, traceable clinical imaging. Derecho y Cambio Social. 2026;23(88):e4842. https://doi.org/10.54899/dcs.v23i88.4842.
19. Švec JG, Granqvist S. Guidelines for selecting microphones for human voice production research. Am J Speech Lang Pathol. 2010;19(4):356-68. https://doi.org/10.1044/1058-0360(2010/09-0091).

License and reuse

CC BY 4.0

© 2026 Aline Cazante. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). https://creativecommons.org/licenses/by/4.0/

Publisher notice on AI, text and data mining

This article is made available under the Creative Commons Attribution 4.0 International License (CC BY 4.0). Reuse, including automated text and data mining and machine-learning applications, is permitted under the terms of the license, provided that appropriate credit is given, a link to the license is supplied, and any changes are indicated. Users remain responsible for compliance with applicable legal and ethical requirements, including privacy and data-protection obligations.

Article record

Canonical citation
Cazante A. Digital outcome measures in orofacial myofunctional therapy: Toward auditable, telepractice-ready, and clinically anchored assessment. J Digit Health Adv Biomater. 2026;1(1):33–46. https://doi.org/10.67463/w9vkca18.

Assessment Peer-reviewed

Appendix A. Technical checklist for standardized digital acquisition in OMT

Environment and device

- ☐ Camera positioned at eye level
- ☐ Approximate frontal alignment with the facial plane
- ☐ Framing from the top of the head to the base of the neck
- ☐ Fixed capture distance between 50 cm and 1.0 m
- ☐ Minimum resolution of 720p at 30 fps
- ☐ Diffuse frontal lighting
- ☐ No beauty mode, undocumented ultrawide mode, or excessive digital zoom

Task specification

- ☐ Standardized verbal instruction
- ☐ Pre-defined number of repetitions
- ☐ Rest interval between repetitions
- ☐ Pre-defined moment of outcome capture

Signal quality

- ☐ No clinically relevant occlusion of the region of interest
- ☐ Head rotation within protocol limits
- ☐ Stable landmark visibility and tracking
- ☐ File integrity and temporal alignment when multimodal capture is used

Processing pipeline

- ☐ Software or model identified by name and version
- ☐ Landmarking or annotation workflow described
- ☐ Frame-exclusion and trial-exclusion criteria pre-defined
- ☐ Primary digital variable pre-specified

Clinical anchoring

- ☐ Clinical scale or reference standard declared in advance
- ☐ Human raters blinded when applicable
- ☐ Digital-versus-human comparison rule specified

Governance

- ☐ Raw video retention policy defined
- ☐ Derived feature storage policy defined
- ☐ Access control specified
- ☐ Audit trail maintained
- ☐ Destruction timeline or retention rule declared

Appendix B. Cross-validation matrix for digital variables and clinical reference standards

This matrix translates clinically plausible digital variables into pre-specified validation comparisons suitable for prospective OMT research.

Digital variable: Interlabial distance at rest

Standardized task
Resting lip posture for 10 seconds

Comparison method
Correlation and agreement analysis

Clinical reference standard
OMES-E or AMIOFE plus blinded clinician assessment

Primary criterion
Intersession stability

Digital variable: Time to re-establish lip seal

Standardized task
Standardized oral opening followed by closure

Comparison method
Mean absolute difference / ICC

Clinical reference standard
Stopwatch-based timing by 2 blinded raters

Primary criterion
Pre-defined acceptable error

Digital variable: Perioral asymmetry index

Standardized task
Repeated functional movement task

Comparison method
Correlation and clinical-category discrimination

Clinical reference standard
Structured clinician symmetry score

Primary criterion
Repetition consistency

Digital variable: Task-related variability metric

Standardized task
Repeated functional sequence

Comparison method
Association with functional category

Clinical reference standard
Clinician judgment of motor stability

Primary criterion
Sensitivity to change