

# Scaling Pretraining Tasks in CodeT5 Enhances Alignment with Human Code Security Judgments

Assignee Research

June 8, 2026

## Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: Does increasing the scale of pretraining tasks in CodeT5 improve alignment with human expert judgments on novel code security flaws. 13 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.7/10. 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: Does increasing the scale of pretraining tasks in CodeT5 improve alignment with human expert judgments on novel code security flaws?.

## 2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.7/10.

## 3 Results

16 papers retrieved. 13 claims extracted; 1 independently verified. Quality review score: 4.7/10.

## 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 a corpus comprising over 11 million ECG samples.	×	0.09
State space models (SSM) are confirmed as the superior architecture choice across all pretraining paradigms compared to	×	0.09
CPC (Contrastive Predictive Coding) shows the strongest and most transferable representations across diverse clinical ta	✓	0.17
data2vec consistently lags behind other pretraining methodologies across all evaluation modes and scaling regimes.	×	0.08
Scaling behavior is most clearly identified for CPC and JEPA pretraining methodologies.	×	0.13
Lower pretraining loss correlates with small residual errors in downstream tasks.	×	0.07
Structured state space models have shown superior performance on long sequences in supervised ECG settings in prior work	×	0.14
The S4 backbone with model dimension 512 consistently outperforms larger configurations (768, 1024) and alternative back	×	0.02
All models in the study operate at 240 Hz on 12-lead ECG inputs.	×	0.04
The default backbone adopted for the study is a 4-layer S4 with dimension 512.	×	0.00
The CNN stem used in all models consists of four convolutional layers with batch normalization.	×	0.03
The Transformer backbone variant uses RoPE positional encoding and GELU activations.	×	0.03
The study evaluates five self-supervised pretraining objectives spanning contrastive, predictive, and clustering-based p	×	0.10

## References

- <http://arxiv.org/abs/2411.15497v3>
- <http://arxiv.org/abs/2403.13430v2>
- <http://arxiv.org/abs/2605.12241v1>