European Heart Journal (2010) **31**, 2041–2048
doi:10.1093/eurheartj/ehq189

Is carotid intima media thickness useful for individual prediction of cardiovascular risk? Ten-year results from the Carotid Atherosclerosis Progression Study (CAPS)

Matthias W. Lorenz*, Carolin Schaefer, Helmuth Steinmetz, and Matthias Sitzer

Department of Neurology, Johann Wolfgang Goethe-University, Schleusenweg 2-16, Frankfurt am Main D-60528, Germany

Received 17 January 2010; revised 5 May 2010; accepted 17 May 2010; online publish-ahead-of-print 8 June 2010

Aims Carotid intima media thickness (cIMT) is an intermediate phenotype of early atherosclerosis that independently predicts vascular events. It is often suggested that cIMT be used as a screening tool to select subjects with an elevated event risk. Whether cIMT adds information to traditional risk models has so far received little investigation.

Methods and results The 10-year follow-up of 4904 subjects from the Carotid Atherosclerosis Progression Study (CAPS) without pre-existing vascular disease included cardiovascular events and total mortality. Using Cox models and reclassification statistics, we investigated the usefulness of cIMT in individual risk prediction beyond the Framingham and the SCORE models, using risk strata of 0–5, 5–10, 10–20, and ≥20% over 10 years. Carotid intima media thickness was significantly and independently predictive for cardiovascular events. Compared with a model using the Framingham risk factors, a second model that included the common carotid-IMT led to the reclassification of 357 subjects (8.1%). In 107 subjects (30.0%), this reclassification was correct as confirmed with the actual outcome over 10 years. Net reclassification improvement was −1.41% (P = NS); integrated discrimination improvement was 0.04% (P = NS). More subjects were shifted to lower than to higher risk categories by the inclusion of cIMT. Analyses including other endpoint definitions, other carotid segments, and the SCORE risk model for baseline prediction did not result in consistently better risk prediction with cIMT.

Conclusion Despite cIMT being predictive for cardiovascular endpoints, it did not consistently improve the risk classification of individuals. Carotid intima media thickness may not be useful for the risk stratification of individuals in the general population.

Keywords Carotid arteries • Intima media thickness • Atherosclerosis • Myocardial infarction • Individual risk prediction • Reclassification

Introduction

Carotid intima media thickness (cIMT) is a measure of early atherosclerosis and vascular remodelling that can be assessed quickly, non-invasively, and cheaply with high-resolution ultrasound. It is correlated with all traditional vascular risk factors and regarded as an 'intermediate phenotype' of atherosclerosis or a marker of subclinical organ damage. Carotid intima media thickness independently predicts cardiovascular events.[1]

It has often been proposed that cIMT be used as a screening tool to identify subjects at high vascular risk who need intensive risk factor management. Even when cIMT is a significant and independent predictor of events, this does not necessarily mean that its clinical usefulness substantially exceeds that of traditional risk models, like the Framingham Risk Score (FRS)[2] or the SCORE model.[3] Using appropriate statistical methods to address these issues is important; the methodology is only just beginning to keep pace with clinical needs. New methods have recently

* Corresponding author. Tel: +49 69 6301 83059, Fax: +49 69 6301 5628, Email: matthias.lorenz@em.uni-frankfurt.de

been proposed to evaluate and compare predictive risk models.[4–8]

Current reviews, as instructed by the US Preventive Services Task Force, conclude that cIMT does not yet qualify as a clinically useful tool for risk stratification in individuals.[9] We intend to contribute to this important question using state-of-the-art reclassification statistics.

Methods

All members of a German primary healthcare scheme ($n = 32\,708$), living within a 50 km radius of five study sites in western Germany, were invited to participate. Within a pre-defined time limit, 6962 people (21.3%) agreed to take part. They then underwent ultrasound examination at baseline to determine the IMT in several segments of the carotid arteries. From them, 5056 (from four out of the five study sites) were subsequently followed in order to monitor the incidence of cardiovascular events, stroke, and death. The study was approved by the Ethical Review Committee of the University Hospital of Frankfurt am Main.

A large set of traditional risk factors were determined at baseline, including age, sex, systolic and diastolic blood pressure, high- and low-density lipoprotein cholesterol, diagnosis of diabetes mellitus, and cigarette smoking. Details of the risk-factor assessment protocol have already been published.[10,11]

To study a population in the context of primary prevention, all subjects with previous events before the baseline visit (myocardial infarction or stroke) were excluded; the remaining population comprised 4904 subjects.

Ultrasound imaging

Ultrasound imaging and cIMT image analysis methods have been described in detail in earlier publications.[10] In brief, ultrasonic examinations were performed with a 7.5–10.0 MHz linear array transducer (P700SE; Phillips Medical System). Using antero-oblique insonation, far-wall carotid IMT was visualized bilaterally at three sites: the common carotid artery (CCA-IMT, 20–60 mm proximally from the flow divider), the carotid bifurcation (BIF-IMT, 0–20 mm proximally from the flow divider), and the internal carotid artery bulb (ICA-IMT, 0–20 mm distally from the flow divider). The images were digitally captured during the systole of a single heartbeat for offline measurement. For 1 in every 100 subjects, vertical and horizontal calibration measurements were carried out with an ultrasound quality-assurance phantom. Carotid IMT measurements were performed offline using automated imaging-processing software as reported previously.[10]

The average intraclass correlation coefficient (ICC) for inter-observer reliability was 0.97 (95% CI 0.96–0.98; $P < 0.001$), and the ± 2 SD of the difference between two observers varied between 0.03 and 0.06 mm.[10] Furthermore, the intra-observer test–retest reliability testing revealed an ICC of 0.93 (95% CI 0.91–0.94; $P < 0.001$) and the ± 2 SD of the difference between the first and second examination varied between 0.04 and 0.06 mm.[10]

In addition to cIMT, the presence of plaques in both internal carotid arteries was assessed. A plaque was defined as a focal protrusion into the vessel lumen.

Follow-up

The clinical endpoints reported here were recorded within the scope of the 10-year follow-up.

In a first step, follow-up events were identified from the primary healthcare scheme records. Every time medical care was sought, the healthcare scheme was notified of the date and corresponding ICD code. The healthcare scheme also holds data on death and other reasons for subjects dropping out of the study. These data were screened for cardiovascular events and death.

In a second step, all study participants whose records included a relevant code were contacted by a letter which included a questionnaire about cardiovascular events. Additionally, they were asked to send medical reports, sign informed consent forms so that their files might be reviewed, and name their current physicians. With the participant's written consent, the physicians were contacted to complete the medical records with a focus on the qualifying event. All records were reviewed blind for the cIMT measurements results. Subjects that did not answer our letter were contacted by mail twice more.

As an alternative source of information, the ICD codes were critically reviewed. In the majority of cases, multiple ICD codes were related to the same event [e.g. from the referring practitioner, from the hospital files (at admission and at discharge), and from the rehab unit]. A diagnosis was confirmed when it was the most frequently used category and if it was coded at least twice. In subjects where each ICD code category existed only once, the primary diagnosis at discharge from the emergency hospital was selected.

Statistical analysis

Event risks were assessed with Cox regression models based on censored data. A linear relation between the cIMT variables and risk was assessed by plotting the residuals against the cIMT variable, showing a fair approximation. The use of linear cIMT terms in Cox models is further justified by previous analyses showing that in most studies, the linear model corresponds well to the categorized analysis for higher IMT values.[1] To compare cIMT with established tools for risk prediction, we used two sets of risk factors as 'standard models'. The 'Framingham risk set' includes all risk factors of the FRS[2] (age, sex, systolic and diastolic blood pressure, low- and high-density lipoprotein cholesterol, diabetes, and smoking); the 'SCORE risk set' includes age, sex, systolic blood pressure, total cholesterol, and smoking.

For the assessment of reclassification measures, we applied methods suggested by Cook.[6,8] To compare the additional value of cIMT, risk prediction models including traditional risk factors were compared with models including cIMT. c-Statistics were calculated both for the models with and without cIMT. According to the predicted risk over 10 years, subjects were classified as 'very low risk' (0–5% event risk in 10 years), 'low risk' (5–10%), 'intermediate risk' (10–20%), or 'high risk' ($\geq$20%) with both models, respectively. Reclassification tables were constructed to determine the numbers of subjects who were reclassified by the cIMT models. To assess whether the reclassification of a group of subjects was correct, the observed event risk was calculated as a Kaplan–Meier estimate. When the observed risk was closer to the new than to the old risk category, reclassification was correct. With the Hosmer–Lemeshow test (for reclassified groups of at least 20 subjects), the quality of the old and new models was compared. From the reclassification tables, the number and proportion of correctly reclassified subjects were calculated.

For the calculation of the net reclassification improvement (NRI) and the integrated discrimination improvement (IDI),[8] no censored data are allowed. Therefore, the follow-up was truncated after 7 years; all subjects censored up to this point in time ($n = 57$, 1.1%) were ignored.

Normality of variables was assessed with QQ plots. Normally distributed data were described with mean and SD and skewed data

with median and inter-quartile range. All statistical tests were performed two-sided on a level of significance of 0.05. All statistical calculations were performed using the SAS® 9.1 software package. Reclassification statistics were assessed with the SAS macros published by Cook[8] and macros modified by the first author of this paper.

Results

The distribution of the traditional risk factors, IMT in three carotid segments, and the prevalence of ICA plaques at the baseline visit are displayed in *Table 1*. Additionally, the distributions of the Framingham 10-year risk of a cardiovascular disease (CVD) event, and of the SCORE 10-year risk of a fatal CVD event, are displayed in this table. All continuous variables followed the normal distribution in a good approximation, except the Framingham and SCORE risks and the cIMT variables. The normality of the cIMT variables improved when logarithmized. However, to ease interpretability and practicability, we decided to use cIMT values untransformed. In sensitivity analyses, a subset of analyses was repeated with logarithmized cIMT values.

For 371 subjects, a code for a cardiovascular event was used at least once. All these subjects were contacted by mail and 168 (45.3%) responded. Forty-six (27.4%) reported to having had myocardial infarction. With such a low response rate, we could use the incomplete data set only to validate the reviewed ICD codes at best. The rate of myocardial infarction in the responding subjects did not differ significantly from the rate in non-responding subjects (χ^2 homogeneity test, $P = 0.331$). Without evidence of systematic bias, we chose to use the reviewed ICD codes for the entire cohort. During a follow-up period (mean 8.5 years, SD 0.9 years, range 7.1–10.0 years), 73 subjects (1.5%) suffered myocardial infarction, 271 (5.5%) angina pectoris or myocardial infarction, and 72 subjects (1.5%) died.

Table 1 Descriptive statistics of vascular risk factors

Risk factor	*n* missing (%)	Mean ± SD or proportion
Age (years)	0	49.7 ± 12.9
Sex	0	48.2% males
Systolic blood pressure (mmHg)	2 (0.04)	127.4 ± 17.0
Diastolic blood pressure (mmHg)	6 (0.12)	77.2 ± 10.2
Antihypertensive medication	0	18.1%
Total cholesterol (mmol/L)	7 (0.14)	5.78 ± 1.10
LDL cholesterol (mmol/L)	18 (0.37)	3.34 ± 0.93
HDL cholesterol (mmol/L)	496 (10.11)	1.54 ± 0.44
Lipid-lowering medication	0	4.5%
Diabetes	0	2.5%
Anti-diabetic medication	0	1.8%
Smoking	10 (0.20)	21.3%
Platelet inhibition or anticoagulation	0	2.6%
CCA-IMT[a] (mm)	10 (0.20)	0.72 ± 0.15
BIF-IMT[a] (mm)	21 (0.43)	0.90 ± 0.32
ICA-IMT[a] (mm)	76 (1.55)	0.76 ± 0.30
ICA plaques	0	5.8%
Framingham risk (10-year CVD)	512 (10.44)	5% (7%)[b]
SCORE risk (10-year fatal CVD)	22 (0.45)	0.83% (2.76%)[b]

[a]Mean of left and right sides.
[b]Median (inter-quartile range).

Multiple risk models were assessed to compare the prediction of individual risk of models with traditional risk factors alone with models including cIMT additionally (*Table 2*). As can be seen in the fourth column, cIMT proved to be a significant and independent predictor of events in most models. The c-statistics, corresponding to the area under a receiver operating characteristic (ROC) curve, improved only marginally with cIMT. The Hosmer–Lemeshow test showed a lack of fit for all models examined ($P < 0.001$). The different criteria of reclassification statistics were fulfilled only infrequently and inconsistently. All NRI and IDI values were very small, partly negative and (with the exceptions of model pairs 21 and 23 for NRI with $P = 0.049$ and 0.012, respectively) not statistically significant. *Tables 3* and *4* exemplarily show the reclassification tables for the two model pairs 2 and 17 (printed with red shading in *Table 2*). *Table 3* compares the risk prediction of CVD based on the Framingham risk factors with one which contains cIMT additionally. Only 30% of the reclassified subjects were reclassified correctly. *Table 4* displays the change in risk classification (risk of death) from a SCORE-based model to a model including the SCORE risk factors plus cIMT. Here, 61.4% of all reclassified subjects were reclassified correctly. Over all, 24 model pairs evaluated, the percentage of correctly reclassified subjects varies ~50% (mean 59.3%). In both example reclassification tables, the number of subjects who are shifted to a lower risk category is at least as high as the number of subjects shifted to a higher risk category. *Figure 1* gives the continuous extension of *Table 3*, comparing the actual predicted risks of both models.

To study the influence of subject selection, the analyses were repeated in two subcohorts where subjects with ICA plaques, or subjects with previous medication, respectively, (blood pressure, lipid-lowering, anti-diabetic, or platelet inhibition/anticoagulation) were excluded. To assess the effects of medication more specifically, the above-listed medication variables were included in a third set of models. A subset of analyses was repeated with logarithmized cIMT variables. All these models (data not shown) reproduced the overall pattern of a lack of convincing reclassification benefit by cIMT.

Discussion

Since its first description in 1986,[12] cIMT has evolved into a widely used assessment tool of subclinical atherosclerosis in clinical practice. Its widespread use may be explained by its non-invasiveness, the easy access to high-resolution ultrasound, and the rapid development of the ultrasound imaging quality. Compared with simpler measures, like the ankle-brachial index, cIMT has the psychological advantages of an imaging tool, as patients and physicians tend to believe in visible structures rather than abstract concepts.

A large proportion of coronary events including sudden cardiac death in the general population occur in previously asymptomatic subjects and are often unanticipated. This highlights the necessity

Table 2 Reclassification statistics for the added value of carotid intima media thickness in risk models with different endpoints, carotid segments, and explanatory variables

Endpoint	Model pair no.	Carotid segment	Risk factors included in risk model	Adjusted HR (*P*-value) of cIMT (per 0.1 mm) (Cox model)	c-Statistics for model without cIMT	c-Statistics for model with cIMT	*n* reclassified (*n* total)	*n* correctly reclassified (%)	NRI by cIMT (%)	IDI by cIMT (%)
CVD soft definition[a]	1	CCA-IMT	SCORE RF set	**1.081 (0.015)**	0.718	0.722	258 (4878)	112 (43.4)	0.29	0.00
	2	CCA-IMT	Framingham RF set	**1.113 (0.002)**	0.719	0.724	357 (4384)	107 (30.0)	−1.41	0.04
	3	BIF-IMT	SCORE RF set	**1.052 (0.001)**	0.718	0.725	435 (4866)	188 (43.2)	2.88	0.20
	4	BIF-IMT	Framingham RF set	**1.047 (0.004)**	0.719	0.725	364 (4372)	226 (62.1)	0.14	0.13
	5	ICA-IMT	SCORE RF set	**1.033 (0.031)**	0.718	0.723	280 (4811)	171 (61.1)	1.99	0.07
	6	ICA-IMT	Framingham RF set	**1.037 (0.025)**	0.719	0.722	267 (4319)	118 (44.2)	1.62	0.07
CVD hard definition[b]	7	CCA-IMT	SCORE RF set	1.105 (0.070)	0.742	0.744	68 (4878)	31 (45.6)	1.31	0.07
	8	CCA-IMT	Framingham RF set	1.093 (0.157)	0.735	0.741	25 (4384)	0 (0)	0.14	0.06
	9	BIF-IMT	SCORE RF set	**1.067 (0.011)**	0.742	0.756	107 (4866)	44 (41.1)	2.17	0.08
	10	BIF-IMT	Framingham RF set	**1.061 (0.034)**	0.735	0.752	63 (4372)	20 (31.7)	−3.45	0.04
	11	ICA-IMT	SCORE RF set	**1.071 (0.003)**	0.742	0.757	103 (4811)	45 (43.7)	3.40	0.12
	12	ICA-IMT	Framingham RF set	**1.071 (0.005)**	0.735	0.745	89 (4319)	38 (42.7)	4.93	0.17
Death[c]	13	CCA-IMT	SCORE RF set	1.043 (0.458)	0.798	0.798	0 (4878)	0 (—)	−2.28	−0.08
	14	CCA-IMT	Framingham RF set	1.048 (0.458)	0.798	0.795	28 (4384)	28 (100.0)	−2.36	−0.07
	15	BIF-IMT	SCORE RF set	1.036 (0.200)	0.798	0.793	67 (4866)	0 (0)	−0.72	−0.05
	16	BIF-IMT	Framingham RF set	1.038 (0.192)	0.798	0.796	60 (4372)	60 (100.0)	−2.29	−0.07
	17	ICA-IMT	SCORE RF set	**1.051 (0.038)**	0.798	0.795	114 (4811)	70 (61.4)	6.05	0.25
	18	ICA-IMT	Framingham RF set	**1.056 (0.025)**	0.798	0.792	80 (4319)	46 (57.5)	1.10	0.39
Combined endpoint[d]	19	CCA-IMT	SCORE RF set	**1.063 (0.041)**	0.731	0.733	213 (4878)	80 (37.6)	1.42	-0.02
	20	CCA-IMT	Framingham RF set	**1.089 (0.011)**	0.729	0.732	272 (4384)	186 (68.4)	−2.05	0.00
	21	BIF-IMT	SCORE RF set	**1.046 (0.001)**	0.731	0.735	413 (4866)	202 (48.9)	4.01	0.17
	22	BIF-IMT	Framingham RF set	**1.043 (0.004)**	0.729	0.733	303 (4372)	132 (43.6)	3.59	0.12
	23	ICA-IMT	SCORE RF set	1.026 (0.070)	0.731	0.732	201 (4811)	92 (45.8)	4.10	0.07
	24	ICA-IMT	Framingham RF set	**1.030 (0.049)**	0.729	0.732	215 (4319)	191 (88.8)	0.10	0.10

The highlighted line reports the model shown in *Table 3* and *Figure 1*. Results indicating substantial improvement of risk classification are printed in bold type.
[a]Pectoral angina or myocardial infarction.
[b]Myocardial infarction.
[c]All-cause mortality.
[d]Pectoral angina or myocardial infarction or death.

Table 3 Reclassification table comparing 10-year risk strata for a risk model including the Framingham risk set for cardiovascular disease (soft definition) with and without carotid intima media thickness

Model without cIMT	Model with cIMT				Total	Reclassified into new risk category		
	0–5%	5–10%	10–20%	≥20%		Lower	Higher	Total
0–5%								
At risk	2513	70	0	0	2583	—	70	70
Events	153	2	0	0	155	—	2	2
No event	2360	68	0	0	2428	—	68	68
Observed risk[a] (%)	7.20	2.94	—	—	7.12	—	—	—
5–10%								
At risk	107	1109	73	0	1289	107	73	180
Events	5	55	5	0	65	5	5	10
No event	102	1054	68	0	1224	102	68	170
Observed risk[a] (%)	5.09	5.87	6.85	—	5.85	—	—	—
10–20%								
At risk	0	85	370	22	477	85	22	107
Events	0	2	14	1	17	2	1	3
No event	0	83	365	21	460	83	21	104
Observed risk[a] (%)	—	2.38	5.64	25.00	5.38	—	—	—
≥20%								
At risk	0	0	10	25	35	10	—	10
Events	0	0	0	1	1	0	—	0
No event	0	0	10	24	34	10	—	10
Observed risk[a] (%)	—	—	0	4.17	3.03	—	—	—
Total								
At risk	2620	1264	453	47	4384	202	165	367
Events	158	59	19	2	238	7	8	15
No event	2462	1205	434	45	4146	195	157	352
Observed risk[a] (%)	7.12	5.48	5.71	8.79	6.61	—	—	—

Red background refers to a decrease in risk stratum; blue background to an increase in risk stratum.
[a]Kaplan–Meier estimate.

to close the gaps that are left between risk assessment with traditional risk factors and the real event risk. Carotid intima media thickness—as a biomarker halfway between risk factors and organ damage—has raised hopes of bridging this gap and helping to prevent clinical events.

Recommendations to use cIMT assessment in clinical practice have been given early but in an unspecified manner, leaving the practitioner uncertain of how to use cIMT with several questions remaining unanswered. Which threshold values should be used? How should the risk factor management be altered in cases of advanced subclinical atherosclerosis? The problems of standardization of cIMT assessment are discussed elsewhere.[13]

Several publications have discussed approaches to modify risk classification with cIMT.[14–17] The most frequently suggested method is to use 'vascular age' calculated from regression equations including cIMT instead of chronological age in risk models like the FRS.[15–17] However, these approaches have never been validated in longitudinal cohort studies. Proving the clinical usefulness of a new biomarker in this context is anything but trivial from the methodological aspect, and the statistical methods are only beginning to keep pace with the clinical needs.[5–8] Today, most experts agree that approaches using ROC analyses and c-statistics have severe flaws, and that newer methods like the reclassification calibration statistics are more appropriate.[4–8]

To the best of our knowledge, only two publications have addressed this issue with adequate statistical methods to date. Baldassare *et al.*[18] have assessed traditional risk factors and cIMT in a cohort of 286 dyslipidaemic subjects and followed them up for 5 years. They found that in the intermediate risk category (FRS 10–20% in 10 years), a proportion of subjects could be successfully reclassified into the high-risk category with cIMT (cIMT from the age-specific 60% quantile in men, or from the 80% quantile in women). Despite Baldassare *et al.*[18] not using reclassification calibration statistics, they adopted the method of a reclassification table and validated the reclassification with follow-up data. A general population was investigated in only one opportunity, and these results have never been published in a full paper. Nambi

Table 4 Reclassification table comparing 10-year risk strata for a risk model including the SCORE risk set for death with and without carotid intima media thickness

Model without cIMT	Model with cIMT				Total	Reclassified into new risk category		
	0–5%	5–10%	10–20%	≥20%		Lower	Higher	Total
0–5%								
At risk	4515	47	1	0	4563	—	48	48
Events	57	3	0	0	60	—	3	3
No event	4458	44	1	0	4503	—	45	45
Observed risk[a] (%)	1.42	6.38	0.00	—	1.47	—	—	—
5–10%								
At risk	44	132	15	0	191	44	15	59
Events	1	2	1	0	4	1	1	2
No event	43	139	14	0	187	43	14	57
Observed risk[a] (%)	20.00	3.76	6.67	—	8.09	—	—	—
10–20%								
At risk	0	23	27	4	54	23	4	27
Events	0	0	5	1	6	0	1	1
No event	0	23	22	3	48	23	3	26
Observed risk[a] (%)	—	0.00	53.21	25.00	26.45	—	—	—
≥20%								
At risk	0	0	0	3	3	0	—	0
Events	0	0	0	0	0	0	—	0
No event	0	0	0	3	3	0	—	0
Observed risk[a] (%)	—	—	—	0.00	0.00	—	—	—
Total								
At risk	4559	202	43	7	4811	67	67	134
Events	58	5	6	1	70	1	5	6
No event	4501	197	37	6	4741	66	62	128
Observed risk[a] (%)	1.49	4.16	43.98	16.67	1.96	—	—	—

Red background refers to a decrease in risk stratum; blue background to an increase in risk stratum.
[a]Kaplan–Meier estimate.

et al.[19] presented a congress abstract in 2007 where a subset of the Atherosclerosis Risk In Communities (ARIC) study was used. The cohort was classified into risk categories with threshold values of 6, 10, and 20% CVD risk over 10 years. The classification was done with traditional risk factors and one model additionally including cIMT; the reclassification table showed that a large proportion of the cohort was reclassified, and the prediction of observed CVD events improved modestly, as shown with the Hosmer–Lemeshow test. Why these data have not been published in a peer-reviewed journal is unknown to the authors of this paper.

The data from the 10-year follow-up of Carotid Atherosclerosis Progression Study (CAPS) show the results of 24 model pairs assessed with reclassification tables and reclassification calibration statistics. The large number of models were chosen to avoid the impression of arbitrarily (and *post hoc*) selected results and are related to the number of clinical endpoints and carotid segments. Among these, single models result in impressive reclassification tables where the proportion of reclassified subjects is considerable and the majority of reclassifications was correct (e.g. model pairs 4 or 24). However, other models are less impressive and the average percentage of correctly reclassified subjects over all model pairs is only 59%. The NRI is significant in only 2 of 24 model pairs, inconsistent with the above-given results; and the IDI is never significant. Considering the multiple testing problem, these results are thoroughly unconvincing.

Keeping in mind the clinical context, we have to be aware of the practical consequences of reclassification. Current concepts, as formalized in the current European guidelines,[20] suggest using a CVD mortality of 5% over 10 years (or an event risk of 10% over 10 years) as threshold for intensive risk factor management and a CVD mortality of 10% over 10 years (or an event risk of ~20% over 10 years) as threshold for the prescription of daily low-dose aspirin.[20] The reclassification table shown in *Table 3* reports only 22 individuals out of 4384 (0.5%) to be shifted over the 'aspirin threshold' of 20%, and 95 out of 4384 subjects (2.2%) over the 'intensified RF management threshold' of 10%. The effect of intensified RF management is difficult to quantify, but as a crude approximation we may assume that the daily intake of aspirin reduces the

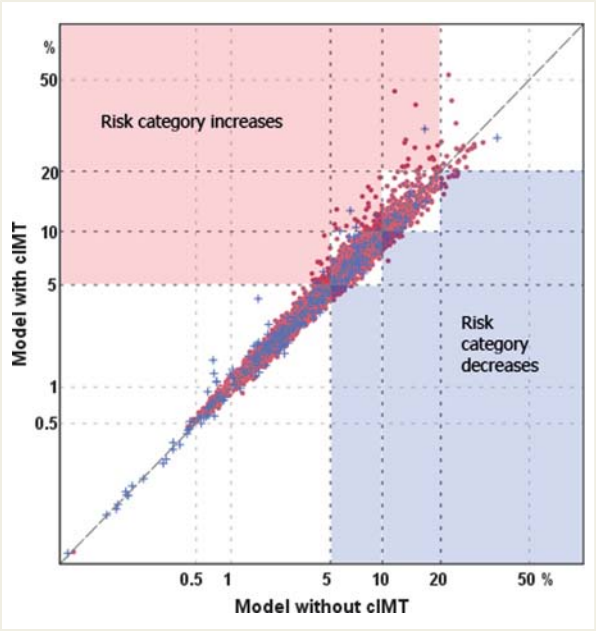

Figure 1 Predicted 10-year risk for cardiovascular disease (soft definition) from models including the Framingham risk set with and without the common carotid artery-intima media thickness. The diagonal line represents the line of identity. The vertical lines show the cutpoints of the risk strata. Subjects with events are displayed with a blue cross; subjects without an event are represented by red circles.

risk of a CVD event by one-third. Given an event risk of 25% over 10 years of this specific subcohort (*Table 3*), we may be able to reduce this absolute risk by 8.3%. To prevent one CVD event over 10 years, we would have to screen 2410 subjects with cIMT measurement. This simple calculation still neglects the fact that the reclassification statistics are not statistically significant.

Another important issue is to look at the direction of reclassification. In both of the reclassification tables displayed, the number of subjects who were downgraded by cIMT exceeds the number of those who were upgraded. More than half of the clinical use of high-resolution ultrasound would be to keep a few individuals from (unjustified) treatment of risk factors.

At first view, these results may appear to contradict the fact that cIMT is independently predictive of future clinical events. It is therefore important to highlight the fundamental difference of risk prediction in populations and in individuals. To (independently) predict event risk in populations, a biomarker needs to have some association with event risk unexplained by other risk factors. This independent fraction of an association may well be very weak; often large populations are needed to verify such effects. To predict event risk in individuals in a meaningful way, a biomarker needs to explain such a large proportion of the event risks that a relevant proportion of subjects cross a treatment threshold in the same direction, when the new biomarker is added to risk factors of a predictive model. We are aware that in such relatively complex study designs, simple choices may have a considerable impact on the results. We therefore performed multiple sensitivity analyses to study the effects of excluding subjects with plaques, individuals on relevant medication, and the effect of medication on the risk models. In all models resulting from these sensitivity analyses, we found results very much similar to those that we reported. We therefore have good reason to believe that these results are robust and valid.

However, our study has several limitations. One is that there was no follow-up visit to assess the occurrence of clinical events. The assessment of clinical events is based on health insurance data, which may not be accurately coded, and questionnaires, which were returned by only 45% of subjects who had an event coded. Nevertheless, our approach to validate the health insurance codes with the questionnaire data still makes the best use of the available data.

As the endpoint definitions rely on ICD codes, it is difficult to assess endpoints like 'fatal CVD' or 'vascular death' because the coding quality in cases of death is often poor and the causal link between event and death is not coded. Therefore, the endpoints used in the original SCORE model had to be substituted by others.

With 24 pairs of risk models and 5 statistical tests, the multiple testing problem arises. However, when even without the adjustment of the significance level no convincing or significant tests are shown, these negative results are not corrupted by the multiple testing problem.

One of the premises to interpret a Cox risk model is that the independent variable of primary interest is linearly associated with risk. In the current data set, there was no evidence that the form of this association was anything other than linear. However, previous analyses of multiple large data sets[1] showed that the cIMT risk association can sometimes be nonlinear, but resembles a linear term at least for higher cIMT. Therefore, minor imprecision may not be excluded but a major mistake can be ruled out.

In summary, reclassification analyses of the 10-year follow-up of the CAPS do not substantiate that cIMT is clinically useful for risk classification of individuals in the context of primary prevention. One reason for this important limitation may be that even when small groups with extreme cIMT values are selected, the absolute risk that can be predicted with cIMT in population samples rarely exceeds 20% in 10 years.[21] For risk populations, cIMT may still be a valid tool.[18]

Before giving practical recommendations, our (negative) results need to be confirmed by other workgroups in other population-based studies. With data from only one large population cohort, we are unable and unwilling to discourage any efforts to use cIMT as a screening tool. However, such recommendations must be well founded on adequately designed studies in population cohorts. Even more important (and perhaps more promising), a wide variety of risk cohorts needs to be evaluated with equal prudence, like hypertensive, dyslipidaemic, or (pre-)diabetic subjects.

Acknowledgements

We are indebted to Dr Michael Skutta, Dr Jens Moll, Dr Donata Ruhkamp, and Dr Alexandra Buehler for professional data collection.

Funding

This study was supported by grants from the Stiftung Deutsche Schlaganfall-Hilfe (German Stroke Foundation).

Conflict of interest: none declared.

References

1. Lorenz MW, Markus HS, Bots ML, Rosvall M, Sitzer M. Prediction of clinical cardiovascular events with carotid intima-media thickness—a systematic review and meta-analysis. *Circulation* 2007;**115**:459–467.
2. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;**97**:1837–1847.
3. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, De Bacquer D, Ducimetière P, Jousilahti P, Keil U, Njølstad I, Oganov RG, Thomsen T, Tunstall-Pedoe H, Tverdal A, Wedel H, Whincup P, Wilhelmsen L, Graham IM, SCORE Project Group. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;**24**:987–1003.
4. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007;**115**:928–935.
5. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;**27**:157–172. Discussion, 207–212.
6. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem* 2008;**54**:17–23.
7. Cui J. Overview of risk prediction models in cardiovascular disease research. *Ann Epidemiol* 2009;**19**:711–717.
8. Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann Intern Med* 2009;**150**:795–802.
9. Helfand M, Buckley DI, Freeman M, Fu R, Rogers K, Fleming C, Humphrey LL. Emerging risk factors for coronary heart disease: a summary of systematic reviews conducted for the U.S. Preventive Services Task Force. *Ann Intern Med* 2009;**151**:496–507.
10. Sitzer M, Markus HS, Mendall MA, Liehr R, Knorr U, Steinmetz H. C-reactive protein and carotid intimal medial thickness in a community population. *J Cardiovasc Risk* 2002;**9**:97–103.
11. Jerrard-Dunne P, Sitzer M, Risley P, Steckel DA, Buehler A, von Kegler S, Markus HS, Carotid Atherosclerosis Progression Study. Interleukin-6 promoter polymorphism modulates the effects of heavy alcohol consumption on early carotid artery atherosclerosis: the Carotid Atherosclerosis Progression Study (CAPS). *Stroke* 2003;**34**:402–407.
12. Pignoli P, Tremoli E, Poli A, Oreste P, Paoletti R. Intimal plus medial thickness of the arterial wall: a direct measurement with ultrasound imaging. *Circulation* 1986;**74**:1399–1406.
13. Touboul PJ, Hennerici MG, Meairs S, Adams H, Amarenco P, Bornstein N, Csiba L, Desvarieux M, Ebrahim S, Fatar M, Hernandez Hernandez R, Jaff M, Kownator S, Prati P, Rundek T, Sitzer M, Schminke U, Tardif JC, Taylor A, Vicaut E, Woo KS, Zannad F, Zureik M. Mannheim carotid intima-media thickness consensus (2004–2006). An update on behalf of the Advisory Board of the 3rd and 4th Watching the Risk Symposium, 13th and 15th European Stroke Conferences, Mannheim, Germany, 2004, and Brussels, Belgium, 2006. *Cerebrovasc Dis* 2007;**23**:75–80.
14. Bard RL, Kalsi H, Rubenfire M, Wakefield T, Fex B, Rajagopalan S, Brook RD. Effect of carotid atherosclerosis screening on risk stratification during primary cardiovascular disease prevention. *Am J Cardiol* 2004;**93**:1030–1032.
15. Stein JH, Fraizer MC, Aeschlimann SE, Nelson-Worel J, McBride PE, Douglas PS. Vascular age: integrating carotid intima-media thickness measurements with global coronary risk assessment. *Clin Cardiol* 2004;**27**:388–392.
16. Gepner AD, Keevil JG, Wyman RA, Korcarz CE, Aeschlimann SE, Busse KL, Stein JH. Use of carotid intima-media thickness and vascular age to modify cardiovascular risk prediction. *J Am Soc Echocardiogr* 2006;**19**:1170–1174.
17. Junyent M, Zambón D, Gilabert R, Núñez I, Cofán M, Ros E. Carotid atherosclerosis and vascular age in the assessment of coronary heart disease risk beyond the Framingham risk score. *Atherosclerosis* 2008;**196**:803–809.
18. Baldassarre D, Amato M, Pustina L, Castelnuovo S, Sanvito S, Gerosa L, Veglia F, Keidar S, Tremoli E, Sirtori CR. Measurement of carotid artery intima-media thickness in dyslipidemic patients increases the power of traditional risk factors to predict cardiovascular events. *Atherosclerosis* 2007;**191**:403–408.
19. Nambi V, Chambless L, Folsom AR, Hu Y, Mosley TH, Volcik K, Boerwinckle E, Ballantyne CM. Carotid intima-media thickness improves coronary heart disease risk prediction in the atherosclerosis risk in communities (ARIC) study. (Congress Abstract). *Circulation* 2007;**116**:II_809.
20. Graham I, Atar D, Borch-Johnsen K, Boysen G, Burell G, Cifkova R, Dallongeville J, De Backer G, Ebrahim S, Gjelsvik B, Herrmann-Lingen C, Hoes A, Humphries S, Knapton M, Perk J, Priori SG, Pyorala K, Reiner Z, Ruilope L, Sans-Menendez S, Op Reimer WS, Weissberg P, Wood D, Yarnell J, Zamorano JL, Walma E, Fitzgerald T, Cooney MT, Dudina A, Vahanian A, Camm J, De Caterina R, Dean V, Dickstein K, Funck-Brentano C, Filippatos G, Hellemans I, Kristensen SD, McGregor K, Sechtem U, Silber S, Tendera M, Widimsky P, Zamorano JL, Altiner A, Bonora E, Durrington PN, Fagard R, Giampaoli S, Hemingway H, Hakansson J, Kjeldsen SE, Larsen ML, Mancia G, Manolis AJ, Orth-Gomer K, Pedersen T, Rayner M, Ryden L, Sammut M, Schneiderman N, Stalenhoef AF, Tokgözoglu L, Wiklund O, Zampelas A, European Society of Cardiology (ESC); European Association for Cardiovascular Prevention, Rehabilitation (EACPR); Council on Cardiovascular Nursing; European Association for Study of Diabetes (EASD); International Diabetes Federation Europe (IDF-Europe); European Stroke Initiative (EUSI); International Society of Behavioural Medicine (ISBM); European Society of Hypertension (ESH); European Society of General Practice/Family Medicine (ESGP/FM/WONCA); European Heart Network (EHN). European guidelines on cardiovascular disease prevention in clinical practice: executive summary. Fourth Joint Task Force of the European Society of Cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of nine societies and by invited experts). *Eur J Cardiovasc Prev Rehabil* 2007;**14**(Suppl. 2):E1–E40.
21. Simon A, Megnien JL, Chironi G. The value of carotid intima-media thickness for predicting cardiovascular risk. *Arterioscler Thromb Vasc Biol* 2010;**30**:182–185.