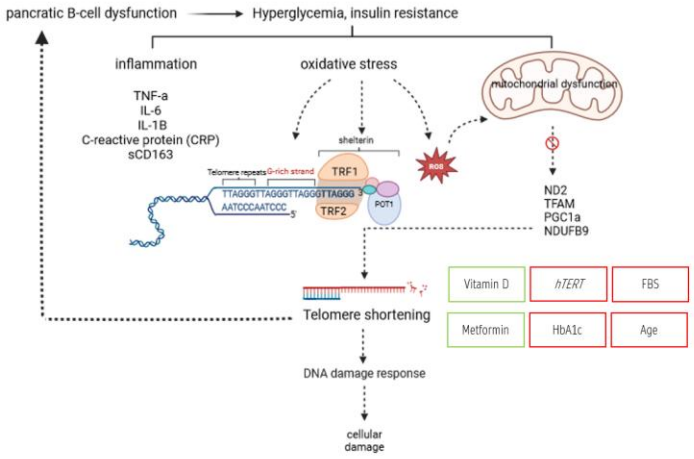
Study design: This cross-sectional analytical study was conducted on 166 patients with type 2 diabetes, who were classified into four groups based on the severity of diabetic retinopathy, including NDR, NPDR, PDR, and Reg PDR. The study aimed to identify distinct patterns of relative telomere length and *hTERT* gene expression across different DR severity stages to understand their potential prognostic capability. Results demonstrated a significant association between telomere length, *hTERT* gene expression, and DR severity, suggesting their potential as predictive biomarkers. Based on these promising results, we further developed a machine learning algorithm (using Supervised training and the gradient boosting classifier algorithm) for staging patients using their clinical, biochemical, and molecular data to classify DR stages.

Key Findings: This thesis demonstrates that telomere length and catalytic core of Telomerase (*hTERT*) expression significantly differ across stages of diabetic retinopathy (DR) ($p<0.01$, $p<0.05$, respectively). By integrating molecular, biochemical, and demographic data, we were able to discriminate between DR stages and track disease progression with greater precision. Strong correlations were observed between telomere length and HbA1c ($0.013<p<0.017$), fasting blood sugar (FBS) ($0.022<p<0.028$), and vitamin D levels ($p=0.002$). Diabetes treatment strategies also influenced telomere length, with combination therapies (e.g., insulin + metformin) showing distinct profiles (longer length of telomere, $p=0.01$). After statistical adjustment and finding the most correlated features (using the Ordinary least squares regression method), we further integrated telomere length with biochemical and demographic variables to develop a predictive ML-based clinical model for effectively predicting disease stages in diabetic retinopathy. Overall, the predictive model achieved an accuracy of 70%, sensitivity of 78%, and specificity of 89% for DR severity classification (using supervised approach and Gradient Boosting Classifier algorithm).

Clinical Importance: Such stratification allows earlier identification of patients at risk of advancing to more severe stages and enables timely adaptation of therapeutic strategies. These findings not only deepen our understanding of the molecular pathogenesis of DR but also pave the way for the implementation of machine learning approaches in predicting disease stages.

Key methods: DNA/RNA extraction, Electrophoresis, Spectrophotometry, Primer design, PCR, cDNA synthesis, qPCR, Elisa, cell culture, Extensive Bioinformatics, Statistical Analysis, Machine learning predictive modelling

DRAWING CONNECTIONS



ACCURACY: 0/70

SENSITIVITY: 0/785

SPECIFICITY: 0/896

PPV: 0/833

NPV: 0/941