Check for updates

RESEARCH ARTICLE

REVISED Cardiovascular risk prediction in India: Comparison of the original and recalibrated Framingham prognostic models in urban populations. [version 2; peer review: 2 approved]

Priti Gupta[1], David Prieto-Merino [ID][2,3], Vamadevan S. Ajay [ID][1], Kalpana Singh [ID][1], Ambuj Roy[4], Anand Krishnan [ID][4], K.M. Venkat Narayan[5], Mohammed K. Ali [ID][5], Nikhil Tandon[4], Dorairaj Prabhakaran[1,6], Pablo Perel [ID][2]

[1]Centre for Chronic Disease Control, New Delhi, 110016, India
[2]Centre for Global Chronic Conditions, London School of Hygiene & Tropical Medicine, London, UK
[3]Applied Statistical Methods in Medical Research Group, Universidad Catolica San Antonio de Murcia (UCAM), Murcia, Spain
[4]All India Institute of Medical Sciences, New Delhi, 110029, India
[5]Rollins School of Public Health, Atlanta, USA
[6]Public Health Foundation of India, Gurgaon, Haryana, India

Abstract

Introduction: Cardiovascular diseases (CVDs) are the leading cause of death in India. The CVD risk approach is a cost-effective way to identify those at high risk, especially in a low resource setting. As there is no validated prognostic model for an Indian urban population, we have re-calibrated the original Framingham model using data from two urban Indian studies.

Methods: We have estimated three risk score equations using three different models. The first model was based on Framingham original model; the second and third are the recalibrated models using risk factor prevalence from CARRS (Centre for cArdiometabolic Risk Reduction in South-Asia) and ICMR (Indian Council of Medical Research) studies, and estimated survival from WHO 2012 data for India. We applied these three risk scores to the CARRS and ICMR participants and estimated the proportion of those at high-risk (>30% 10 years CVD risk) who would be eligible to receive preventive treatment such as statins.

Results: In the CARRS study, the proportion of men with 10 years CVD risk > 30% (and therefore eligible for statin treatment) was 13.3%, 21%, and 13.6% using Framingham, CARRS and ICMR risk models, respectively. The corresponding proportions of women were 3.5%, 16.4%, and 11.6%. In the ICMR study the corresponding proportions of men were 16.3%, 24.2%, and 16.5% and for women, these were 5.6%, 20.5%, and 15.3%.

Conclusion: Although the recalibrated model based on local

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 02 Dec 2019	✓ view	✓ view
version 1 17 Apr 2019	? view	? view

1. **Kay Tee Khaw** [ID], University of Cambridge, Cambridge, UK

2. **Katrina K. Poppe** [ID], University of Auckland, Auckland, New Zealand

Any reports and responses or comments on the article can be found at the end of the article.

population can improve the validity of CVD risk scores our study exemplifies the variation between recalibrated models using different data from the same country. Considering the growing burden of cardiovascular diseases in India, and the impact that the risk approach has on influencing cardiovascular prevention treatment, such as statins, it is essential to develop high quality and well powered local cohorts (with outcome data) to develop local prognostic models.

Keywords
CVD risk score prognostic model, the re-calibrated Framingham risk score

---

Corresponding author: Pablo Perel (Pablo.Perel@lshtm.ac.uk)

Introduction

Currently, cardiovascular diseases (CVDs) account for two-thirds of the total non-communicable disease (NCD) burden in India[1]. According to the 2016 Global Burden of Disease study, ischemic heart disease was the leading cause of the Disability-adjusted life years (DALYs), measured to be 3062 per 100,000 population in India[2]. Also, the all-age death rate increased significantly between 1990 and 2016 for ischaemic heart disease (percentage change 54·5%), and CVDs are the leading cause of death in most parts of India[2,3]. Age adjusted prevalence of CVDs have also increased in India[4,5]. Indians are affected by CVDs at a younger age compared to their European counterparts, with more than 50% CVDs deaths occurring before the age of 70[6–8]. The World Health Organization (WHO) had estimated that, due to the burden of CVDs, India had lost 237 billion dollars over ten years (2005–2015)[9].

CVDs risk approach is a cost-effective means to identify those at high risk so that immediate short and long-term preventive steps can be followed to mitigate the risk[10]. Risk stratification approach has been primarily found to be cost-effective in resource-poor settings[11].

Although risk factor effect can be similar across populations, the estimated cardiovascular disease risk from risk models differs substantially across populations. This is mainly because of the different "baseline incidence of the risk model outcome" and prevalence of the different risk factors across populations. Also, a meta-analysis based on 17 population-based cohorts worldwide has shown that ethnicity modifies the association between risk factors and cardiovascular disease[12]. Another study from the United Kingdom has shown the CVD risk prediction model to be inaccurate in the South Asian group as compared to white Europeans[13]. Studies have also proved that the Framingham risk prediction model underestimates the CVD risk in Asian Indians and socioeconomically deprived individuals[14,15].

The Framingham risk equation is a well-established and widely used method to measure cardiovascular disease (CVD) risk, but was developed with a white US-based population several decades ago and so, there is a need to re-calibrate it when applying it in other populations. Recalibrating a risk equation to

a new population involves estimating the average values of the risk factors and the average risk of CVD. These values are used as the reference values in the risk model equations. To the best of our knowledge, only one study has recalibrated the Framingham risk equation in India, and this was for a rural population[16]. As there is no validated prognostic model for an Indian urban population, we have re-calibrated the original Framingham model. In this paper, we report the Framingham model recalibration to an Indian urban population using data from two studies: CARRS (Centre for cArdiometabolic Risk Reduction in South-Asia), and ICMR (Indian Council of Medical Research). We compare the 10-year predictions of CVD fatal and non-fatal events produced by the original Framingham model and the recalibrated models and describe the potential impact of the recalibration on the proportion of the population eligible for treatment as recommended by current WHO guidelines.

Methods

Data sources

We used data from two studies:

a) The CARRS Cohort study was a population-based sample of urban adults in Chennai, New Delhi and Karachi established to assess the prevalence and incidence of cardio-metabolic diseases and their risk factors. Its details have been published previously[17], in brief, participants were selected in each city using multi-stage cluster random sampling with the Kish method[18] to select only one man and one woman aged 20+ from each randomly selected household. Here we used baseline data from the cross-sectional survey conducted between October 2010, and December 2011 with mortality follow up through June 2014.

b) The ICMR study was a cross-sectional survey conducted to estimate CVD risk factor prevalence in the National Capital Region of India (Delhi and Ballabgarh) in 2010–2012[19]. Multi-stage cluster random sampling was used for the primary sampling unit (household) selection. Data were collected on sociodemographic characteristics, CVD risk factors, treatment status, and measurements of height, weight, hip and waist circumference, and blood pressure. Fasting blood glucose (FBG) and lipids measurements were done using fasting venous blood. Here we have included data only from the urban area of Delhi.

Risk calculation

Below we described the different steps we conducted to recalibrate the Framingham score

Step 1: Prognostic index or linear predictor calculation (i.e. Xi): We first calculated each individual Framingham "score" for CVD events in the next 10 years. This is a weighted sum of the individual's characteristics using the Framingham weights (see Table 1).

Step 2: Reference individual survival calculation (S_0): We then obtained the 2012 yearly mortality rates from CVD causes for India from WHO (0.003489 for men and 0.002646 for women). We assumed that the ratio of non-fatal to fatal events was 2:1 so the yearly rates of total fatal and non-fatal

events was estimated as 3*0.003489 = 0.010467 for men and 3*0.002646 = 0.007938 for women. We assumed a constant ratio, and therefore the probabilities of not having events in 10-years were $(1 − 0.010467)^{10} = 0.90013$ for men and $(1 − 0.007938)^{10} = 0.923396$ for women.

Step 3: Reference individual score calculation (X_0): We then calculated three risks for each individual considering different values of scores and survivals: M1) using Framingham's reference score and survival, M2) using a reference score (X_0) derived from CARRS and estimated survival (S_0) derived from WHO 2012 data for India (see above), and M3) using reference score (X_0) derived from ICMR and the estimated survival (S_0) derived from WHO 2012 data for India. We estimated the score of the "average individual" in the population by multiplying the averages of the variables (in the log scale for the continuous variables) by the original Framingham coefficients and adding the values for all risk factors.

Step 4: Estimation of risks with different models: We calculated the risks for everyone in the CARRS and the ICMR datasets with three models (M1, M2, M3) using different combinations of reference score (X_0) and survival probabilities (S_0).

Comparison of risk and treatment

With each of the three risk calculations (M1. M2 and M3), we have stratified individuals in three different risk categories (<10%, 10–30%, and >30) which are commonly used for treatment recommendations for antihypertensive and statins. To see how recalibration with one or another data set affects the proportion of individuals treated, we have compared the proportion of individuals in the third risk category (>30%) between the different models. We reported the study following the TRIPOD statement[20]. A completed TRIPOD statement is available from OSF[21]. We used statistical package R, version 3.5.1 (2018-07-02)[1] for all our analysis

Results
Population characteristics
The CARRS study had data from 16,287 participants, but only 11,407 of those had the data needed to calculate the Framingham risk score (5,151 men and 6,256 women). The ICMR study had 3,075 individuals, but only 2,401 had all the data needed to calculate the score (1,089 men and 1,312 women).

In Table 2 we show the summaries statistics for each of the variables used in the Framingham score calculated in all the individuals that provided data for each variable separately.

In Table 3 below we report the reference scores using the Framingham, CARRS and ICMR populations, for this we have used the means of the log of the variables (which is not the same as the log of the mean). For example, for age, we first calculated a new variable "log(age)" for every single individual. Then we calculated the mean of this mean Mean[log(age)] =3.72033. Table 4 shows an example of the calculations using the CARRS population means.

Figure 1 shows the distribution of the recalibrated Framingham scores by sample and sex. Women have on average higher scores than men, and the Framingham ICMR recalibrated score has slightly higher means than the Framingham CARRS recalibrated in each sex.

Finally, in Table 5 we summarize the risk of the participants in both, CARRS and ICMR, estimated with the three different models for each sex/cohort. We present the mean risk and the distribution of the individuals in the three risk categories stated above (0–10%, 10–30%, and > 30%).

Table 1. Coefficients from simplified Framingham model[22].

	MEN	WOMEN
Log of Age	3.11296	2.72107
Log of Body Mass Index	0.79277	0.51125
Log of SBP* if not treated	1.85508	2.81291
Log of SBP if treated	1.92672	2.88267
Smoking (0=No / 1=Yes)	0.70953	0.61868
Diabetes (0=No / 1=Yes)	0.53160	0.77763

* SBP: systolic blood pressure.

Table 2. Descriptive of risk factors in the two data sets *: mean (standard deviation).

	MEN			WOMEN		
Data	CARRS	ICMR	All	CARRS	ICMR	All
Age*	42.9 (13.7)	47.2 (13.0)	43.6 (13.7)	41.1 (12.5)	45.2 (13.0)	41.8 (12.7)
BMI*	24.2 (4.6)	24.7 (6.1)	24.3 (4.9)	26.5 (5.5)	26.5 (5.5)	26.5 (5.5)
SBP* untreated	124.9 (17.7)	129.9 (19.4)	125.7 (18.0)	116.5 (17.9)	122.3 (19.6)	117.5 (18.4)
SBP* treated	139.7 (21.1)	143.9 (23.8)	140.7 (21.8)	136.5 (23.8)	143.6 (24.5)	138.1 (24.1)
Treatment	8.6%	13.8%	9.4%	13.6%	19.0%	14.5%
Smokers	27.1%	28.8%	27.3%	1.6%	3.3%	1.9%
Diabetics	26.0%	12.5%	23.5%	25.5%	10.7%	22.6%

BMI: body mass index, SBP: systolic blood pressure ***NOTE: We have removed patients with prevalent CVD from the CARRS and ICMR datasets***

Table 3. Reference scores using Framingham, CARRS and ICMR populations and reference survival times from Framingham and WHO.

Data	Male	Female
Average scores (X_0)		
Framingham	23.93880	26.01450
CARRS	23.36817	25.31639
ICMR	23.76358	25.66680
Reference Survival (S_0)		
Framingham	0.88431	0.94833
WHO 2012	0.90013	0.92340

CARRS: Centre for cArdiometabolic Risk Reduction in South-Asia; ICMR: Indian Council of Medical Research; WHO: World health organization

Table 4. Example of reference score calculations using means from CARRS population.

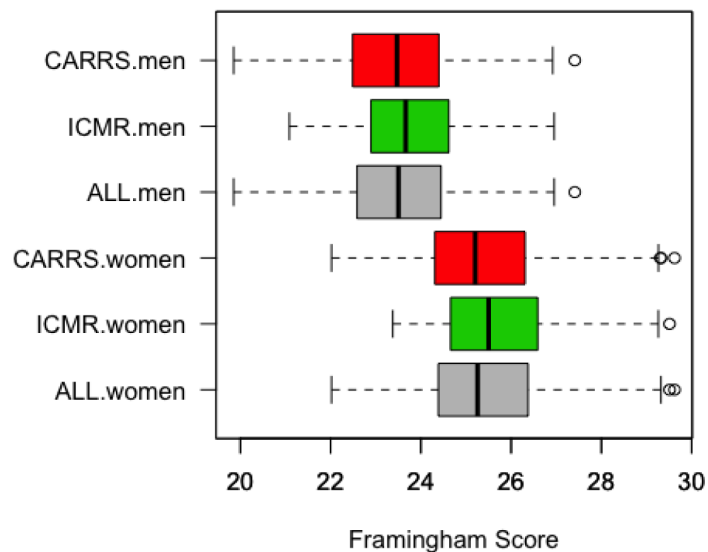|	MEN	WOMEN
	Coefficient * mean	Coefficient * Mean
Log of Age	3.11296*3.72033	2.72107*3.67602
Log of Body Mass Index	0.79277*3.17187	0.51125*3.25623
Log of SBP if untreated	1.85508*4.81871*0.8975	2.81291*4.74785*0.8553
Log of SBP if treated	1.92672*4.92633*0.1025	2.88267*4.90344*0.1447
Smoking	0.70953*0.26817	0.61868*0.01607
Diabetes	0.53160*0.27187	0.77763*0.26391
Sum	**SM = 23.42494**	**SW = 25.35218**

Figure 1. Box-plots of the distribution of the Framingham scores by sex and study. CARRS: Centre for cArdiometabolic Risk Reduction in South-Asia; ICMR: Indian Council of Medical Research.

Table 5. Means of estimated risks and distribution by risk categories of individuals in each cohort with different estimation models: M-1) Framingham M-2) F-CARRS recalibrated and M-3) F- ICMR recalibrated. F: Framingham; CARRS: Centre for cArdiometabolic Risk Reduction in South-Asia; ICMR: Indian Council of Medical Research.

Data/model	MEN				WOMEN			
	Mean	<=10%	10–30%	>30%	Mean	<=10%	10–30%	>30%
CARRS participants:								
M1) Framingham	0.131	58.7%	29.2%	**12.1%**	0.060	82.0%	14.9%	**3.1%**
M2) F-CARRS	0.182	47.2%	32.2%	**20.6%**	0.149	60.2%	23.6%	**16.1%**
M3) F-ICMR	0.133	58.2%	29.4%	**12.4%**	0.114	67.6%	21.5%	**11.0%**
ICMR participants:								
M1) Framingham	0.155	53.4%	30.7%	**16.0%**	0.079	76.5%	17.9%	**5.6%**
M2) F-CARRS	0.213	41.6%	32.6%	**25.8%**	0.188	51.7%	27.4%	**20.9%**
M3) F-ICMR	0.157	53.0%	30.5%	**16.5%**	0.145	60.6%	24.1%	**15.3%**

Effect of the recalibrated prognostic model on treatment

According to the WHO guidelines individuals with a risk of fatal or non-fatal cardiovascular event > 30% should get treatment with statins[23]. In Figure 2 below we plot the difference in the proportion of individuals (by sex and cohort) that should be eligible for treatment according to the different models. For example, as shown in the red bar to the top-left graph of Figure 2, if we used model M2 instead of M1, about 8% more men in CARRS will be eligible for treatment. This can be calculated from the difference of 21.0% – 13.3% in Table 5 (first two rows for men). The pattern is very similar in both datasets within each sex. In men model M2 increases the proportion of men that should be treated compared to both model M1 and model M3, and M1 and M3 categorize men very similarly. For women, models M2 and M3 also increase the proportion of women that should be treated in comparison to M1, but in addition model, M2 also categorize more women than M3 as eligible to be treated.

Discussion

In this study, we calculated the 10-year Framingham CVD score in two cohorts of Indian urban populations (CARRS and ICMR) by using coefficients from a simplified Framingham model. We then predicted the risk in each cohort using three different models: one with the original Framingham reference coefficients and two recalibrated with the average risk factors prevalence in each of the datasets and the CVD mortality estimations for India from WHO-2012 data.

The average 10-year CVD risk estimates calculated using the Framingham recalibrated equation with the CARRS data was substantially higher than the original Framingham equation for both men and women, but the recalibrated equation using the ICMR the averages were only distinctly higher for women but not for men. A previous study in rural India, also found that the Framingham score underestimates in comparison with the one recalibrated with national data[16]. Other studies in South

Asian Indian populations have also shown higher CVD incidence in comparison to the predicted by Framingham risk score[24–27].

The overall CVD risk score is used to inform clinical decisions to start treatment to lower blood pressure and statins. The thresholds recommended vary according to the guideline and settings. For example, the WHO guidelines recommend that individuals with a 30% 10-year risk of a fatal or non-fatal cardiovascular event should start with statins. Our study exemplifies that by using the prognostic model recalibrated with the CARRS data, there will be a substantial increase in the proportion of men and women that would be eligible for treatment with statins in comparison to the original Framingham risk score. However, by using the prognostic model recalibrated with the ICMR data, there would only be a substantial increase in the proportion of women that would be eligible for treatment.

To the best of the authors' knowledge, this is the first large community-based study to recalibrate Framingham risk score in an urban population in India. One of our strengths is that the data are representative of their respective cities and that we used two different cohort studies. The main limitation is that we cannot check if the re-classification of the recalibrated model is indeed an improvement in risk prediction comparison with the original Framingham score because of the lack of cardiovascular events in the existing cohorts. Also, for the recalibration we used WHO mortality data which includes a population with and without previous CVD, which might have overestimated the risk, however, until large and robust cohorts ,with detailed outcome assessment, become available this is the best available dataset that can be used to estimate the expected mortality in patients without CVD.

Early identification and initiation of intensive primary prevention among individuals with high risk of CVDs are critically important in reducing the CVD burden in India. Although,
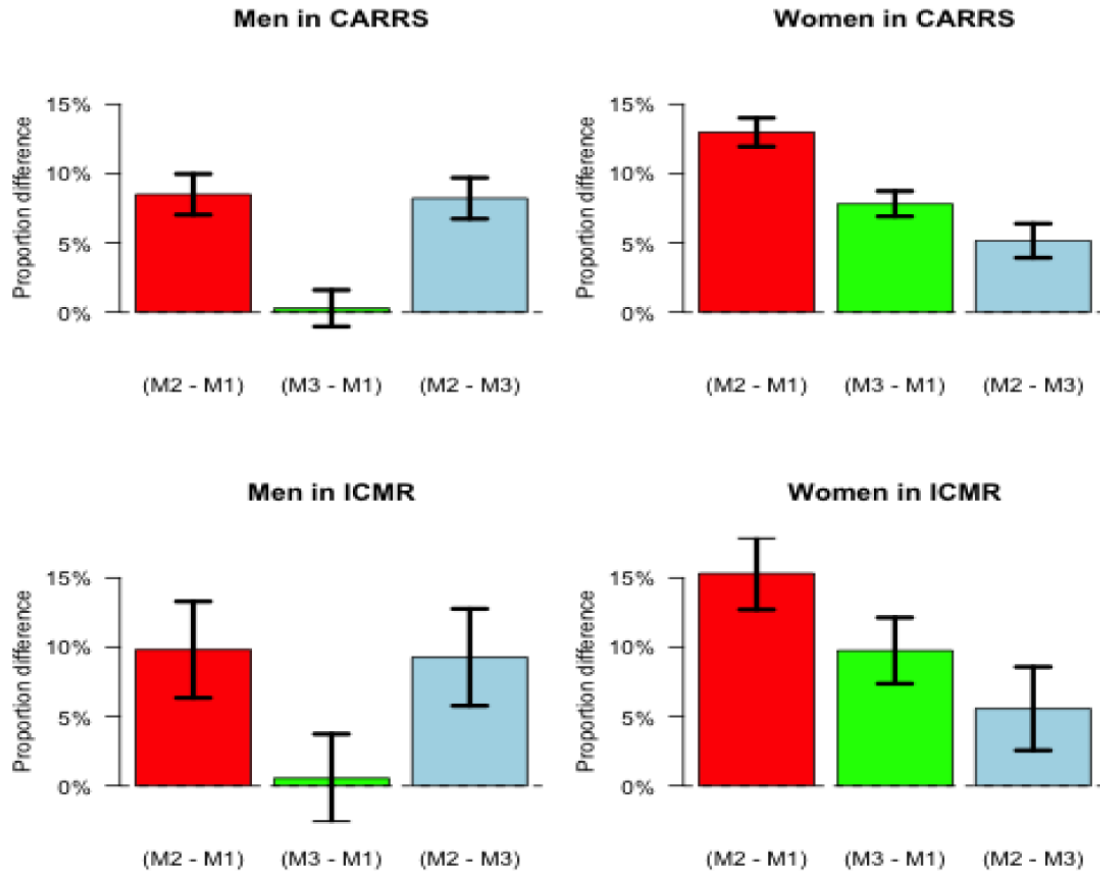
Figure 2. Difference in the proportion of treated individuals (risk > 30%) between the estimation of the three methods (M2-M1, M3-M1, and M2-M3). Bars reach the difference in proportions and segments represent 95% confidence intervals for the difference. M-1) Framingham M-2) F-CARRS recalibrated and M-3) F- ICMR recalibrated. F: Framingham; CARRS: Centre for cArdiometabolic Risk Reduction in South-Asia; ICMR: Indian Council of Medical Research

almost all the major recent international guidelines including the National Institute for Health and Care Excellence (NICE) 2014 guidelines, World Health Organization (WHO) 2007 guidelines, European Society of Cardiology (ESC) 2016 guidelines and the 2017 American College of Cardiology (ACC)/American Heart Association (AHA) guidelines and national guideline unanimously recommend assessment of cardiovascular risk[18,28–31], their adoption in primary prevention is suboptimal[31–33]. Few of common barriers for its decreased use are; lack of national guidelines, too many choices for CVD risk score, the uncertainty of validity of these risk score model in local context, time-consuming and lack of adjustment for the treatment[34,35]. Recalibrated model based on local population can improve the validity of the risk score model and reduce the perceived barriers of physician related to the local validity and enhance the use of CVD prediction model in the clinical setting for primary prevention. However, our study shows that even recalibrated models using data from the same country could be indeed very different and therefore it is vital to recalibrate models applying relevant local data (reflecting as best as possible local prevalence and overall mortality data). With the increasing use of technology, a possible approach could be to develop risk calculators in which local prevalent data and local incidence data is easily uploaded, and a "tailored" recalibrated model is provided for each setting. However, in the long-term future studies should develop CVD prognostic models using high quality and well powered local cohorts (with outcome data) and evaluate their implementation and impact.

Valid and reliable local prognostic models will also be key to evaluate different high risk prevention strategies.

Ethics and consent
CARRS (Centre for cArdiometabolic Risk Reduction in South-Asia) study was approved by the Institutional Review Boards (IRBs) of the Public Health Foundation of India, New Delhi (approval number: IRB00006658), All India Institute of Medical Sciences, New Delhi, Madras Diabetes Research Foundation, Chennai, India, Aga Khan University, Karachi, Pakistan, and Emory University, Atlanta, USA. ICMR (Indian Council of Medical Research) study was approved by All India Institute of Medical Sciences, New Delhi, ethics approval number is IEC/OP-05/06/09/10.

Data availability

Underlying data

This paper uses datasets from other studies. We gained access to these datasets by requesting permission from the PIs of the concerned studies. Therefore, data sets cannot be shared publicly.

Researchers who are interested can take permission from Dr. Prabhakaran who is the PI for both the studies (email id: dprabhakaran@ccdcindia.org) to access the data.

Extended data: We have added two supplementary figures:
- S1 Figure: This figure demonstrates contribution of risk factors to mean score. https://doi.org/10.6084/m9.figshare.11282900.v1

- S2 Figure: This figure shows predicted risk according to recalibration. https://doi.org/10.6084/m9.figshare.11282960.v1

Reporting guidelines

Open Science Framework: TRIPOD statement 'Cardiovascular risk prediction in India: Comparison of the original and recalibrated Framingham prognostic models in urban populations'. https://dx.doi.org/10.17605/OSF.IO/NXMZQ[21]

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

References

1. Patel V, Chatterji S, Chisholm D, *et al.*: **Chronic diseases and injuries in India.** *Lancet.* 2011; **377**(9763): 413–28.
   PubMed Abstract | **Publisher Full Text**

2. India State-Level Disease Burden Initiative Collaborators: **Nations within a nation: variations in epidemiological transition across the states of India, 1990-2016 in the Global Burden of Disease Study.** *Lancet.* 2017; **390**(10111): 2437–60.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. Office of the registrar general GOI: **Report on medical certification of cause of death 2009.** New Delhi; 2009 [cited 2017 Jul 17].
   Reference Source

4. Institute for Health Metrics and Evaluation (IHME): **GBD Compare Data Visualization.** Seattle, WA: IHME, University of Washington. 2017; [cited 2019 Nov 11].
   Reference Source

5. Prabhakaran D, Anand S, Gaziano TA, *et al.*: **Cardiovascular, Respiratory, and Related Disorders.** The International Bank for Reconstruction and Development / The World Bank; 2017; [cited 2019 Nov 11].
   PubMed Abstract

6. Joshi P, Islam S, Pais P, *et al.*: **Risk factors for early myocardial infarction in South Asians compared with individuals in other countries.** *JAMA.* 2007 [cited 2018 May 18]; **297**(3): 286–94.
   PubMed Abstract | **Publisher Full Text**

7. Xavier D, Pais P, Devereaux PJ, *et al.*: **Treatment and outcomes of acute coronary syndromes in India (CREATE): a prospective analysis of registry data.** *Lancet.* 2008; **371**(9622): 1435–42.
   PubMed Abstract | **Publisher Full Text**

8. Leeder S, Mph JP, Harikrishnan S, *et al.*: **A race against time: The challenge of cardiovascular disease in developing economies.** 2nd ed. New Delhi: Centre for Chronic Disease Control; 2014 [cited 2017 Jul 17].
   Reference Source

9. World health organization: **Global atlas on cardiovascular disease prevention and control.** WHO. Geneva: World Health Organization; 2011 [cited 2017 Jul 17].
   Reference Source

10. Lloyd-Jones DM: **Cardiovascular risk prediction: basic concepts, current status, and future directions.** *Circulation.* 2010; **121**(15): 1768–77.
    PubMed Abstract | **Publisher Full Text**

11. World Health Organization: **Integrated Management of Cardiovascular Risk.** Geneva; 2002 [cited 2017 Jul 17].
    Reference Source

12. Gijsberts CM, Groenewegen KA, Hoefer IE, *et al.*: **Race/Ethnic Differences in the Associations of the Framingham Risk Factors with Carotid IMT and Cardiovascular Events.** *PLoS One.* 2015; **10**(7): e0132321.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. Quirke TP, Gill PS, Mant JW, *et al.*: **The applicability of the Framingham coronary heart disease prediction function to black and minority ethnic groups in the UK.** *Heart.* 2003; **89**(7): 785–6.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

14. Kanjilal S, Rao VS, Mukherjee M, *et al.*: **Application of cardiovascular disease risk prediction models and the relevance of novel biomarkers to risk stratification in Asian Indians.** *Vasc Health Risk Manag.* 2008; **4**(1): 199–211.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

15. Brindle PM, McConnachie A, Upton MN, *et al.*: **The accuracy of the Framingham risk-score in different socioeconomic groups: a prospective study.** *Br J Gen Pract.* 2005; **55**(520): 838–45.
    PubMed Abstract | **Free Full Text**

16. Chow CK, Joshi R, Celermajer DS, *et al.*: **Recalibration of a Framingham risk equation for a rural population in India.** *J Epidemiol Community Health.* 2009; **63**(5): 379–85.
    PubMed Abstract | **Publisher Full Text**

17. Nair M, Ali MK, Ajay VS, *et al.*: **CARRS Surveillance study: design and methods to assess burdens from multiple perspective.** *BMC Public Health.* 2012; **12**(1): 701.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. World Health Organization: **Prevention of Cardiovascular Disease Guidelines for assessment and management of cardiovascular risk WHO Library Cataloguing-in-Publication Data.** Geneva; 2007 [cited 2017 Jul 17].
    Reference Source

19. Prabhakaran D, Roy A, Praveen PA, *et al.*: **20-Year Trend of CVD Risk Factors: Urban and Rural National Capital Region of India.** *Glob Heart.* 2017 [cited 2018 Aug 3]; **12**(3): 209–17.
    PubMed Abstract | **Publisher Full Text**

20. Collins GS, Reitsma JB, Altman DG, *et al.*: **Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement.** *Ann Intern Med.* 2015 [cited 2018 May 11]; **162**(1): 55–63.
    PubMed Abstract | **Publisher Full Text**

21. Gupta P: **Cardiovascular risk prediction in India: Comparison of the original and recalibrated Framingham prognostic models in urban populations.** 2019. http://www.doi.org/10.17605/OSF.IO/NXMZQ

22. D'Agostino RB Sr, Vasan RS, Pencina MJ, *et al.*: **General cardiovascular risk profile for use in primary care: the Framingham Heart Study.** *Circulation.* 2008; **117**(6): 743–53.
    PubMed Abstract | **Publisher Full Text**

23. World Health Organization: **HEARTS: Technical package for cardiovascular disease management in primary health care: evidence-based treatment protocols.** Geneva; 2018.
    Reference Source

24. Bhopal R, Fischbacher C, Vartiainen E, *et al.*: **Predicted and observed cardiovascular disease in South Asians: application of FINRISK, Framingham and SCORE models to Newcastle Heart Project data.** *J Public Health (Oxf).* 2005 [cited 2017 Jul 20]; **27**(1): 93–100.
    PubMed Abstract | **Publisher Full Text**

25. Tillin T, Hughes AD, Whincup P, *et al.*: **Ethnicity and prediction of cardiovascular disease: performance of QRISK2 and Framingham scores in a U.K. tri-ethnic prospective cohort study (SABRE--Southall And Brent REvisited).** *Heart.* 2014 [cited 2018 Jun 8]; **100**(1): 60–7.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

26. Cappuccio FP, Oakeshott P, Strazzullo P, *et al.*: **Application of Framingham risk estimates to ethnic minorities in United Kingdom and implications for primary

prevention of heart disease in general practice: cross sectional population based study. *BMJ.* 2002 [cited 2018 May 18]; **325**(7375): 1271.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

27. Cappuccio FP, Barbato A, Kerry SM: **Hypertension, diabetes and cardiovascular risk in ethnic minorities in the UK.** *Br J Diabetes Vasc Dis.* 2003 [cited 2018 May 18]; **3**(4): 286–93.
Publisher Full Text

28. National Institute for Health and Care Excellence: **Cardio Cardiovascular disease: risk assessment vascular disease: risk assessment and reduction, including lipid and reduction, including lipid modification.** Clinical guideline. 2014 [cited 2017 Jul 17].
Reference Source

29. Piepoli MF, Hoes AW, Agewall S, *et al.*: **2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts) Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR).** *Atherosclerosis.* 2016; **252**: 207–74.
PubMed Abstract | **Publisher Full Text**

30. Whelton PK, Carey RM, Aronow WS, *et al.*: **2017 ACC/AHA/AAPA/ABC/ACPM/AGS/ APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation,**

and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Hypertension.* 2018 [cited 2018 May 11]; **71**(6): e13–115.
PubMed Abstract | **Publisher Full Text**

31. Wardhan H, Nath RK: **National Programme for Prevention and Control of Diabetes, Cardiovascular Disease and Stroke Guidelines on Assessment and Management of Cardiovascular Risk for Medical Officers.** 2009 [cited 2018 May 11].
Reference Source

32. World health organization: **HEARTS: Technical package for cardiovascular disease management in primary health care.** 2016 [cited 2018 May 11].
Reference Source

33. Enas E, Dharmarajan T: **The Lipid Association of India Expert Consensus Statement 2016: A sea change for management of dyslipidemia in Indians.** *J Clin Prev Cardiol.* 2016 [cited 2018 May 11]; **5**(2): 62–66.
Publisher Full Text

34. George CE, Ramadas D, Norman G, *et al.*: **Barriers to cardiovascular disease risk reduction: Does physicians' perspective matter?** *Indian Heart J.* 2016 [cited 2018 Jun 8]; **68**(3): 278–85.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

35. Liew SM, Doust J, Glasziou P: **Cardiovascular risk scores do not account for the effect of treatment: a review.** *Heart.* 2011 [cited 2018 Jun 8]; **97**(9): 689–97.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Open Peer Review

Current Peer Review Status: ✓ ✓

Version 2

Reviewer Report 26 May 2020

https://doi.org/10.21956/wellcomeopenres.17104.r37239

✓   **Kay Tee Khaw** [iD]

Gonville and Caius College, University of Cambridge, Cambridge, UK

The authors have addressed and clarified the points raised in the review, in particular, the distinction between cardiovascular disease or coronary heart disease in the risk prediction algorithms, and provided explanations for the somewhat unexpected sex differences. Their main conclusion remains important and highly relevant, which is the need to have relevant local population data when making clinical or public health policy.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiologist.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 16 December 2019

https://doi.org/10.21956/wellcomeopenres.17104.r37238

✓   **Katrina K. Poppe** [iD]

Department of Medicine, University of Auckland, Auckland, New Zealand

Thank you for considering my comment about people in the cohorts with CVD.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

- -

?   **Katrina K. Poppe** (iD)

¹ Department of Medicine, University of Auckland, Auckland, New Zealand
² Department of Medicine, University of Auckland, Auckland, New Zealand

Overall
This paper has a clear and valid rationale, and clearly describes the approach used to recalibrate the available risk score to the two Indian cohorts. The discussion is clear and I agree with the main conclusion that "it is essential to develop high quality and well powered local cohorts (with outcome data) to develop local prognostic models".

Main comments
The authors note that there is "no validated prognostic model for an Indian urban population" and so focus on recalibrating a Framingham equation developed in a completely different population. In 2018, the VIEW/PREDICT group in New Zealand published sex-specific risk scores for the primary prevention of CVD (Pylypchuk *et al.* Lancet 2018)[1]. As the scores were derived from a multi-ethnic cohort, they included a categorical variable for ethnicity, allowing risk to vary for the five largest defined ethnic groups in the country, one of which was Indian (derived from 14,188 Indian women and 20,232 Indian men). Although the meaning of urban and rural will differ by country, the PREDICT cohort will be predominantly urban. I would suggest that the PREDICT risk scores are a more appropriate starting point for assessing risk in an Indian population.

The CARRS and ICMR cohorts are population studies derived from sampling individuals per household, thus each cohort will include a mixture of people with and without CVD. I could not see anywhere that people with CVD were excluded, so my following comments assume they have remained in the cohorts.
  ○ The Framingham equation being referred to in this study is for people without CVD, and from the risk factors included, seems to be the 2008 score by D'Agostino *et al*.[2] (the equation is not referenced in the current manuscript). What proportion of each cohort already had CVD and what was the composition of the disease? E.g., proportions with stroke vs MI vs peripheral vascular disease vs… etc. Could a difference in that composition have led to the difference in risk estimates using recalibrated risk scores in the two cohorts?

○ The reference survival (So) was based on WHO estimated country-level mortality rates. This would be all-cause mortality, yet I believe the fatal component of the outcome predicted in the 2008 Framingham score is "coronary death". It is fair enough to then estimate the non-fatal events from a ratio, but it will multiplicatively amplify the non-fatal CVD event rate if based on all-cause deaths.

○ The country-level mortality from WHO will include people from rural and urban areas, and people with and without existing CVD. Either of these factors may reduce its relevance to the cohorts being studied, and/or the risk score being investigated.

A wording suggestion - the "Framingham score calculation (ie. Xi)" introduced on page 3 is more standardly called the prognostic index or linear predictor.

As the authors note, a key limitation of the available data is that the predicted risks cannot be compared to actual event rates in the cohorts, so there is no way of knowing which version of the score is accurate, and if recalibration is even needed. This further supports the need for development of high quality local cohorts with outcome data.

References
1. Pylypchuk R, Wells S, Kerr A, Poppe K, et al.: Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study. *The Lancet*. 2018; **391** (10133): 1897-1907 Publisher Full Text
2. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, et al.: General cardiovascular risk profile for use in primary care: the Framingham Heart Study.*Circulation*. 2008; **117** (6): 743-53 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have

significant reservations, as outlined above.

Author Response 25 Nov 2019

Priti Gupta

We thank the opportunity to respond to the reviewer very positive and helpful comments, which we have now responded point-by-point and, by doing so, we believe it has improved our manuscript

Overall
This paper has a clear and valid rationale, and clearly describes the approach used to recalibrate the available risk score to the two Indian cohorts. The discussion is clear and I agree with the main conclusion that "it is essential to develop high quality and well powered local cohorts (with outcome data) to develop local prognostic models".

Main comments
The authors note that there is "no validated prognostic model for an Indian urban population" and so focus on recalibrating a Framingham equation developed in a completely different population. In 2018, the VIEW/PREDICT group in New Zealand published sex-specific risk scores for the primary prevention of CVD (Pylypchuk *et al.* Lancet 2018)[1]. As the scores were derived from a multi-ethnic cohort, they included a categorical variable for ethnicity, allowing risk to vary for the five largest defined ethnic groups in the country, one of which was Indian (derived from 14,188 Indian women and 20,232 Indian men). Although the meaning of urban and rural will differ by country, the PREDICT cohort will be predominantly urban. I would suggest that the PREDICT risk scores are a more appropriate starting point for assessing risk in an Indian population.

The CARRS and ICMR cohorts are population studies derived from sampling individuals per household, thus each cohort will include a mixture of people with and without CVD. I could not see anywhere that people with CVD were excluded, so my following comments assume they have remained in the cohorts.
  ○ The Framingham equation being referred to in this study is for people without CVD, and from the risk factors included, seems to be the 2008 score by D'Agostino *et al.*[2] (the equation is not referenced in the current manuscript). What proportion of each cohort already had CVD and what was the composition of the disease? E.g., proportions with stroke vs MI vs peripheral vascular disease vs... etc. Could a difference in that composition have led to the difference in risk estimates using recalibrated risk scores in the two cohorts?
  ○ The reference survival (So) was based on WHO estimated country-level mortality rates. This would be all-cause mortality, yet I believe the fatal component of the outcome predicted in the 2008 Framingham score is "coronary death". It is fair enough to then estimate the non-fatal events from a ratio, but it will multiplicatively amplify the non-fatal CVD event rate if based on all-cause deaths.
  ○ The country-level mortality from WHO will include people from rural and urban areas, and people with and without existing CVD. Either of these factors may reduce its relevance to the cohorts being studied, and/or the risk score being investigated.
We thank for the reviewer for these thoughtful comments. We can confirm that the

original analyses excluded patients with previous CVD in the ICMR cohort, but we have not excluded them in the CARRS cohort. Therefore, following the suggestion from the reviewer we have re-done the analysis (results remain very similar) and had made this clear in the methods. We have also updated all tables and figures accordingly. We agree that WHO data will include data with and without existing CVD but this is the best dataset that it is available to estimate the CVD mortality in India. While large and robust cohorts are unavailable re-calibration of existing models using the existing local data is the best approach we can take. We have added this as a potential limitation in the discussion.

A wording suggestion - the "Framingham score calculation (ie. Xi)" introduced on page 3 is more standardly called the prognostic index or linear predictor.

We have updated the text following the reviewer's suggestion on page 3

Competing Interests: No competing interests were disclosed.

? **Kay Tee Khaw** (iD)
[1] Gonville and Caius College, University of Cambridge, Cambridge, UK
[2] Gonville and Caius College, University of Cambridge, Cambridge, UK

General comments
Estimating an individual's absolute risk of cardiovascular disease within the next 5 or 10 years is the basis of the "individual high risk" strategy to prevention of cardiovascular disease; those individuals identified at highest risk are targeted for established preventive interventions such as medications to lower blood cholesterol or blood pressure levels to reduce cardiovascular disease risk.

Various risk algorithms are used, most of them based on some variant of the original Framingham cardiovascular risk score, using classical cardiovascular risk factors blood cholesterol, blood pressure, smoking, body mass index and diabetes as well as age and sex.

The accuracy of risk algorithms for predicting cardiovascular disease risk in an individual therefore depends on a number of factors:
Firstly, we require coefficients of risk for the different risk factors, that is how much the level or presence of a risk factor may increase future cardiovascular disease risk. While most may use a standard coefficient for each risk factor such as blood pressure, or smoking there may be variants

on this such as different coefficients for different age groups, or men and women.
Secondly we require the incidence of cardiovascular disease over the next ten years; again, these will differ in men and women and in different age groups as well as different populations, and indeed over time with secular trends.
In addition, in terms of public health implications, the proportion or numbers in the population who might be classified as high risk will also depend on the prevalence of the relevant risk factors in the different groups.
The risk algorithms generally used in clinical and public health, such as the Framingham cardiovascular disease risk score are largely derived from data based originally on the Framingham Study or more recently variants of algorithms from other prospective studies mostly based in Western countries.
The authors of this study make the very strong argument that though the burden of cardiovascular disease in India is huge, and increasing, there is a dearth of data on the use of such risk algorithms as applied to the Indian population, and no validated prognostic model for an Indian urban population

They make the laudable attempt to recalibrate the original Framingham model using data from two urban Indian studies. They estimated three risk score equations using three different models. The first model was based on the Framingham original model, the second and third were recalibrated models using risk factor prevalence from the CARRS and ICMR studies and estimated survival from WHO 2012 data for India. They applied the three risk scores to the CARRS and ICMR participants and estimated the proportions of those at high risk (>30% 10 year CVD risk). The estimates of the proportions with such risk varied greatly using the different models.

Their main conclusion was that their study exemplifies the variation between recalibrated models using different data from the same country. They state that it is essential to develop high quality and well powered local cohorts with outcome data to develop local prognostic models. While I would agree with the overall conclusions of the authors and support strongly the need for locally relevant population studies, there are a number of points that should be clarified in the text.

Issues to consider in the manuscript
The authors in this exercise state that they have recalibrated the original Framingham model using data from two urban Indian studies. They estimated three risk score equations using the first model based on the Framingham original model, the second and third were recalibrated models using risk factor prevalence from two Indian studies, the CARRS and ICMR studies and estimated survival from WHO 2012 data for India.

A general comment: the text mostly refers to CVD (cardiovascular disease) risk but sometimes refers to CHD (coronary heart disease) risk. These are not the same: CVD includes CHD but is a larger category encompassing other conditions such as stroke etc so absolute rates are higher, and the relationship to risk factors is somewhat different. On page 3, there is a statement that the Framingham risk equation is a... method to measure coronary heart disease (CHD) risk but then later under Step 1 the Framingham score it is stated that this is it was calculated for CVD risk. Then in Step 2 the authors state that they obtained 2012 yearly mortality rates for CHD from India. Were the estimates for CVD or CHD? This materially affects the estimates of proportion with any given 10 year absolute risk. In particular, there is a large male:female excess for coronary heart disease but generally stroke rates are more similar between men and women. This needs to be clarified throughout the manuscript.

Most approaches estimating the proportion of those with a given 10 year CVD (or CHD) risk in a defined population (e.g. >30% 10 years or >20% 10 years) generally apply the Framingham coefficients to derive individual Framingham risk scores. The proportion with a high score thus relies both on the coefficients used and the prevalence of risk factors in the population which may vary in different populations. This can be done with cross sectional data in which data on risk factors are available using the Framingham risk score. This was done in Step 1 by the authors using cross sectional data from the two Indian studies: CARRS and ICMR.

The authors then proceeded to develop a model based on 2012 yearly mortality rates for CHD (as stated in the text rather than CVD) for India from WHO (page 3, Step 2). From what I understand, they then derived the average probability of not having events in 10 years for men and women separately. It would seem that because a longitudinal population cohort with individual follow up and endpoints was not available, the WHO data were used to approximate the data to estimate the score of the "average individual in the population by multiplying the averages of the variables by the original Framingham coefficients and adding the values for all risk factors".

Then in Step 4 they calculated the risks for everyone in the CARRS and ICMR data sets with the 3 models using different combinations of reference score and survival probabilities.
I found Steps 2-4 rather hard to follow and it would particularly have been helpful to know more about the assumptions they used in applying the WHO 2012 mortality dataset (apart from the CHD/CVD distinction). For example, they state they used the total CHD annual mortality- was this age standardised, or if not, was this for the whole population or a subset of the population? Presumably the CARRS and ICMR studies encompassed the whole adult age range but it would be helpful to have some more information about the age distribution the rates of cardiovascular disease are so strongly age related. The estimates of the proportion of the population at high cardiovascular risk over 10 years must surely depend hugely on the age distribution of the population, and the Framingham algorithms were derived on a slightly older age distribution than 20year + (CARRS).

Table 1 shows coefficients from the simplified Framingham model. There are a number of references to the Framingham model cited which I have looked up, and I may have missed this, but it would be helpful to cite the exact reference for coefficients used in this table. I was rather surprised for example, in the sex specific coefficients that the women had different coefficients from men for many of the risk factors, whereas in many of the Framingham algorithms that I have seen, the sexes are combined for risk factors with a weighting for male vs female. Body mass index, age, and smoking appear to be more weighted in men whereas blood pressure, and diabetes are more highly weighted in women. This is relevant when considering results.

The tables are clearly presented showing the risk factor distribution in the two population datasets from India (Table 2) and reference scores and survivals in Table 3. The examples of reference score calculations using means from the CARRS population in Table 4 is also of interest with similar differential weighting of the risk factors in men and women as with the Framingham study. However, the results that were surprising were that women in Tables 4 and Figure 1 appeared to have higher Framingham scores. In Table 5, however, it seems that while there was a much higher proportion of men with >30% risk compared to women using the Framingham risk score, the proportions of men and women were much more similar using the F-CARRS or F-ICMR derived scores. Given that the prevalence of risk factors in women in both CARRS and ICMR were generally

lower in women than men apart from body mass index (Table 1), and that women had lower mortality (CHD or CVD) than men this does not really make sense. I wonder whether this may be a consequence of the modelling, (CHD/CVD mortality) or insufficient account of the age distribution. Perhaps they could check their models.

Minor points
The introduction makes a strong argument about the importance of cardiovascular disease in India. However, some of the statements may need some nuance. The statement that the all age death rate increased significantly between 1990 and 2016 for ischemic heart disease. This may reflect both increasing age specific mortality from heart disease or major demographic shifts that is ageing of the population as ischemic heart disease rates increase with age and the numbers and proportion of older people in India have increased over that time period. (In the USA, absolute CVD deaths have increased over the last few decades despite declining age specific or age standardised rates, simply because of the increased numbers of older people). Similarly though, the statement that Indians are affected by CVDs at a younger age compared to their European counterparts may well be true, the statement that more than 50% CVD deaths occur before the age of 70 years may reflect the much younger age distribution of the population – comparisons of risk are not robust without appropriate denominators.

Though the methods used and estimates of proportions at different levels of absolute risk can be discussed, the authors have made the point clearly. I think this manuscript indeed does illustrate and highlight the extraordinary discrepancy between the large and increasing burden of cardiovascular disease in emerging economies globally and the paucity of locally relevant data; they have provided an example of where such local data can be used and the great need for relevant evidence and support of ongoing population studies to inform policy and practice.
As an aside, though not the focus of this manuscript, whatever the model used, the estimates indicate the very high proportion of the population with high absolute cardiovascular risk and challenges of the individual high risk strategy in countries such as India. The discussion is careful and considered in terms of the various guidelines for identifying high risk individuals for preventive interventions but perhaps might also point out the value of local population studies for providing the evidence base for mass preventive strategies.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
No

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiologist.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 25 Nov 2019

Priti Gupta

We thank the opportunity to respond to the reviewers' very positive and helpful comments, which we have now responded point-by-point and, by doing so, we believe it has improved our manuscript

REVIEWER 1
- A general comment: the text mostly refers to CVD (cardiovascular disease) risk but sometimes refers to CHD (coronary heart disease) risk. These are not the same: CVD includes CHD but is a larger category encompassing other conditions such as stroke etc so absolute rates are higher, and the relationship to risk factors is somewhat different. On page 3, there is a statement that the Framingham risk equation is a... method to measure coronary heart disease (CHD) risk but then later under Step 1 the Framingham score it is stated that this is it was calculated for CVD risk. Then in Step 2, the authors state that they obtained 2012 yearly mortality rates for CHD from India. Were the estimates for CVD or CHD? This materially affects the estimates of proportion with any given 10-year absolute risk. In particular, there is a large male: female excess for coronary heart disease but generally stroke rates are more similar between men and women. This needs to be clarified throughout the manuscript.

Response: We thank the reviewer for identifying this inconsistency. The Framingham score we used estimates the probability of CVD and that is what we meant and accordingly we have obtained the corresponding India CVD 2012 mortality rate. We have now edited the manuscript accordingly.
- Most approaches estimating the proportion of those with a given 10 year CVD (or CHD) risk in a defined population (e.g. >30% 10 years or >20% 10 years) generally apply the Framingham coefficients to derive individual Framingham risk scores. The proportion with a high score thus relies both on the coefficients used and the prevalence of risk factors in the population which may vary in different populations. This can be done with cross sectional data in which data on risk factors are available using the Framingham risk score. This was done in Step 1 by the authors using cross sectional data from the two Indian studies: CARRS and ICMR.
  The authors then proceeded to develop a model based on 2012 yearly mortality rates for CHD (as stated in the text rather than CVD) for India from WHO (page 3, Step 2). From what I understand, they then derived the average probability of not having events in 10 years for men and women separately. It would seem that because a longitudinal population cohort with individual follow up and endpoints was not

available, the WHO data were used to approximate the data to estimate the score of the "average individual in the population by multiplying the averages of the variables by the original Framingham coefficients and adding the values for all risk factors". Then in Step 4 they calculated the risks for everyone in the CARRS and ICMR data sets with the 3 models using different combinations of reference score and survival probabilities. I found Steps 2-4 rather hard to follow and it would particularly have been helpful to know more about the assumptions they used in applying the WHO 2012 mortality dataset (apart from the CHD/CVD distinction). For example, they state they used the total CHD annual mortality- was this age standardised, or if not, was this for the whole population or a subset of the population? Presumably the CARRS and ICMR studies encompassed the whole adult age range but it would be helpful to have some more information about the age distribution the rates of cardiovascular disease are so strongly age related. The estimates of the proportion of the population at high cardiovascular risk over 10 years must surely depend hugely on the age distribution of the population, and the Framingham algorithms were derived on a slightly older age distribution than 20year + (CARRS).

Response: We thank the reviewer for the comments, and we hope the answer below clarified the procedure we followed
1. **We did not develop a "*de novo*" risk score model (i.e. we cannot calculate coefficients for the different risk factors because, as the reviewer points out, we have no follow up and outcomes), we use instead the coefficients from the existing Framingham model. What we did is to re-calibrate this model to the Indian data that we are going to predict.**
2. **To re-calibrate a model in a new dataset we need a "reference individual", i.e. an individual with a known value of Framingham score and known risk of having a cardiovascular event.The risk of any other individual is calculated depending on the difference between that individual's score and the reference individual's score.The question is then, how to find the reference individual. In reality no one has a reference individual in our data, so we have to make assumptions. One option is to assume that the population CV risk estimated by the WHO studies would be the risk of a person with average score in the Framingham model in our Indian sample. And so the reference individual will be a person with score produced by setting all variables to the average value and we will assume that their risk is precisely the population CVD risk estimated by WHO.**
3. **WHO estimates of death rates were calculated standardizing by age with the WHO standard population for individuals between 30 and 70 years old.**
   ○ Table 1 shows coefficients from the simplified Framingham model. There are a number of references to the Framingham model cited which I have looked up, and I may have missed this, but it would be helpful to cite the exact reference for coefficients used in this table. I was rather surprised for example, in the sex specific coefficients that the women had different coefficients from men for many of the risk factors, whereas in many of the Framingham algorithms that I have seen, the sexes are combined for risk factors with a weighting for male vs female. Body mass index, age, and smoking appear to be more weighted in men whereas blood pressure, and diabetes are more highly weighted in women. This is relevant when considering results.

Response: We have indeed used the Framingham sex specific coefficients available on supplementary table 1 of D'Agostino *et al*. General Cardiovascular Risk Profile for Use in Primary Care The Framingham Heart Study. Circulation. 2008;117: 743-753. We have now made it clearer and included this reference in the methods and tables i.e new reference number 3.

  ○ The tables are clearly presented showing the risk factor distribution in the two population datasets from India (Table 2) and reference scores and survivals in Table 3. The examples of reference score calculations using means from the CARRS population in Table 4 is also of interest with similar differential weighting of the risk factors in men and women as with the Framingham study. However, the results that were surprising were that women in Tables 4 and Figure 1 appeared to have higher Framingham scores. In Table 5, however, it seems that while there was a much higher proportion of men with >30% risk compared to women using the Framingham risk score, the proportions of men and women were much more similar using the F-CARRS or F-ICMR derived scores. Given that the prevalence of risk factors in women in both CARRS and ICMR were generally lower in women than men apart from body mass index (Table 1), and that women had lower mortality (CHD or CVD) than men this does not really make sense. I wonder whether this may be a consequence of the modelling, (CHD/CVD mortality) or insufficient account of the age distribution. Perhaps they could check their models.

Response: We thank the reviewer for this comment, and we understand the results appear to be counter-intuitive if one considers that women seem to have lower levels of risk factor than men in this population. However, it is important to consider the coefficients multiplying the risk factors when we calculate the score. For example, the coefficients for SBP, both treated and untreated, are higher in women than men, and this is a key factor that adds much weight to the score. Even if women have lower SBP mean than men, when multiplied by the coefficient this increases more their score than in men. To illustrate the contribution of each risk factor we have added a figure in the appendix (see Figure-S1). It can be seen that age and BMI actually contribute to increase the average score of men more than for women, but SBP (both treated and untreated) contribute more to increase the score in women. Adding up all these contributions women end up with a higher mean score than men in both data sets.

Finally, the conversion of the score to a risk factor depends on the re-calibration (the choice of "reference individual" as explained in our response to your question 2). If the score of the reference individual is lowered but while their risk is kept the same, the predicted risks of all other individuals in the data will increase. To illustrate this, we have created another plot (see Figure-S2) where we show how the same distribution of risk scores is converted into different distribution of predicted risks depending on the recalibration model.

Minor points
  ○ The introduction makes a strong argument about the importance of cardiovascular disease in India. However, some of the statements may need some nuance. The statement that the all age death rate increased significantly between 1990 and 2016 for ischemic heart disease. This may reflect both increasing age specific mortality from heart disease or major demographic shifts that is ageing of the population as

ischemic heart disease rates increase with age and the numbers and proportion of older people in India have increased over that time period. (In the USA, absolute CVD deaths have increased over the last few decades despite declining age specific or age standardised rates, simply because of the increased numbers of older people). Similarly though, the statement that Indians are affected by CVDs at a younger age compared to their European counterparts may well be true, the statement that more than 50% CVD deaths occur before the age of 70 years may reflect the much younger age distribution of the population – comparisons of risk are not robust without appropriate denominators.

Response: We agree with the reviewer that there are two issues regarding the statement that CVD affect Indian as a younger age. We have now made it clear that it related both to a demographic age distribution of age plus an increased in age standardized age for CVD.
We have added following line with reference:
"Age adjusted prevalence of CVDs have also increased in India(1,2)"

- Though the methods used and estimates of proportions at different levels of absolute risk can be discussed, the authors have made the point clearly. I think this manuscript indeed does illustrate and highlight the extraordinary discrepancy between the large and increasing burden of cardiovascular disease in emerging economies globally and the paucity of locally relevant data; they have provided an example of where such local data can be used and the great need for relevant evidence and support of ongoing population studies to inform policy and practice. As an aside, though not the focus of this manuscript, whatever the model used, the estimates indicate the very high proportion of the population with high absolute cardiovascular risk and challenges of the individual high risk strategy in countries such as India. The discussion is careful and considered in terms of the various guidelines for identifying high risk individuals for preventive interventions but perhaps might also point out the value of local population studies for providing the evidence base for mass preventive strategies.

Response: We thank the reviewer for the positive comments about our discussion and we agree about the implications of local studies for providing evidence for preventive strategies. We have now edited the discussion section accordingly and inlcuded following lines:
"Also, for the recalibration we used WHO mortality data which includes a population with and without previous CVD, which might have overestimated the risk, however, until large and robust cohorts ,with detailed outcome assessment, become available this is the best available dataset that can be used to estimate the expected mortality in patients without CVD."
"Valid and reliable local prognostic models will also be key to evaluate different high risk prevention strategies."

Competing Interests: No competing interests were disclosed.