

# Scaling Synthetic Tabular Pretraining Data and Zero-Shot Classification Performance

Assignee Research

June 7, 2026

## Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How does increasing the volume of synthetic tabular pretraining data affect zero-shot classification accuracy on standard structured data benchmarks for tabular foundation models. 7 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Pretraining Strategies and Scaling for ECG Foundation Models: A Systematic Study. Research question: How does increasing the volume of synthetic tabular pretraining data affect zero-shot classification accuracy on standard structured data benchmarks for tabular foundation models?.

## 2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.2/10.

## 3 Results

13 papers retrieved. 7 claims extracted; 0 independently verified. Quality review score: 3.2/10.

## 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

## 5 Extracted Claims

Claim	Verified	Confidence
The study covers five different pretraining methodologies trained on over 11M samples.	×	0.15
State space models are confirmed as the superior architecture choice across all pretraining paradigms.	×	0.11
CPC shows the strongest and most transferable representations across diverse clinical tasks.	×	0.11
Data2vec consistently lags behind across all evaluation modes and scaling regimes.	×	0.03
Lower pretraining loss correlates with small residual errors in downstream tasks.	×	0.05
The S4 backbone with model dimension 512 consistently outperforms larger and alternative configurations.	×	0.01
The study investigates five self-supervised pretraining objectives spanning contrastive, predictive, and clustering-base	×	0.10

## References

- <http://arxiv.org/abs/2512.03307v1>
- <http://arxiv.org/abs/2402.01204v4>
- <http://arxiv.org/abs/2605.12241v1>