Advisory report: temporal encoding for FA sequences

Evidence-based recommendations following the literature analysis

Tymo van Rijn
CMI – Informatica (1057297)

March 2026

Abstract

This document turns the findings of `02_Analysis/Analysis_Report_Temporal_Encoding.pdf` into actionable advice for the next design and realisation phases. The core message is that temporal information should be represented explicitly, that binning remains a legitimate baseline, and that design choices should be compared with at least one alternative encoding rather than pursued in isolation. The ongoing Temporal Separability Benchmark (TSB), documented in `02_Analysis/week_5_6.pdf` and `02_Analysis/Temporal_Separability_Benchmark-Study-Proposal.pdf`, is positioned as the empirical instrument to rank binning configurations *within* each foundation model before committing to full sequence training.

1 Purpose and scope

The literature analysis showed that "temporal encoding" is broader than temporal binning alone: recent work consistently treats time (intervals, irregular spacing, multi-scale structure) as a first-class signal rather than as plain frame order. This advisory report answers a practical question for the project: *what should be built next, and how should success be judged*, given limited supervision time and compute.

The recommendations below are deliberately short. They are meant to be read alongside the analysis PDF (sources and nuance) and the week 5–6 slides (TSB protocol, sanity checks, and design refinements).

2 What the literature implies (condensed)

Across the reviewed studies, four transferable patterns recur:

- **Time-aware embeddings.** Represent elapsed time or visit intervals with learned embeddings that are added to, or fused with, visual features (longitudinal transformers, T-Rep-style thinking, L-MAE).
- **Irregular timing.** Medical sequences are rarely uniformly spaced; positional encodings should not assume equal Δt when real timestamps exist.
- **Change-focused objectives.** Contrastive or reconstruction objectives that stress *evolution* between timepoints can complement naive sequence stacking.
- **Multi-scale structure.** Some signals live at short horizons (phase-to-phase) and others at longer spans; hierarchical or cycle-aware encodings are relevant when complexity is justified.

For fluorescein angiography (FA), the clinical takeaway is unchanged: phases and leakage dynamics are inherently temporal, so a static final-frame shortcut can discard discriminative evidence even when single-frame accuracy looks acceptable.

3 Recommendations

1. **Keep binning as a baseline, not as the only hypothesis.** Uniform or quantile binning with mean-pooled (or otherwise aggregated) embeddings remains a reasonable engineering baseline because it is simple to implement and to debug. It should not, however, be the only encoding under consideration.

2. **Add at least one explicit time signal in the representation.** As a feasible next step aligned with the literature, implement a *time-embedding* path: associate each frame or bin with a scalar time (e.g. elapsed seconds since injection or exam start) and map it to a vector that is fused with the RETFound-Green (or chosen backbone) features before the temporal module. This is the lowest-friction way to test "explicit time" without replacing the whole pipeline.

3. **Use the TSB to de-risk binning choices before large LSTM or transformer runs.** The benchmark proposal defines a linear probe (LDA with Fisher separability) over bin-level vectors, followed by downstream sequence training on the same configuration. Rank configurations **within each foundation model**; treat cross-model comparison of raw probe magnitudes as secondary because embedding scale and pretraining domain differ.

4. **Treat strong early–late separability as necessary but not sufficient.** The week 5–6 work shows that embeddings support an approximately monotonic axis from early to late frames and that separability increases with temporal distance (first vs last outperforms first vs second). That validates that "something temporal" is present in the representation. It does not, by itself, prove that a sequence model will translate that into better pathology classification—hence the planned Spearman correlation between probe rankings and downstream metrics (ROC-AUC, AP) for H1.

5. **Revisit aggregation inside bins if the probe is insensitive.** If mean-pooling washes out discriminative frames, document and test alternatives (e.g. centre frame or attention pooling) as a controlled variable rather than an ad hoc change.

6. **Document negative outcomes.** If a time-embedding variant or a binning strategy does not improve validation metrics, record it briefly (hypothesis, setup, result). Graders reward honest scope control as much as success stories.

4 Success criteria (advice-level)

For the next milestone, "success" should mean:

- At least two encoding strategies (e.g. binning-only vs binning + elapsed-time embedding) are implemented with matched data protocol and comparable compute.

- Reported metrics include discrimination on the clinical target task *and*, where applicable, agreement between LDA-based rankings and downstream ranking (TSB H1).

- Design rationale cites both literature (this advisory chain) and empirical diagnostics (timestamp quality, probe sanity checks).

5 Evidence map

Artifact	Role
`02_Analysis/Analysis_Report_Temporal_Encoding.pdf`	Literature synthesis; motivation to broaden beyond binning.
`03_Advise/Advisory_Report_Temporal_Encoding.tex`	This document; turns synthesis into decisions and criteria.
`02_Analysis/Temporal_Separability_Benchmark-Study-Proposal.pdf`	Formal experiment: RQ, H1–H3, protocol sketch.
`02_Analysis/week_5_6.pdf`	Progress evidence: LDA intuition, TSB method refinements, sanity checks, within-model scope decision.