Comparison of Leucocyte Telomere Length, Atherosclerotic Cardiovascular disease risk and the output of a Deep learning Cardiac BioAge risk prediction tool: A UK Biobank Cohort study

Ehsan Vaghefi¹,³*, Songyang An¹,³, Rini Corbett³, David Squirrell²,³

1. Department of Optometry and Ophthalmology, University of Auckland, New Zealand
2. Department of Ophthalmology, Greenlane Clinical Centre, Auckland District Health Board, Auckland
3. Toku Eyes, Auckland, New Zealand

* Corresponding author

Associate Professor Ehsan Vaghefi

6 Clayton Street, New Market
Auckland
New Zealand, 1023

e.vaghefi@auckland.ac.nz

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract.

Significance.
That a retinal image based Deep Learning (DL) Cardiac BioAge Model may be a useful novel tool that can personalise an individual’s risk of Atherosclerotic cardiovascular disease (ASCVD) events.

Purpose.
To determine whether the results issued by our DL Cardiac BioAge model are consistent with the known trends of cardiovascular disease (CVD) risk and the biomarker Leucocyte Telomere Length, in a cohort of individuals from the UK Biobank.

Methods.
Individuals were divided by sex, ranked by Z adjusted log T/S Leucocyte Telomere length (LTL) and then grouped into deciles. The retinal images were then presented to the DL model and individual’s Cardiac BioAges determined. Individuals within each LTL decile was then ranked by Cardiac BioAge, and the mean of the CVD risk biomarkers in the top and bottom quartiles compared. The relationship between an individual’s Cardiac BioAge, the CVD biomarkers and LTL were determined using traditional correlation statistics.

Results.
The DL Cardiac BioAge model was able to accurately stratify individuals by the traditional CVD risk biomarkers, and for both males and females those issued with a Cardiac BioAge in the top quartile of their chronological peer group had a significantly higher mean SBP, HbA1C and 10-year Pooled Cohort Equation ASCVD scores compared to those individuals in the bottom quartile. Cardiac BioAge was associated with LTL shortening for both males and females. (Males: -0.220, *P* <0.001; Females: -0.174, *P* <0.001)

Conclusion
In this small cohort study increasing CVD risk; as assessed by both traditional biomarkers, ASCVD risk scoring and a DL Cardiac BioAge CVD risk model, was inversely related to LTL. At a population level our data supports the growing body of evidence that suggests that LTL shortening is a surrogate marker for increasing CVD risk and that this risk can be captured by our novel DL Cardiac BioAge model.
Key Words.
Deep learning, Cardiac BioAge, Leucocyte Telomere Length, Cardiovascular risk

Introduction:

Atherosclerotic Cardiovascular disease (ASCVD) is the commonest cause of hospitalization and premature death in the US. The risk of an individual experiencing an ASCVD event includes both non-modifiable variables; age, sex, and ethnicity, and modifiable variables such as diabetes, hypertension, hyperlipidemia, and smoking. Across a population the risk of experiencing an ASCVD event varies greatly, and risk-based equations have therefore been developed to identify those who are at greatest risk of an ASCVD event risk so that treatments can be instigated appropriate to the individual’s risk. The landmark Framingham Heart Study was the first to demonstrate that multivariable equations could identify an individual's ASCVD risk with far greater accuracy than the existing metrics based solely on blood pressure and cholesterol. Since the Framingham-based equations were first published, other equations have been developed designed to serve different populations with refined accuracy, but in all these equations chronological age remains one of the strongest predictors of ASCVD risk. However, it is increasingly recognised that biological age, the concept that individuals age at different rates, impacts both the risk of an individual developing a chronic disease and the severity to which they are impacted by it. That individuals age at different rates is not currently captured by traditional regression-based risk equations as they assume all individuals in the population age at the same rate. The traditional equations are therefore insensitive to the variation in individual risk that results from biological aging, and it has therefore been suggested that biological markers such as Leukocyte telomere length (LTL) and DNA methylation rates could be used to refine the prediction of an individual’s ASCVD risk.

Currently LTL testing is impractical for population screening as it is an expensive test which, as a biological specimen requires careful handling within specialised laboratories. There is then arguably a need to develop other novel techniques that can personalise the ASCD risk scores produced by the traditional ASCVD risk equations. The retina is unique in being the only part of the vascular system that is visible by non-invasive means. It is now recognised that Deep Learning (DL) can extract data from retinal images to augment the traditional
means of estimating ASCVD risk\textsuperscript{15,16}. We have developed an DL model (Cardiac BioAge) designed to identify those individuals within a tight chronological age band who are at increased risk of cardiac events. The aim of the study was to determine whether the results issued by our Cardiac BioAge model are consistent with the known trends of CVD risk; including the traditional biomarkers; hypertension, HbA1C, ASCVD risk and the novel biomarker LTL, in a cohort of individuals from the UK Biobank.

**Methods.**

Only those individuals in the UK Biobank who had LTL measurements, fundus images and relevant cardiovascular biomarker data; (Systolic, TChol/HDL, HbA1C, smoking status and diabetes status) were used in this study. The data from the UK Biobank can be accessed via a direct request to the UK Biobank (IRB UOA-862 99); and was obtained using approved data management and data transfer protocols. The research presented in this study and the datasets used in it, adhered to the principles outlined in the Declaration of Helsinki. The composition of the subset of the UK Biobank dataset used in this study is shown in Table 1. Participants in the UK Biobank were recruited from a UK general population with only approximately 5\% of the UK Biobank population self-identified as having diabetes “diagnosed by doctor. 33,370 individuals in the UK Biobank had LTL data, retinal images, and relevant biodata. The demographics and the distribution of the Z adjusted log T/S LTL’s of this individual is shown in Table 1 and Figure 1.

*Table 1: The demographic and risk factor makeup of the UK Biobank dataset used in this study.*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>56</td>
<td>8.3</td>
</tr>
<tr>
<td>Systolic Blood Pressure (mmHg)</td>
<td>134</td>
<td>18</td>
</tr>
<tr>
<td>Diastolic Blood Pressure (mmHg)</td>
<td>82</td>
<td>10.1</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>5.43</td>
<td>0.6</td>
</tr>
<tr>
<td>Total cholesterol(mg/dL)</td>
<td>220</td>
<td>43.7</td>
</tr>
<tr>
<td>HDL cholesterol (mg/dL)</td>
<td>57</td>
<td>15.1</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>BMI</td>
<td>27.2</td>
<td>4.7</td>
</tr>
<tr>
<td>Sex at birth</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>45%</td>
<td>55%</td>
</tr>
<tr>
<td>Current Smoker</td>
<td>TRUE</td>
<td>FALSE</td>
</tr>
<tr>
<td></td>
<td>11.9%</td>
<td>88.1%</td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>4.6%</td>
<td>95.4%</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>White</td>
<td>92.9%</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>2.8%</td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td>1.0%</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>3.3%</td>
</tr>
</tbody>
</table>

Figure 1 The distribution of the Z adjusted log T/S ratio LTL distribution for males and females in the UK Biobank.

Mean Z adjusted log T/S ratio LTL Males -0.105 (1.006) v Females 0.067 (0.992); P<0.001.

The Cardiac BioAge DL Model.
The cardiac BioAge model has been described previously\textsuperscript{17}. In brief, this model produces a predicted Pooled Cohort Equation (PCE) derived ASCVD risk score\textsuperscript{18} and then determines how far that individuals score deviates from that of their chronological peers. The cardiac BioAge score, in effect informs whether the individuals 10-year ASCVD risk is greater or less than that of their chronological peers, and is derived using the following methodology:
The degree of similarity of the individual’s DL model predicted ASCVD risk with the ASCVD risk score obtained from other similar aged people is determined.

The average (chronological) age of that individuals’ closest neighbours, based on the condition of the retina as expressed by multi-dimensional features extracted from the DL model, is determined.

The mean expected age for a person with the individual’s predicted ASCVD risk is determined.

These data are then combined in a probabilistic manner to give a final value for the cardiac BioAge.

Individuals were divided into males and females and each cohort was then ranked by Z adjusted log T/S LTL and grouped into deciles; shortest to longest. The retinal images were then presented to the DL model and individual’s Cardiac BioAge determined and individuals within each LTL decile were then ranked by Cardiac BioAge. Next, the mean of following variables; Systolic BP, HbA1C, PCE 10-year ASCVD risk score, derived from those individuals in the top quartile and the bottom quartile were compared and the relationship between these biomarkers and Z adjusted log T/S LTL determined using traditional correlation statistics.

Results.

The chronological age and the Cardiac BioAge issued by the DL model of males ranked by mean Z adjusted log T/S LTL decile are shown in Table 2.

Table 2. Chronological and Cardiac BioAge, categorised by mean Z adjusted log T/S LTL ratio deciles for males and females:

<table>
<thead>
<tr>
<th>Z adjusted log T/S LTL ratio Decile</th>
<th>Males</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Z adjusted log T/S LTL ratio*</td>
<td>Mean Chronological Age (yrs)</td>
</tr>
<tr>
<td>0</td>
<td>-0.4242</td>
<td>68.1</td>
</tr>
<tr>
<td>10</td>
<td>-0.3501</td>
<td>65.5</td>
</tr>
<tr>
<td>20</td>
<td>-0.2431</td>
<td>63.4</td>
</tr>
</tbody>
</table>
Correlation (Pearson) Males: Chronological age to LTL -0.231 (P <0.001); Cardiac BioAge to LTL -0.220 (P <0.001). Correlation (Pearson) Females: Chronological age to LTL -0.197 (P <0.001), Cardiac BioAge to LTL -0.174 (P <0.001)

The Systolic Blood pressure, HbA1C and 10-year PCE ASCVD risk score of those in individuals issued with a Cardiac BioAge score in the top quartile (Q1) and those issued with a Cardiac BioAge score in the bottom quartile of their chronological peer group, is shown in Table 3 and Figure 2. For both males and females, the mean systolic BP, mean HbA1C and the mean 10-year PCE ASCVD risk score, were all significantly higher in those individuals in the highest quartile of Cardiac BioAge compared to those in the lowest quartile of Cardiac BioAge (P <0.001) Table 3. LTL were significantly shorter in males compared to females (Figure 1) and in both males and females, a worsening biomarker profile, an increased CVD risk and increasing Cardiac BioAge were inversely correlated with LTL (Table 2, Figure2).

Table 3: Mean Biomarkers for individuals in the top quartile of Cardiac BioAge (Q1) compared to the bottom Quartile (Q4) for their chronological peer group.

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q1 Cardiac BioAge Mean (Sd)</td>
<td>Q4 Cardiac BioAge Mean (Sd)</td>
</tr>
<tr>
<td>Cardiac BioAge</td>
<td>57 (2.0)</td>
<td>54 (8.4)</td>
</tr>
<tr>
<td>SBP (mmHg)</td>
<td>140 (0.8)</td>
<td>135 (1.1)</td>
</tr>
<tr>
<td>HbA1c (mmol/mol)</td>
<td>37 (0.5)</td>
<td>35 (0.6)</td>
</tr>
<tr>
<td>10-year ASCVD PCE risk score</td>
<td>8% (1.0)</td>
<td>7% (1.2)</td>
</tr>
</tbody>
</table>
**Figure 2. Distribution of the mean biomarkers for male and female individuals in the top quartile of Cardiac BioAge (TopQ) compared to the bottom Quartile (BotQ) for their chronological peer group, categorised by Z adjusted log T/S LTL decile.**

Correlation between relevant biomarker and Z adjusted log T/S LTL: Males: SBP -0.045 (\(P < 0.001\)), HbA1c -0.049 (\(P < 0.001\)), 10-year ASCVD PCE risk score -0.137 (\(P < 0.001\)). Females: SBP-0.0619 (\(P < 0.001\)), HbA1c -0.046 (\(P < 0.001\)), 10-year ASCVD PCE risk score -0.117 (\(P < 0.001\)).

**Discussion.**

In this cross-sectional study using a subset of individuals drawn from the UK Biobank, we found that our DL Cardiac BioAge model was able to accurately stratify individuals by the traditional CVD risk biomarkers, and that for both males and females that an elevated Cardiac BioAge score was associated with LTL shortening. In line with traditional ASCVD models, both males and females who were issued with a Cardiac BioAge in the top quartile of their chronological peer group had a significantly higher mean SBP, HbA1C and 10-year PCE ASCVD scores compared to those individuals in the bottom quartile. Consistent with other published data, we observed that a worsening biomarker profile and an increased ASCVD risk was associated with LTL shortening for both males and females\(^{11,12,14,19,20}\).

The rate of telomere shortening throughout life is determined by both endogenous (genetic) and external (non-genetic) factors\(^{21,22}\). Although the exact mechanism behind LTL shortening and CVD remains to be elucidated, it has been hypothesised that LTL may reflect both an individual's cumulative inflammatory exposure and oxidative stress and their genetically determined capacity to repair vascular damage\(^{11}\). Multiple confounding risk factors and reverse causation means that any observed associations between LTL and CVD should be interpreted with caution, however studies using Mendelian randomisation provide compelling evidence for a relationship between LTL shortening and an increased risk of both atherosclerotic cardiovascular events and an increased risk of hypertension\(^{12,23,24}\). Irrespective of the mechanism by which LTL shorten, our findings indicate that whether assessed by traditional Biomarkers, or LTL, our novel Cardiac BioAge DL model could be used as a novel surrogate biomarker for cardiac Health.
A critical component to the provision of Public Health is accurate risk prediction as it enables focused preventative action and tailored disease management strategies. Accurate disease prediction also facilitates “Precision medicine” the identification of subgroups within a wider population who have different disease risks, prognosis, and responses to treatment due to differences in their habitat, physiology, and other characteristics\textsuperscript{25}. The utilisation of novel biomarkers, like LTL testing, in tandem with tradition measures of ASCVD risk such as the PCE equation, is one approach to the provision of more “precise” interventions to reduce the risk of ASCVD events. Although such approaches have merit, the cost, and practicalities of providing LTL testing remains a significant barrier to its widespread adoption. Arguably one of the more important contributions that AI will play within the health care ecosystem is to improve the effectiveness and precision of the traditional population-based risk predictions\textsuperscript{26, 27}. Optometry has a unique and critical role as the community provider of quality public eye health care and through clinical examination and targeted investigation, Optometrists have been providing Precision eye health care for decades. Until now Optometry’s focus has been on ocular disease, but with the emergence of AI algorithms, such as our Cardiac BioAge DL model, which have been designed specifically to stratify an individual’s risk of systemic disease by examining retinal images, there is now the potential for Optometrists to extend the provision of precision medicine beyond their traditional scope of practice. As retinal photographs are routinely captured in Optometry practices these algorithms can be deployed without significant additional investment in primary care, a feature which makes these technologies particularly relevant to low-resource settings. As such these tools have the potential to not only make ASCVD risk assessment more affordable and accessible but may also improve patient outcomes by both facilitating early intervention and increasing the public awareness of this issue.

The principal limitation of this study is that it is a cross sectional study of individuals in the UK Biobank whose LTL, retinal images and traditional biomarkers were measured at one point of time. As such, although our DL model can categorise individuals into low and high-risk groups for CVD based on the traditional biomarkers, and we observed the expected relationship between worsening biomarkers, increasing Cardiac BioAge and LTL shortening, it is not possible to determine whether these associations are causative. Further studies, using longitudinal data are therefore required to elucidate the relationship between LTL, aging, and “disease”. As is the case in all studies of DL algorithms, further studies of the algorithm in other datasets are required to demonstrate whether the results are generalisable to other
datasets and other races and ethnicities beyond those found in the UK Biobank. Finally, although LTL may be inversely associated with CVD risk in large population-based studies, the validity of this metric as a clinical tool to predict an individual's CVD risk and its magnitude remains to be proven²⁸.

**Conclusion.**

In this small cohort study conducted on individuals in the UK Biobank who had relevant Biomarker data, retinal images, and LTL data, increasing CVD risk; as assessed by both traditional biomarkers, ASCVD risk scoring and a DL Cardiac BioAge CVD risk model, was inversely related to LTL. At a population level our data supports the growing body of evidence that suggests that LTL shortening is a surrogate marker for increasing CVD risk and that this risk can be captured by suitably trained DL models. If these results can be replicated in other large population-based datasets, DL models like ours may offer the potential to significantly improve access to ASCVD risk detection strategies as the risk predictions these models produce do not require multiple clinical and laboratory assessments to generate an individual’s ASCVD risk score.

**Disclosures**

Ehsan Vaghefi and David Squirrell are the co-founders of Toku Eyes. Songyang An is a member of the AI engineering team at Toku Eyes. Rini Corbett: Employment- Toku Eyes

**Sources of funding**

Toku Eyes has funded this study.

**References:**


20. Schneider CV, Schneider KM, Teumer A et al. Association of telomere length with risk of


