NON-INVASIVE MARKERS OF SUBCLINICAL ATHEROSCLEROSIS FOR PREDICTING A PRIMARY CARDIOVASCULAR EVENT: A RAPID SYSTEMATIC REVIEW
NON-INVASIVE MARKERS OF SUBCLINICAL ATHEROSCLEROSIS FOR PREDICTING A PRIMARY CARDIOVASCULAR EVENT: A RAPID SYSTEMATIC REVIEW

DOMINIQUE ROBERFROID, LORENA SAN MIGUEL, DOMINIQUE PAULUS
Non-invasive markers of subclinical atherosclerosis for predicting a primary cardiovascular event: a rapid systematic review

Authors: Dominique Roberfroid (KCE), San Miguel Lorena (KCE), Dominique Paulus (KCE)

Reviewers: Mattias Neyt (KCE), Roos Leroy (KCE)

External experts: Johan De Sutter (Universiteit Gent), Bernhard Gerber (UCL), Geert Goderis (KU Leuven), Alexandre Persu (UCL), Ernst Rietzschel (Universiteit Gent), Muriel Sprynger (CHU-ULG), Jean-Claude Wautrecht (ULB)

External validators: Christian Brohet (UCL), Olivier Descamps (Polyclinique du Centre Hospitalier de Jolimont), Guy De Backer (Universiteit Gent)

Acknowledgements: Nicolas Fairon (KCE), Luc Hourlay (KCE)

Other reported interests: Membership of a stakeholder group on which the results of this report could have an impact: Muriel Sprynger (unpaid participation to a study (ESAOTE) and collaboration with EACVI (European Association of cardiovascular Imaging)

Payments to speak, training remuneration, subsidised travel or payment for participation at a conference: Geert Goderis (MSD Symposium)

Participation in scientific or experimental research as an initiator, principal investigator or researcher: Olivier Descamps (clinical studies, Sanofi, AMGEN), Ernst Riezshecl (PI Asklepios study)

Layout: Ine Verhulst

Disclaimer:

- The external experts were consulted about a (preliminary) version of the scientific report. Their comments were discussed during meetings. They did not co-author the scientific report and did not necessarily agree with its content.
- Subsequently, a (final) version was submitted to the validators. The validation of the report results from a consensus or a voting process between the validators. The validators did not co-author the scientific report and did not necessarily all three agree with its content.
- Finally, this report has been approved by common assent by the Executive Board.
- Only the KCE is responsible for errors or omissions that could persist. The policy recommendations are also under the full responsibility of the KCE.
# TABLE OF CONTENTS

- **LIST OF FIGURES** ...............................................................................................................................................2
- **LIST OF TABLES** .................................................................................................................................................2
- **LIST OF ABBREVIATIONS** ..................................................................................................................................4

**SCIENTIFIC REPORT** ........................................................................................................................................5

1 **BACKGROUND** ..............................................................................................................................................5

  1.1 THE NEED FOR MORE ACCURATE PREDICTION.........................................................................................5
  1.2 USEFUL STATISTICS.........................................................................................................................................8
  1.3 OBJECTIVES ....................................................................................................................................................8

2 **INCREMENTAL PREDICTIVE VALUE** ............................................................................................................10

  2.1 METHODS .......................................................................................................................................................10
  2.2 RESULTS .........................................................................................................................................................11
    2.2.1 Flow-mediated dilation .........................................................................................................................12
    2.2.2 Ankle-Brachial Index (ABI) ................................................................................................................13
    2.2.3 Coronary artery calcification score (CAC) ............................................................................................19
    2.2.4 Carotid intima-media thickness (cIMT) ................................................................................................22
    2.2.5 Carotid plaques .......................................................................................................................................24
  2.3 DISCUSSION ..................................................................................................................................................27

3 **CLINICAL BENEFIT OF IMPROVED RISK PREDICTION** .............................................................................28

  3.1 OBJECTIVES ..................................................................................................................................................28
  3.2 METHODS .......................................................................................................................................................28
  3.3 RESULTS .........................................................................................................................................................29
  3.4 DISCUSSION ..................................................................................................................................................29

4 **REVIEW OF EVIDENCE ON COST-EFFECTIVENESS** ....................................................................................29

  4.1 METHODS .......................................................................................................................................................29
    4.1.1 Search strategy ........................................................................................................................................29
    4.1.2 Selection procedure .............................................................................................................................30
    4.1.3 Selection criteria ....................................................................................................................................30
4.2 OVERVIEW OF ECONOMIC EVALUATIONS

4.2.1 Type of economic evaluation

4.2.2 Time frame of analyses and discounting

4.2.3 Population

4.2.4 Intervention and comparator

4.2.5 Cost and outcome inputs

4.2.6 Modelling

4.2.7 Results

4.2.8 Sensitivity analysis

4.2.9 Conflict of interest

4.2.10 Discussion

4.2.11 Conclusion

5 DISCUSSION

BIBLIOGRAPHY

LIST OF FIGURES

Figure 1 – Flow chart of evidence retrieval

LIST OF TABLES

Table 1 – Non-invasive markers of subclinical atherosclerosis studied in this report
Table 2 – Research question: NRI
Table 3 – Flow-mediated dilation
Table 4 – Ankle-brachial index
Table 5 – Excluded ABI studies and reasons
Table 6 – Excluded studies on aPWV and reasons
Table 7 – aPWV
Table 8 – Coronary artery calcium score
Table 9 – Excluded studies on CAC
Table 10 – Studies on cIMT excluded on abstracts
Table 11 – cIMT
Table 12 – Excluded studies on Carotid Plaques
<table>
<thead>
<tr>
<th>Table No.</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Table 13 – NRI &amp; CNRI of carotid plaques</td>
<td>26</td>
</tr>
<tr>
<td>14</td>
<td>Table 14 – NRI and CNRI from markers added to the Framingham model to predict a first cardiovascular event</td>
<td>27</td>
</tr>
<tr>
<td>15</td>
<td>Table 15 – Research question on clinical effectiveness of atherosclerosis markers</td>
<td>28</td>
</tr>
<tr>
<td>16</td>
<td>Table 16 – Selection criteria for economic evaluations</td>
<td>30</td>
</tr>
<tr>
<td>17</td>
<td>Table 17 – Overview of economic evaluations of CVD risk markers for individuals with no previous history of cardiovascular disease/event</td>
<td>31</td>
</tr>
</tbody>
</table>
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>ABBREVIATION</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABI</td>
<td>Ankle-brachial index</td>
</tr>
<tr>
<td>aPWV</td>
<td>Aortic pulse wave velocity</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>CP</td>
<td>Carotid plaques</td>
</tr>
<tr>
<td>cIMT</td>
<td>Carotid Intima-media thickness</td>
</tr>
<tr>
<td>CAC</td>
<td>Coronary artery calcium score</td>
</tr>
<tr>
<td>CHD</td>
<td>Coronary Heart Disease</td>
</tr>
<tr>
<td>CNRI</td>
<td>Clinical Net Reclassification Improvement</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular Disease</td>
</tr>
<tr>
<td>FMD</td>
<td>Flow-mediated dilation</td>
</tr>
<tr>
<td>FRS</td>
<td>Framingham Risk Score</td>
</tr>
<tr>
<td>GRADE</td>
<td>Grading of Recommendations Assessment, Development and Evaluation</td>
</tr>
<tr>
<td>HDL</td>
<td>High Density Lipoprotein</td>
</tr>
<tr>
<td>HTA</td>
<td>Health Technology Assessment</td>
</tr>
<tr>
<td>ICER</td>
<td>Incremental Cost Effectiveness Ratio</td>
</tr>
<tr>
<td>LYG</td>
<td>Life Years Gained</td>
</tr>
<tr>
<td>MI</td>
<td>Myocardial Infarction</td>
</tr>
<tr>
<td>NRI</td>
<td>Net Reclassification Improvement</td>
</tr>
<tr>
<td>QALY</td>
<td>Quality Adjusted Life Years</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised Controlled Trial</td>
</tr>
<tr>
<td>SCORE</td>
<td>Systematic COronary Risk Evaluation</td>
</tr>
</tbody>
</table>
1 BACKGROUND

1.1 The need for more accurate prediction

Cardio-vascular diseases (CVD) remain the most important cause of mortality in our population. Primary prevention is thus crucial. CVD have a long asymptomatic latent period, which provides an opportunity for early preventive interventions. Current trends in primary prevention of CVD emphasize the need to treat individuals based on their global cardiovascular risk. Risk prediction models, such as the SCORE model (Systematic Coronary Risk Evaluation) used in Belgium and other European countries, or the Framingham score (FRS) primarily used in the USA, are key components of prevention strategies by allowing the identification and appropriate management of vulnerable individuals. Following the assessment of the absolute individual CVD risk, individuals will be classified at low risk (<1% of CVD death at 10 years with the SCORE, or <10% of CVD event at 10 years for the Framingham score), intermediate risk (1-5% in SCORE, 10-20% in FRS), or high-risk (>5% in SCORE, >20% in FRS). These risk categories bear a clinical meaning, as there is a risk threshold above which treatment is recommended and below which it is not. For example, in Belgium statins are currently reimbursed only if SCORE ≥5%. Of course, any threshold is arbitrary and clinical skills remain central to adapt risk evaluation and management according to each individual situation, but risk stratification can indicate if the clinical follow-up should be more or less intensive.

However, there is increasing recognition of the imprecision of risk classifications generated by these prediction models. With a cut-off of 5% in 10-year mortality risk, the sensitivity of the SCORE model is 52% (13% for women, 60% for men) and its specificity is 85% (98% for women, 76% for men). With a SCORE calibrated for the Belgian epidemiology, the sensitivity and specificity are 77% (60% in women, 85% in men) and 72% (83% in women, 61% in men). The intermediate-risk group actually represents a composite of higher-risk individuals, for whom more aggressive therapy might be indicated, and lower-risk individuals in whom CVD might not occur.

Overall, CVD remains the first cause of death (29.7% vs. 27.2% for cancer), although in men cancer is the first cause of mortality (cancer: 30.5% of deaths; CVD: 27.3%). In women CVD is the main cause of death (cancer: 23.9%; CVD: 32.1%). Source: Statbel 2011
be managed with lifestyle measures alone. The performance of prediction models could be improved by integrating novel markers of cardiovascular risk.

A more accurate evaluation of CVD risk could also have a substantial impact on the incidence of CVD, since the highest number of CVD events occurs in individuals classified in the intermediate or lower risk groups.

New markers of CVD are increasingly numerous. In a previous KCE report, we reviewed the added predictive value of a series of serum biomarkers, such as the C-Reactive Protein (CRP) or natriuretic peptides. We assess here the added predictive value of another set of emerging risk markers, i.e. markers of subclinical atherosclerosis such as flow-mediated dilatation, aortic pulse wave velocity, ankle-brachial index, carotid intima-media thickness, and coronary artery calcium score (Table 1). These markers are not the only ones available (e.g. cardiac stress test, retinal vessel calibers, pericardial adipose tissue), but they are the most studied ones and an initial scoping review demonstrated that the evidence on incremental predictive value of the other markers of sub-clinical atherosclerosis is marginal. Moreover, we focused on non-invasive markers.

All these markers are associated with the occurrence of CVD. For instance, a recent meta-analysis including 45,828 individuals reported that the hazard ratio per 0.1-mm difference of common CIMT was 1.08 (95%CI, 1.05-1.10) for myocardial infarction and 1.12 (95% CI, 1.10-1.15) for stroke. But the crucial question before considering the adoption of these markers in routine clinics is whether the information provided adds to the information already based on the traditional risk factors, e.g. the Framingham score or the SCORE. A new marker should demonstrate its ability to correctly reclassify individuals into clinically meaningful risk categories that support clinicians to improve clinical decision-making.
<table>
<thead>
<tr>
<th>Table 1 – Non-invasive markers of subclinical atherosclerosis studied in this report</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Flow-mediated dilation (FMD)</strong></td>
</tr>
<tr>
<td><strong>Aortic pulse wave velocity (aPWV)</strong></td>
</tr>
<tr>
<td><strong>Ankle-brachial index (ABI)</strong></td>
</tr>
<tr>
<td><strong>Carotid Intima-media thickness (cIMT)</strong></td>
</tr>
<tr>
<td><strong>Carotid plaques (CP)</strong></td>
</tr>
<tr>
<td><strong>Coronary artery calcium score (CAC)</strong></td>
</tr>
</tbody>
</table>
1.2 Useful statistics

As explained in our previous KCE report\(^5\), the most sensible and useful summary statistics regarding the added predictive value of integrated models is the Net Reclassification Index (NRI). NRI summarizes the net proportion of individuals with “correct” reclassification (e.g., those who develop events who were up-classified, and those who do not develop events who were down-classified) and “incorrect” reclassification (those who develop events who were down-classified, and those who do not develop events who were up-classified)\(^6\).

\[
NRI = (P_{up|D=1} - P_{down|D=1}) - (P_{up|D=0} - P_{down|D=0})
\]

where “D” denotes the event indicator, “up” an up-reclassification and “down” a down-reclassification. The null hypothesis of \(NRI=0\) can be formally tested with a simple asymptotic test\(^b\):

\[
Z = \frac{NRI}{\sqrt{\frac{(P_{up|D=1} - P_{down|D=1})}{nD = 1} + \frac{(P_{up|D=0} - P_{down|D=0})}{nD = 0}}}
\]

For example, if among 100 individuals developing a cardiovascular event classified at intermediate risk by FRS, 10 are correctly reclassified at high risk by the model FRS+marker and 5 incorrectly reclassified at low-risk, then the % of net correct reclassification in individuals with events will be 5\% ((10-5)/100). The % of net correct reclassification is then also computed in individuals remaining asymptomatic, and the NRI corresponds to the sum of ‘NRI-event’ and ‘NRI-non-event’. These 2 metrics, which can be tested separately, should also be reported in assessment studies for fuller interpretation\(^17\). In general, no more than 3 categories are recommended (high risk individuals who need treatment, intermediate risk individuals for whom the need of treatment is unclear, low risk individuals who do not need treatment). One difficulty of the NRI is its dependence upon the number of risk categories and the choice of cut-off points. When established risk categories are used in clinical practice to adapt the strategy of risk reduction, an NRI computed on these categories will be the most informative statistics\(^17\). In other instances, a “category-free” NRI, which does not depend on the existence of fixed risk categories, has been proposed recently\(^17\). There is no generally accepted cut-off for a clinically important NRI as it depends on the relative benefits and risks of overtreatment versus undertreatment\(^16\).

Some authors have also introduced the concept of clinical NRI (CNRI), i.e. the amount of reclassification observed only in individuals classified in the intermediate-risk category by the reference prediction model\(^19\). It calculates the amount of improvement offered by a strategy where only the individuals for whom the treatment decision could be changed by measuring a biomarker are considered. Usually, there is a risk threshold above which treatment is recommended and below which it is not\(^3\). For very-low-risk patients, it is rational that neither testing nor treatment is needed, whereas for high-risk patients, treatment is indicated without further testing, because no test result would reduce their estimated risk below the treatment threshold. This assumes a 2-step screening strategy where individuals would first be classified based on the reference prediction model of CVD risk, and the risk marker is then measured only in intermediate-risk individuals. Patients who have a risk that is either just above or just below a treatment threshold might be moved across the threshold and have their treatment changed by the ascertainment of additional risk information.

1.3 Objectives

We aimed at assessing the added predictive value, as measured by the net reclassification index, of non-invasive markers of subclinical atherosclerosis when measured in addition to traditional risk prediction models in asymptomatic individuals with no history of CVD, i.e. in primary prevention (Table 2).

---

\(^b\) Macros/program files for calculating IDI and NRI using Stata, SAS, and R can be found at [http://www.ucr.uu.se/downloads](http://www.ucr.uu.se/downloads)
Table 2 – Research question: NRI

<table>
<thead>
<tr>
<th>P (patient)</th>
<th>Individuals with no history of cardiovascular disease/event</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (Intervention)</td>
<td>CVD risk assessed by the measurement of the carotid intima-media thickness (CIMT) OR coronary artery calcium (CAC) scores OR brachial flow-mediated dilation (FMD) OR ankle-brachial index (ABI) OR aortic pulse wave velocity (PWV) OR carotid plaque (CP) measured in addition to traditional risk prediction models Systematic COronary Risk Evaluation (SCORE) OR Framingham model</td>
</tr>
<tr>
<td>C (comparison)</td>
<td>CVD risk assessed by traditional risk prediction models (Systematic COronary Risk Evaluation (SCORE) OR Framingham model)</td>
</tr>
<tr>
<td>O (outcome)</td>
<td>Net reclassification improvement for cardiovascular disease/coronary heart disease</td>
</tr>
<tr>
<td>S (settings)</td>
<td>Any</td>
</tr>
<tr>
<td>S (study type)</td>
<td>Prediction modelling studies</td>
</tr>
</tbody>
</table>

For markers with a substantial NRI, we aimed at assessing the effectiveness (see chapter 3) and cost-effectiveness (see chapter 4) of using such markers in clinical settings. Reclassification based on the results of a marker should result in a clinical management different to what it would otherwise have been, and that is effective and cost-effective in reducing the risk for incident CVD. Such evidence should be generated by high-quality RCTs and sound economical evaluations.

We acknowledge that the evaluation of screening tools requires to also evaluate other parameters such as the acceptability by practitioners and patients, the feasibility, and the safety. However, effectiveness and cost-effectiveness are considered priorities, i.e. if they are not fulfilled all other criteria are futile.
2 INCREMENTAL PREDICTIVE VALUE

2.1 Methods

We screened electronic databases (MEDLINE, EMBASE, the Cochrane Library, and the CRD databases) to retrieve evidence on the added predicted value (NRI) of carotid intima-media thickness (CIMT), carotid plaques (CP), coronary artery calcium (CAC) scores, brachial flow-mediated dilation (FMD), ankle-brachial index (ABI), or aortic pulse wave velocity (PWV). The search strategies can be found in an Appendix. We also manually screened the references of the included papers. There was no restriction on language. As Pencina et al. published their ground-breaking paper on NRI in 2008, we searched for systematic reviews and original studies published from January 2008, up to September 2014.

As a first step, we searched for published systematic reviews. In case a recent good-quality systematic review existed, we extracted the results from that study, and searched for original studies published after the search date reported in that systematic review.

Any original study combining the 3 following criteria was considered eligible.

- Prospective study assessing the predictive increments of adding a marker (or combination of markers) to established risk prediction models based on conventional risk factors in asymptomatic individuals. Studies assessing combination of markers were eligible provided that the increment of individual markers was reported.
- Reporting the NRI, or presentation of data allowing the computation of NRI
- Reporting on prediction of first cardiovascular event (fatal or non-fatal) in the general population, i.e. reporting useful information in the field of primary CVD prevention. CVD was defined as coronary heart disease (fatal on non-fatal myocardial infarction) and stroke unless specified differently. CHD was defined as coronary death or myocardial infarction unless specified differently.

Exclusion criteria were defined as follows:

- Cross-sectional studies comparing risk classification by various prediction models
- Editorial or conference abstracts
- Studies carried out in symptomatic individuals
- Studies carried out in diabetic patients, as diabetes classifies automatically patients at high risk of CVD
- Studies carried out exclusively in no Western populations
- Studies reporting a crude NRI without 95% CI intervals and/or no test of statistical significance, and without the possibility for us to compute such statistics based on figures provided in the paper. Study population with small numbers of events can give very spurious results (if 2 over 4 events are reclassified correctly, the NRIevents=50%; for an illustration see the study by Berard et al.).

Whenever possible, we reported the NRI and the CNRI, and for each of these metrics we reported the proportion of correct reclassification in individuals developing CVD and in individuals remaining asymptomatic.

The quality appraisal of the systematic review was based on the AMSTAR grid whereas the quality appraisal of primary studies included was based on the NICE checklist for prognostic studies derived from Hayden et al.

Although the GRADE methodology for prognosis studies is not yet established, we rated the quality of the body of evidence for each marker (high, moderate, poor, very poor) based on the number of studies, the quality of individual studies, the consistency of results across studies, the size of the reclassification, and the imprecision around the point estimate.

This was a rapid systematic review. First, the search was focused by the inclusion in the search strategy of the outcome of interest, i.e. the reclassification statistics. Second, in case a recent good-quality systematic review existed, we extracted the results from that study, and searched for original studies published after the search date reported in that systematic review. Third, study selection was done by two reviewers, but data extraction and quality appraisal by only one. Fourth, we extracted from studies only parameters of interest.
2.2 Results
We retrieved 420 hits from MEDLINE, EMBASE, the Cochrane Library, and CRD databases. After removal of duplicates, 350 original references remained, which included 48 reviews and 303 original studies (Ferket 2014 included both a systematic review and an original study). 33 of the reviews were excluded on title and/or abstract mainly because they were not about primary prevention or because they were not genuine systematic reviews. Hence, 15 systematic reviews were eligible, constituting the pool from which we selected the most recent good-quality systematic review for each atherosclerosis marker. Selection of original studies to update the included systematic review is described for each atherosclerosis marker (Figure 1). The quality appraisal of included studies is presented in an appendix.

Figure 1 – Flow chart of evidence retrieval
2.2.1 **Flow-mediated dilation**

The most recent systematic review on the prognostic value of flow-mediated dilation was published by Peters et al. in 2012, with a search strategy until February 2011\textsuperscript{26, 27}. This review was rated moderate quality based on the AMSTAR checklist (see appendix), mainly because the review was restricted to Medline. However, our search strategy run in EMBASE, CRD and the Cochrane library, besides MEDLINE (see description above), did not detect any additional primary studies than the ones included in the systematic review by Peters et al. Therefore, we extracted the results of that review, and in a second step updated the review by searching for papers published since September 2011.

In the review by Peter et al., only one study reported the NRI of flow-mediated dilation\textsuperscript{28} (see Table 3). That study from the MESA cohort\textsuperscript{28} reported quite a substantial net reclassification of 29% of individuals (p<0.0001) when FMD was measured on top of FRS. However, this overall effect was mainly driven by shifting 52% of those with no incident events to a lower risk group at the expense of shifting 23% of those with an incident event to a lower risk group. The NRI in the intermediate risk group was 28% (p<0.0001): 38% of those at intermediate risk without an event were correctly reclassified to the low-risk category, but incorrect down-classification occurred in 10% of the subjects with incident CVD\textsuperscript{27}.

<table>
<thead>
<tr>
<th>Author</th>
<th>Cohort</th>
<th>N</th>
<th>Characteristics population</th>
<th>Risk factors in baseline model</th>
<th>Thresholds intermediate risk</th>
<th>Endpoint</th>
<th>Event rate</th>
<th>Follow up (years)</th>
<th>Prediction (years)</th>
<th>NRI% (95% CI)</th>
<th>CNRI% (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeboah 2009\textsuperscript{28}</td>
<td>MESA</td>
<td>3026</td>
<td>Age: 61 Men: 50%</td>
<td>Refitted FRS\textsuperscript{c}</td>
<td>10-20%</td>
<td>CVD\textsuperscript{d}</td>
<td>6.0</td>
<td>5</td>
<td>10</td>
<td>29\textsuperscript{e} (p&lt;0.0001)</td>
<td>28 (p&lt;0.0001)</td>
</tr>
<tr>
<td>Yeboah 2012\textsuperscript{4}</td>
<td>MESA</td>
<td>1330</td>
<td>Age: 64±9.5 Men: 62.7%</td>
<td>FRS</td>
<td>5-20%</td>
<td>CHD</td>
<td>7.1</td>
<td>7.6</td>
<td>10</td>
<td>NA</td>
<td>2.4\textsuperscript{g} (p=0.024)</td>
</tr>
</tbody>
</table>

\textsuperscript{c} Age, sex, diabetes mellitus, systolic blood pressure, use of blood pressure, medication, smoking, HDL cholesterol, LDL cholesterol, triglycerides, statin use, heart rate

\textsuperscript{d} Myocardial infarction, definite angina, coronary revascularization, resuscitated cardiac arrest, stroke, or CVD death

\textsuperscript{e} 23% of events were down-classified

\textsuperscript{f} Only intermediate-risk individuals

\textsuperscript{g} NRI\textsubscript{events}=0%; NRI\textsubscript{noevents}=2.4%

\textsuperscript{h} NRI\textsubscript{events}=-2.4%; NRI\textsubscript{noevents}=4.7%
Our search strategy yielded 6 references of studies published after September 2011, 5 of which were discarded on the basis of their title and/or abstract. Only one study fitted the inclusion criteria\(^4\). The results of this study were also derived from the MESA cohort, and focused on intermediate-risk individuals only. In that group, adding FMD to FRS resulted in a modest increase of NRI for either CHD or CVD. For CHD, 3.2% of non-events were reclassified, among which 2.4% were correct, and no event was reclassified. For CVD, 5.6% of non-events were reclassified, among which 4.7% were correct, and 2.4% of events were reclassified, none of which were correct. Therefore, the results of the 2 studies included in our review appeared contradictory, although the 2 studies referred to the same cohort. This may be due to different cut-offs limiting the intermediate risk group, different modelling, and a different definition of outcomes. In conclusion, the evidence on the benefit of FMD in addition to FRS is of very low quality (only 2 studies relating to the same cohort; inconsistencies in results). Moreover, the substantial incorrect down-classification of events reported by Yeboah 2009\(^{28}\) is worrisome.

2.2.2 Ankle-Brachial Index (ABI)

The most recent systematic review of evidence was published in September 2013 by Lin et al. with a search strategy up to September 2012\(^{14}\). A more extensive report of the review can be found on the website of the U.S. Preventive Services Task Force (http://www.guideline.gov/content.aspx?id=34783). This systematic review was rated of good quality (9/11 on the AMSTAR scale; a priori protocol not mentioned; assessment of publication bias not reported) (see appendix).

The four heterogeneous risk prediction studies included in Lin's review showed that the magnitude of the NRI was small when the ABI was added to the FRS to predict CHD or CVD events\(^4, 12, 29, 30\) (Table 4). The added prediction value was marginally greater in intermediate-risk individuals, but remained modest.

---

\(^{1}\) The most recent systematic review was published by Ferket et al. in 2014, but the authors did not analysed NRI in their meta-analysis\(^{25}\)
### Table 4 – Ankle-brachial index

<table>
<thead>
<tr>
<th>Author</th>
<th>Cohort</th>
<th>N</th>
<th>Characteristics population</th>
<th>Risk factors in baseline model</th>
<th>Endpoint</th>
<th>Event rate</th>
<th>Follow up (years)</th>
<th>Prediction (years)</th>
<th>NRI, (95% CI) %</th>
<th>CNRI, (95% CI) %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kavousi 2012</td>
<td>Rotterdam</td>
<td>5933</td>
<td>Age: 69.1</td>
<td>Refitted FRS</td>
<td>CHD</td>
<td>5.8</td>
<td>6.8</td>
<td>10</td>
<td>0.6 (-1.8; 2.9)</td>
<td>7.3 (2.9; 11.7)</td>
</tr>
<tr>
<td>Yeboah 2012</td>
<td>MESA</td>
<td>1330k</td>
<td>Age: 64±9.5</td>
<td>FRS</td>
<td>CHD</td>
<td>7.1</td>
<td>7.6</td>
<td>10</td>
<td>NA</td>
<td>3.6 (p=0.036)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Men: 62.7%</td>
<td></td>
<td>CVD</td>
<td>9.2</td>
<td></td>
<td></td>
<td></td>
<td>6.8 (p=0.068)</td>
</tr>
<tr>
<td>Rodondi 2010</td>
<td>Health ABC</td>
<td>2191</td>
<td>Age: 73.5</td>
<td>FRS</td>
<td>CHD</td>
<td>15.8</td>
<td>8.2</td>
<td>7.5</td>
<td>3.3 (0.04;6.5)</td>
<td>7.0 (2.9; 11.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Men: 44.7%</td>
<td></td>
<td>CVD</td>
<td>3.0</td>
<td>14</td>
<td>10</td>
<td>0.8 (p=0.05)</td>
<td>NR</td>
</tr>
<tr>
<td>Murphy 2012</td>
<td>ARIC</td>
<td>11594</td>
<td>Age: 53.8</td>
<td>FRS</td>
<td>CVD</td>
<td>15.8</td>
<td>8.2</td>
<td>7.5</td>
<td>3.3 (0.04;6.5)</td>
<td>7.0 (2.9; 11.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Men: 43.6%</td>
<td></td>
<td>CVD</td>
<td>3.0</td>
<td>14</td>
<td>10</td>
<td>0.8 (p=0.05)</td>
<td>NR</td>
</tr>
<tr>
<td>Fowkes 2014</td>
<td>18 prospective cohort studies</td>
<td>44752</td>
<td>Age: variable</td>
<td>FRS</td>
<td>CHDmen</td>
<td>variable</td>
<td>variable</td>
<td>variable</td>
<td>4.3 (0.0–7.6)</td>
<td>15.9 (6.1–20.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Men: variable</td>
<td></td>
<td>CHDwomen</td>
<td>variable</td>
<td>variable</td>
<td>variable</td>
<td>9.6 (6.1–16.4)</td>
<td>23.3 (13.8–62.5)</td>
</tr>
</tbody>
</table>

**Notes:**
- NRI in events 4.3%; NRI in non-events 2.6%
- In intermediate-risk individuals
- NRIevents=2.1%; NRInon-events =1.5%
- NRIevents=4.1%; NRInon-events=2.7%
- NRIevents=2.2%; NRInon-events=1.1%
- The improvement for non-events was 0.5%
- NRIevents =2.4% ; NRInon-events =1.5%
- NRIevents =16.4% ; NRInon-events =-5.1%
- NRIevents =5.8%; NRInon-events =7.6%
- NRIevents =25%; NRInon-events =13.2%
We updated that systematic review by searching for new evidence published between September 2012 and September 2014. The search returned 26 hits (13 from MEDLINE, 13 from EMBASE, 0 from Cochrane). After removal of duplicates, 18 references remained, 14 of which were discarded on title and/or abstract (including the above-mentioned systematic review and the corresponding AHRQ report), mainly because studies were not conducted in asymptomatic patients. Among the 4 remaining references, 3 were excluded on full-text with reasons presented in Table 5. Therefore, one primary study was added to the evidence base\(^1\). This study included 18 prospective studies amounting to 44,752 individuals, among which the ARIC study, the Rotterdam study, and the Health ABC study, i.e. 3 of the 4 studies included in the review by Lin et al. Details of this good quality study can be found in Appendix 2.4.

### Table 5 – Excluded ABI studies and reasons

<table>
<thead>
<tr>
<th>Reference</th>
<th>Reasons for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ninomiya 2013(^32)</td>
<td>Cut-offs used to delineate risk categories were unusually low (&lt;2.0%; 2.0%-4.0%; &gt;4.0%)</td>
</tr>
<tr>
<td>Matsushita 2013(^33)</td>
<td>Only a conference abstract was available with little details to appraise the quality and because the predicted outcome was a mixture of CHD, strokes, heart failure, and peripheral artery disease</td>
</tr>
<tr>
<td>Ferket 2014(^25)</td>
<td>Not a cohort study. Event rates adjusted for competing risks were obtained by microsimulation.</td>
</tr>
</tbody>
</table>

The NRI was negligible in most studies included by Lin et al\(^14\). On the contrary, the NRI was increased, particularly in women and in individuals at intermediate CVD risk, in the multi-cohort study by Fowkes et al\(^31\), although the latter included 3 of the studies also included in the review by Lin et al. The big sample size and the harmonization of cut-offs of risk categories may partly explain the different results reported by Fowkes et al. However, their results should be interpreted with caution as incorporation of the ABI in the prediction of major coronary events led to no significant when a better performing model based on fitting individual risk covariates was used instead of the FRS\(^t\). The authors of that study suggested that the impact of the ABI is not a fixed phenomenon but is influenced by how well the base risk factor model performs\(^u\). These results suggest that the impact of the ABI is not a fixed phenomenon but is influenced by how well the base risk factor model performs\(^31\). The difference of ABI performance by sex is difficult to explain. The authors of the study referred to chance or to a poorer performance of the FRS in women being compensated by the addition of ABI. It also appears that the results of that study are quite imprecise, with a 95%CI around CNRI going from 6.1% in men to 62.5 in women). Based on these elements, we conclude that the evidence that ABI may yield a substantial CNRI is moderate.

\(^t\) NRIs of 2.0% (95% CI: –2.3; 4.2%; p=0.567) in men and 1.1% (95% CI: –1.9; 4.0%; p=0.483) in women. In only those at intermediate 10–19% risk, NRIs were 7.7% (95%CI 0.0; 13.0%, p=0.049) in men and 2.4% (95% CI: –3.0;10.5%, p=0.275) in women

\(^u\) The corresponding author was contacted on 24 November 2014. Here is his response: “You are correct in that the results you quote are based on the external validation set. These are not based on the FRS but on a newly developed risk factor model and suggest that the ABI has minimal impact if added to a well-functioning risk factor model. Models generally work less well in external validation datasets as the characteristics of the population would be slightly different from those used in the development and internal validation datasets”.
Aortic pulse wave velocity

We retrieved 34 hits from the electronic database, none of which was a systematic review. Greenland at al. had reviewed the evidence on aPWV in 2010 and reported that among the 11 longitudinal studies retrieved, the information on incremental risk stratification had generally not been reported. 24 studies were discarded on title/abstract and 7 studies were discarded on full text (see Table 6 for reasons).

Table 6 – Excluded studies on aPWV and reasons

<table>
<thead>
<tr>
<th>Reference</th>
<th>Reasons for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rodondi 2010</td>
<td>No NRI reported for aPWV. aPWV increased the c-statistics in comparison with a predictive model based on traditional risk factors, but less than the ankle-arm index</td>
</tr>
<tr>
<td>Stea 2010</td>
<td>234 patients, cross-sectional study, conference abstract</td>
</tr>
<tr>
<td>Sehestedt</td>
<td>The NRI was computed for subclinical organ damage as defined by a mix of various measure: ambulatory blood pressure, pulse wave velocity, urine albumin/creatinine ratio, left ventricular mass index and carotid atherosclerotic plaques. The relative contribution of each of these markers is not reported.</td>
</tr>
<tr>
<td>Berard 2011</td>
<td>Conference abstract only (see Berard 2013)</td>
</tr>
<tr>
<td>Berard 2013</td>
<td>Different definitions of outcomes (coronary events vs. cardiovascular death) and different time horizon were used to assess reclassification “Using FRS: high, intermediate and low risks were defined as 10-year risk of coronary heart disease &gt;20%, 10–20% and &lt;10%, respectively. Using FRS+PWW: high, intermediate and low risks were defined as 14-year risk of cardiovascular mortality &gt;5%, 1–5% and &lt;1%”</td>
</tr>
<tr>
<td></td>
<td>• Missing data up to 13.9%</td>
</tr>
<tr>
<td></td>
<td>• Low number of events: 11 cardiovascular deaths</td>
</tr>
<tr>
<td>Kavousi 2012</td>
<td>Conference abstract only (see Kavousi 2012)</td>
</tr>
<tr>
<td>Pompilio 2013</td>
<td>Conference abstract only</td>
</tr>
</tbody>
</table>


We included 3 studies12, 40, 41. The study by Ben-Shlomo et al. comprised 14888 individuals from 16 cohorts, and was included in spite of a number of limitations: 10% of participants had experienced CVD events before the start of the observation period; the thresholds used to define risk categories are unusual (see Table 7), making the extrapolation of results to clinical settings difficult; there was no quality appraisal of the individual studies (see appendix). Pooling studies with very different predictor measurement may also be questionable, but the authors reported no evidence of heterogeneity for any of the outcomes40.

The NRI was modest for CVD and CHD events, but greater than 10 for CVD deaths. The CNRI was around 15% for CVD and CHD events. There was no evidence that the increased risk associated with aPWV was modified by sex, population type, smoking status, renal function, baseline diabetes, or antihypertensive use. However, aPWV was more strongly related to the risk of CHD and stroke in younger participants.

We retrieved 2 additional studies which had not been included in the meta-analysis by Ben-Shlomo et al12, 41. For details on studies, please see Table 7 and Appendix. Results were inconsistent, with non-significant NRI and CNRI in the Rotterdam study12 and high figures in a small study from Portugal. The higher NRI in the study by Pereira41 may be due to the small number of events and potential misclassifications.

In conclusion, given the predominant weight of the study by Ben Shlomo40 in the review, and the methodological limitations of that study (see above), we conclude that there is moderate quality evidence that aPWV could yield a substantial CNRI in routine clinical practice.


\[\text{We contacted the corresponding author to assess the possibility of secondary analysis on the unique population with no history of CVD and using more conventional thresholds to define risk categories (mail sent on January 12 2015). We received no answer so far.}\]
### Table 7 – aPWV

<table>
<thead>
<tr>
<th>Author</th>
<th>Cohort</th>
<th>N</th>
<th>Characteristics population</th>
<th>Risk factors in baseline model</th>
<th>Thresholds intermediate risk</th>
<th>Endpoint</th>
<th>Event rate</th>
<th>Follow up (years)</th>
<th>Predict (years)</th>
<th>NRI, % (95% CI)</th>
<th>CNRI (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ben-Shlomo 2014&lt;sup&gt;40&lt;/sup&gt;</td>
<td>16 individual cohorts</td>
<td>14888&lt;sup&gt;6&lt;/sup&gt;</td>
<td>No report, overall population</td>
<td>FRS+diabetes</td>
<td>Quartiles 2-3 of study population within individual studies</td>
<td>CVD variables</td>
<td>5</td>
<td>5</td>
<td>5.4 (4.4;6.4)&lt;sup&gt;x&lt;/sup&gt;</td>
<td>14.0 (11.4;16.5)</td>
<td></td>
</tr>
<tr>
<td>Kavousi 2012&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Rotterdam</td>
<td>3678</td>
<td>Age: 69±8.5, Men:41%</td>
<td>Refitted FRS</td>
<td>10-20%</td>
<td>CHD</td>
<td>5.8</td>
<td>6.8</td>
<td>12.2 (10.7;13.7)</td>
<td>27.2 (29.6; 37.7)</td>
<td></td>
</tr>
<tr>
<td>Pereira&lt;sup&gt;41&lt;/sup&gt; Portugal</td>
<td>1709</td>
<td>Age:51.7±10.5, Men: 43%</td>
<td>Refitted EuroScore</td>
<td>≤1%-&lt;5%</td>
<td>CVD&lt;sup&gt;y&lt;/sup&gt;</td>
<td>2.1 (n=47)</td>
<td>1.9</td>
<td>2</td>
<td>24.7&lt;sup&gt;z&lt;/sup&gt;</td>
<td>31.1&lt;sup&gt;aa&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>

<sup>w</sup> 10% had CVD events at baseline<br><sup>x</sup> NRI are also reported for 10-year risk prediction, and for stroke. There was no substantial differences in the NRI in the 10-year risk model compared with that using studies included in the 5-year risk model. For an absolute risk of 5% in 5-years the overall NRI was lower: 0.90% for CHD, 1.11% for cardiovascular events. For a threshold defined by the within-study median predicted risk for those experiencing an event, the overall NRI was 2.16% for CHD and 2.15% for cardiovascular events.<br><sup>y</sup> CVD defined as death, stroke, transient ischemic attack, myocardial infarction, unstable angina, peripheral arterial disease, revascularization or renal failure<br><sup>z</sup> Upward and downward reclassification with the addition of PWV to the standard model was seen respectively in 21.2% and 2.1% of participants with cardiovascular events, and in 7.8% and 13.4% of participants without cardiovascular events. The category-free NRI was also determined, being 0.265 (p <0.0001).<br><sup>aa</sup> 20% participants with events being upward reclassified with the inclusion of PWV
2.2.3 Coronary artery calcification score (CAC)

The most recent systematic review on the prognostic value of coronary artery calcification (CAC) was published by Peters et al. in 2012, with a search strategy until February 201126, 27. This review was rated moderate quality based on the AMSTAR checklist (see appendix), mainly because the review was restricted to Medline. However, our search strategy run in EMBASE, CRD and the Cochrane library, besides MEDLINE, did not detect any more additional primary studies than the ones included in the systematic review by Peters et al.

The review by Peters et al. included 9 studies on CAC prediction, 4 of which reported the NRI42-45 (Table 8).
<table>
<thead>
<tr>
<th>Author</th>
<th>Cohort</th>
<th>N</th>
<th>Characteristics population</th>
<th>Risk baseline factors in baseline model</th>
<th>Thresholds intermediate risk</th>
<th>Event rate</th>
<th>Follow-up (years)</th>
<th>Predict (years)</th>
<th>Endpoint</th>
<th>NRI, % (95% CI)</th>
<th>CNRI, % (95% CI)</th>
</tr>
</thead>
</table>
| Elias-Smale 2010<sup>45</sup> | Rotterdam | 2028 | Age: 70±6  
Men: 43% | Refitted FRS | 10-20% | 6.7 | 9.2 | 10 | CHD | 14.0 (p<0.01) | NR |
| Erbel 2010<sup>43</sup> | HNR | 4129 | Age: 59±8  
Men: 47% | FRS | 10-20% | 2.3 | 5.0 | 10 | CHD | 22.4 (p=0.0009) | 21.7<sup>ab</sup> (p=0.0002) |
| Mohlenkamp 2011<sup>42</sup> | HNR | 1934 | Age: 57±7  
Men: 31%  
No statin | Age, sex, TC/HDL ratio, antihypertensive medication | 3-10% | 2.2 | 5.1 | 5 | CVD<sup>cc</sup> | 25.1<sup>dd</sup> (p=0.01) | 43.6<sup>ee</sup> (p<0.001) |
| Polonski 2010<sup>44</sup> | MESA | 5878 | Age: 62±10  
Men: 46% | FRS, race/ethnicity | 3-10% | 3.6 | 5.8 | 5 | CHD<sup>ff</sup> | 25.0 (16.34)<sup>gg</sup> | 54.8<sup>hh</sup> (41.69) |
| Kavousi 2012<sup>12</sup> | Rotterdam | 3678 | Age: 69±8.5  
Men: 41% | Refitted FRS | 10-20% | 5.8 | 6.8 | 10 | CHD | 19.3 (12.5; 26.2) | 39.3<sup>i</sup> (26.8; 51.7) |
| Möhlenkamp 2011<sup>46</sup> | HNR | 3966 | Age: 59±8  
Men: 47% | FRS | 10-20% | 2.3 | 5.1 | 10 | CHD | 23.8 (p=0.0007) | 19.8 (p=0.003) |
| Yeboah 2012<sup>4</sup> | MESA | 1330<sup>jj</sup> | Age: 64±9.5  
Men: 62.7 | FRS | 5-20% | 7.1 | 9.2 | 10 | CHD | NA | 65.9<sup>kk</sup> |

HNR: Heinz Nixdorf Recall; MESA: Multi-Ethnic Study of Atherosclerosis; FRS: Framingham Risk Score; TC: total cholesterol; HDL: high density cholesterol; CACS: coronary artery calcification score

<sup>ab</sup> NRI events=5%; NRInoevents=16.6%
<sup>cc</sup> Primary endpoints for this study included fatal and non-fatal coronary events, cardiovascular mortality, stroke and coronary revascularization
<sup>dd</sup> NRI events=25.6%; NRInoevents=-0.5%
<sup>ee</sup> Computed by us
<sup>ff</sup> MI, CHD death, resuscitated cardiac arrest, definite or probable angina
<sup>gg</sup> NRI events=23%; NRInoevents=2%
<sup>hh</sup> NRI events=29%; NRInoevents=26%
<sup>ii</sup> NRI events=24%; NRInoevents=15.3%
<sup>ij</sup> Individuals at intermediate risk without diabetes mellitus
<sup>kk</sup> NRI events=25.5%; NRInoevents=40.4%
<sup>il</sup> NRI events=10.6%; NRInoevents=36.0%
As the search strategy of the principal systematic review was carried out until September 2011, we searched for primary studies published after that date.

We retrieved 54 hits. We excluded 44 studies on title/abstract, mainly because of cross-sectional design; studies carried out in symptomatic individuals; or testing of non-included tests or multi-biomarkers against CAC. Seven (7) references were excluded on full text, and the reasons for exclusion are presented in Table 9.

Table 9 – Excluded studies on CAC

<table>
<thead>
<tr>
<th>Reference</th>
<th>Reasons for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elias-Smale 201147</td>
<td>Same study population, and same results, as in the study by Elias-Smale 2010 (Rotterdam cohort) 45. NRI total=15%; NRI in events=18%; NRI in non-events=-3%). The study also investigates the added predictive value of CT scan of atherosclerosis in various vascular beds. Aortic arch calcium had a NRI=8%. Carotid calcium had a NRI=9%. CAC displayed the greatest additional predictive value.</td>
</tr>
<tr>
<td>Yeboah 201448</td>
<td>Same study population (MESA cohort) than in Polonski 2010 44 (5.8 years of follow-up) and in Yeboah 2012 4 (7.6 years follow-up; intermediate risk individuals), restricted to non-diabetics people. Only the NRI for CHD in the intermediate risk group is provided, with no confidence interval or statistical test. The study looks at the additional prediction by thoracic aorta calcium (TAC), aortic valve calcification (AVC), mitral annular calcification (MAC), pericardial adipose tissue volume (PAT), and a measure of liver attenuation (LA). The addition of CAC to the FRS had a clinical NRI of 0.547 for incident CHD in individuals at intermediate risk. However, the addition of TAC, AVC, MAC, PAT and LA individually to FRS + CAC resulted in clinical NRI of 0.0236, 0.0258, 0.0187, 0.0124, and 0.0116, respectively, for incident CHD. The addition of CAC to the FRS had a clinical NRI of 0.442 for incident CVD. The addition of TAC, AVC, MAC, PAT and LA individually to FRS+CAC resulted in a clinical NRI of 0.006, 0.030, 0.0130, 0.0037 and 0.0223 respectively for incident CVD</td>
</tr>
<tr>
<td>Nasir 201249</td>
<td>Conference abstract based on BioImage Study (study not described elsewhere?)</td>
</tr>
<tr>
<td>Matsushita 201333</td>
<td>Conference abstract based on MESA cohort, with comparison of individuals with and without chronic kidney disease</td>
</tr>
<tr>
<td>Kavousi 201250</td>
<td>Conference abstract based on data of the Rotterdam cohort presented in full in another paper by Kavousi et al. 12</td>
</tr>
<tr>
<td>Yeboah 201251</td>
<td>A conference abstract on the MESA study (see comments for Yeboah 2014 48). The NRI is reported with no confidence interval or statistical test.</td>
</tr>
</tbody>
</table>
The addition of CAC to the FRS resulted in an NRI of 0.191 for incident CVD and 0.229 for incident CHD.

<table>
<thead>
<tr>
<th>Study</th>
<th>Description</th>
<th>NRI for CHD</th>
<th>NRI for CVD</th>
<th>NRI for stroke/TIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gepner 201352</td>
<td>A conference abstract on the same study population (MESA cohort) than in Polonski 2010 44 (5.8 years of follow-up) and in Yeboah 2012 4 (7.6 years follow-up; intermediate risk individuals) with 8.5% years follow-up. The results differ from those reported by Polonski et al. but whether the same cut-offs of risk categories were used is not described. N=6 779</td>
<td>10.3 (5.2;15.5)</td>
<td>11.0 (6.0; 15.9)</td>
<td>2.8 (-1.2; 6.8)</td>
</tr>
</tbody>
</table>

We included 3 additional original studies4, 12, 46, thus in total 7 studies were included. For more details on studies and quality appraisal, please see the Appendix. However, the 7 papers refer in reality to 3 cohorts. It should be noted that one of the studies42 was carried out in a subgroup (people without indication for statin therapy) of another one43. It was also unclear why the study by Elias-Smale45 included only 2 028 individuals, when Kavousi reported 3 678 in the same Rotterdam study2.

Results were quite consistent across the 3 cohorts, with an NRI amounting to around 20% for CHD, and almost nearly doubled in the intermediate-risk individuals.

In conclusion, there is high-quality evidence that CAC could yield to a substantial NRI and CNRI if applied in routine clinical practice, particularly in individuals at intermediate risk.

2.2.4 Carotid intima-media thickness (cIMT)

Two systematic reviews were published in 2012 which complement each other neatly13, 27mm. Den Ruitjer et al. meta-analysed individual data from 14 population-based cohorts contributing 45 828 individuals13. Their population-based cohorts were retrieved through a search in MEDLINE and EMBASE up to June 2012, and by asking experts. Of the 31 eligible cohorts invited to participate, 14 were included (5 refused to participate or did not respond, 1 reported only maximum cIMT, and 11 were still pending at the time of publication). No details were provided on the non-participating cohorts. This meta-analysis was rated of high-quality with AMSTAR (see Appendix). The review by Peters et al. included 12 original published studies (research up to 7 September 2011 only in MEDLINE). This review was rated of moderate quality based on the AMSTAR checklist (see appendix), mainly because the review was restricted to Medline. Five of the included studies reported the NRI27. Only one of these did not contribute to the meta-analysis by Den Ruitjer: the Framingham cohort which included 2 965 individuals and reported similar results to those of the Den Ruitjer (NRI for Common Carotid Artery=0.0%; NRI for Internal Carotid Artery=7.6%)53

We therefore based our analysis on the results of the meta-analysis by Den Ruitjer et al. which summarized the majority of the existing evidence using uniform definitions of common CIMT, study population, risk categories, and cardiovascular events.

We also searched for primary studies published after the meta-analysis by Den Ruitjer et al (i.e. from 2012 onwards), and retrieved 61 studies published up to September 2014. 59 of them were excluded on title/abstracts. In particular, 14 studies were already included in the meta-analysis by Den Ruitjer (Table 10)

mm The most recent systematic review was published by Ferket et al. in 2014, but the authors did not analysed NRI in their meta-analysis25
Table 10 – Studies on cIMT excluded on abstracts

<table>
<thead>
<tr>
<th>Reference</th>
<th>Reasons for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kavousi 2012\textsuperscript{50} Rotterdam study already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>2</td>
<td>Elias-Smale 2012\textsuperscript{54} Rotterdam study already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>3</td>
<td>Leening 2013\textsuperscript{55} Rotterdam study already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>4</td>
<td>Zavodni 2014\textsuperscript{56} MESA study already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>5</td>
<td>Veeranna 2012\textsuperscript{57} MESA study already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>6</td>
<td>Matsushita 2013\textsuperscript{33} MESA study already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>7</td>
<td>Polak 2013\textsuperscript{58} MESA study already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>8</td>
<td>Yeboah 2012\textsuperscript{4} MESA study already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>9</td>
<td>Gepner 2013\textsuperscript{52} MESA study already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>10</td>
<td>Nambi 2012\textsuperscript{59} ARIC data already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>11</td>
<td>Ziegelbauer\textsuperscript{60} CAPS data already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>12</td>
<td>Gardin 2014\textsuperscript{61} CHS data already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>13</td>
<td>Gardin 2012\textsuperscript{62} CHS data already included in the meta-analysis of Den Ruitjer</td>
</tr>
<tr>
<td>14</td>
<td>Den Ruijter 2013\textsuperscript{63} Hypothetical cohort (simulation)</td>
</tr>
</tbody>
</table>

The 2 remaining studies were excluded on full-text. Nasir 2012\textsuperscript{49} because it was a conference abstract\textsuperscript{nn}. The study by Baldassare et al. based on the data of the IMPROVE study was also excluded because the NRI was provided for various ultrasonographic measurements of the carotid only in comparison with a model composed of FRS and c-IMT\textsuperscript{mean} (average of all mean IMT values)\textsuperscript{64}. Such comparison was beyond the scope of the present study.

\textsuperscript{nn} The NRI was computed for major adverse cardiovascular events (not defined) on 6,808 individuals followed up during a median time of 2.5 years NRI=0.2\% (95\%CI: -2.3\%; 2.8\%) p=0.85
Table 11 – cIMT

<table>
<thead>
<tr>
<th>Author</th>
<th>Cohort</th>
<th>Marker</th>
<th>N</th>
<th>Characteristics population</th>
<th>Risk factors in baseline model</th>
<th>Thresholds intermediate risk</th>
<th>Event rate</th>
<th>Follow-up years</th>
<th>Prediction years</th>
<th>Endpoint</th>
<th>NRI, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Den Ruijter 2012</td>
<td>14 cohorts</td>
<td>cIMT</td>
<td>45 828</td>
<td>Age: 58 (35-75) Men:47.4%</td>
<td>FRS</td>
<td>5%-20%</td>
<td>8.7</td>
<td>11</td>
<td>10</td>
<td>CVD</td>
<td>0.8 (0.1;1.6)</td>
</tr>
<tr>
<td>Polak 2011</td>
<td>Framingham</td>
<td>cIMTpp</td>
<td>2 965</td>
<td>Age: 58±10 Men:44.7%</td>
<td>FRS</td>
<td>6%-20%</td>
<td>10.0</td>
<td>7.2</td>
<td>10</td>
<td>CVDpp</td>
<td>0.076&lt;sup&gt;tt&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

This big meta-analysis (n=45 828) showed that the added value of common cIMT measurements to the Framingham Risk Score in the general population was small (0.8% correctly reclassified) (Table 11). In individuals at intermediate risk, the added value was 3.2% in men and 3.9% in women. The results of the study by Polak et al. 2011 were at odds with those of the meta-analysis, yielding a NRI of 7.6% for a 10-year risk prediction. One reason for this discrepancy may be the extended definition of CVD in the study by Polak which included peripheral arterial disease, angina, and transient ischemic attack. Although the meta-analysis by Ruijter et al. presented some limitations, there is high-quality evidence that cIMT does not improve substantially CVD risk prediction.

2.2.5 Carotid plaques

Two systematic reviews were retrieved. The review by Lee et al. was in fact a narrative review. Therefore we used the review by Peters et al. as the starting point of our review. This systematic review included 6 original studies on carotid plaques, among which 2 reported on the NRI. One of the studies was carried out exclusively in China and was discarded for that reason. The remaining study is reported in Table 13. We searched for
additional studies published after the date of the literature search carried out by Peters et al. We retrieved 106 references, of which 95 were discarded on title and abstract. Among the 11 remaining, 8 were excluded on full-text (see Table 12 for reasons).

Therefore, our review on the added predictive value of carotid plaque over traditional risk markers included 3 studies (Table 13). The NRI was very consistent across studies (around 7%), as well as the CNRI (around 17% for a 10-year risk prediction). Such consistency was surprising given the various ways of defining carotid plaques used in the 3 studies, as well as the differing outcomes. It is worth mentioning that in 2 of the studies, the base models included diabetes, which is a factor classifying de facto individuals at high CVD risk, and therefore potentially limiting the added value of measuring carotid plaques. Results were not imprecised. In conclusion, there is high quality evidence that the presence of carotid plaques improves the CVD risk prediction beyond FRS.

Table 12 – Excluded studies on Carotid Plaques

<table>
<thead>
<tr>
<th>Reference</th>
<th>Reasons for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berard 2013</td>
<td>Different definitions of outcomes (coronary events vs. cardiovascular death) and different time horizon were used to assess reclassification. Using FRS: high, intermediate and low risks were defined as 10-year risk of coronary heart disease &gt;20%, 10–20% and &lt;10%, respectively. Using FRS+PWW: high, intermediate and low risks were defined as 14-year risk of cardiovascular mortality &gt;5%, 1–5% and &lt;1%</td>
</tr>
<tr>
<td></td>
<td>Missing data up to 13.9%</td>
</tr>
<tr>
<td></td>
<td>Low number of events: 12 cardiovascular deaths</td>
</tr>
<tr>
<td>Nambi 2014</td>
<td>The study compares the added predictive values of measuring various carotid segments</td>
</tr>
<tr>
<td></td>
<td>The NRI for carotid plaques cannot be isolated, and the data was already included in the meta-analysis by Den Ruijter et al. on the predicted value of cIMT 13 (see chapter 3.5).</td>
</tr>
<tr>
<td>Gepner 2013</td>
<td>Conference abstract</td>
</tr>
<tr>
<td>Sehestedt 2012</td>
<td>The authors test the NRI of adding sub-clinical organ damage (SOD) to traditional risk factors, but the results for carotid plaques cannot be disentangled from a cluster of indicators of SOD. The data of this study was already included in the meta-analysis by Den Ruijter et al. on the predicted value of cIMT 13 (see chapter 3.5)</td>
</tr>
<tr>
<td>Gardin 2014</td>
<td>Results reporting specifically to carotid plaques were not presented. The data of this study was already included in the meta-analysis by Den Ruijter et al. on the predicted value of cIMT 13 (see chapter 3.5)</td>
</tr>
<tr>
<td>Baber 2014</td>
<td>Conference abstract</td>
</tr>
<tr>
<td>Postley 2013</td>
<td>Conference abstract</td>
</tr>
<tr>
<td>Plichart 2011</td>
<td>Four categories were used to compute the NRI, with unusual cut-offs (&lt;2.3%; 2.3-3.6; 3.6-6.3; &gt;6.3) for predicted 6-year risk. The proportion of net correctly reclassified individuals by the addition of carotid plaques was 8.1%(95%CI=6.4–9.9%; p &lt; 0.001) in those who did not experience a CHD event (n = 4852) and 5.6%(95%CI=−2.3–13.4%; p = 0.17) in those who did experience a CHD event (n = 198), yielding a significant global NRI of 13.7%(95%CI=5.6–21.7%; p &lt; 0.001)</td>
</tr>
<tr>
<td>Ziegelbauer 2012</td>
<td>Results reporting specifically to carotid plaques were not presented. The data of this study was already included in the meta-analysis by Den Ruijter et al. on the predicted value of cIMT 13 (see chapter 3.5)</td>
</tr>
<tr>
<td>Author</td>
<td>Cohort</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>Nambi</td>
<td>ARIC</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Polak</td>
<td>Framingham</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Polak</td>
<td>MESA</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**vv** Four categories of risk were used

**ww** The NRI were higher in women than in men

**xx** Plaques were defined as a maximum intima-media thickness>1.5mm in the internal carotid artery

**yy** coronary heart disease (i.e., a fatal coronary event, myocardial infarction, coronary insufficiency, or angina), a cerebrovascular event (i.e., ischemic stroke, hemorrhagic stroke, or transient ischemic attack), peripheral arterial disease (i.e., intermittent claudication), or heart failure

**zz** NRIevents=4.2% NRIinonevents=3.1%

**aaa** NRIevents=10.3% NRIinonevents=7.2% (computed by us on the basis of the reclassification table provided by Polak et al. in the appendices)

**bbb** Plaques were defined as mean of the maximum ICA IMT. Plaques were also defined with other criteria. The maximum ICA IMT is the indicator yielding the highest NRI
2.3 Discussion

High heterogeneity in study methods hampered the pooling of results. Main results of the review are reported in Table 14. The incremental prediction performance of FMD could not be assessed given the paucity of data and contradictory findings. The reclassification performance of cIMT was low, and this marker could unlikely serve as a useful CVD markers beyond the traditional risk factors. CAC score provided the best improvement in CVD stratification above the FRS. Improvements in CVD risk reclassification with ABI and CP were lower than for CAC, but still reclassified correctly at least 15% of the intermediate risk group. aPWV may be in the same range as ABI and CP, but the evidence was less strong.

Table 14 – NRI and CNRI from markers added to the Framingham model to predict a first cardiovascular event

<table>
<thead>
<tr>
<th>Marker</th>
<th>Studies</th>
<th>Cohorts</th>
<th>N</th>
<th>Quality of evidence</th>
<th>NRI% (95%CI)</th>
<th>CNRI% (95%CI)</th>
<th>Subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Flow-mediated dilation (FMD)</td>
<td>24, 28</td>
<td>1</td>
<td>3 026</td>
<td>Very low</td>
<td>NA (NA)</td>
<td>NA (NA)</td>
<td>-</td>
</tr>
<tr>
<td>2. Ankle-Brachial Index (ABI)</td>
<td>5, 12, 29-31</td>
<td>19</td>
<td>46 082</td>
<td>Moderate</td>
<td>4.3 (0.0; 7.6)</td>
<td>15.9 (6.1; 20.6)</td>
<td>Men</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.6 (6.1; 16.4)</td>
<td>23.3 (13.3; 62.5)</td>
<td></td>
</tr>
<tr>
<td>3. Aortic Pulse Wave Velocity (aPWV)</td>
<td>312, 40, 41</td>
<td>18</td>
<td>20 275</td>
<td>Moderate</td>
<td>4.9 (4.0; 5.9)</td>
<td>14.8 (12.4; 17.1)</td>
<td>Women</td>
</tr>
<tr>
<td>4. Coronary Artery Calcium (CAC)</td>
<td>74, 12, 42-46</td>
<td>3</td>
<td>13 685</td>
<td>High</td>
<td>14.0 (NR)</td>
<td>21.7 (NR)</td>
<td>Lowest value</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.0 (16; 34)</td>
<td>54.8 (41; 69)</td>
<td>Highest value</td>
</tr>
<tr>
<td>5. Carotid Intima-Media Thickness (cIMT)</td>
<td>213, 65</td>
<td>16</td>
<td>48 793</td>
<td>High</td>
<td>0.8 (0.1; 1.6)</td>
<td>3.6 (2.7; 4.6)</td>
<td>-</td>
</tr>
<tr>
<td>6. Carotid Plaques (CP)</td>
<td>365, 68, 72</td>
<td>3</td>
<td>22 924</td>
<td>High</td>
<td>7.7 (2.3; 11.4)</td>
<td>17.7 (10.9; 24.7)</td>
<td>-</td>
</tr>
</tbody>
</table>

Uncertainty around the estimates may partly result from the modelling. This point is well illustrated in the study by Fowkes et al. on ABI. In that study, when a better performing model based on fitting individual risk covariates was used instead of the FRS, incorporation of the ABI in the prediction of major coronary events led to no significant improvement, whereas the CNRI was substantial when added to FRS. Another source of uncertainty results from the heterogeneous composition of study populations. For instance, diabetes automatically puts the individuals at high-risk of CVD. However, this parameter was most of the time not integrated in risk prediction models. This point is further discussed in the general discussion (see chapter 5). Finally, most of the studies shared the same reporting weaknesses: lost-to-follow up, blinding of outcome assessment, and validation of the prediction model in an external cohort were seldom reported.

ccc NA: Not applicable. The 2 studies were performed on the same study population but report contradictory findings.

ddd Results presented are from the study by Fowkes 2014 which had a predominant weight (including 18 cohorts with 44 752 individuals)

eee Results presented are from the study by Ben-Shlomo et al. which had a predominant weight (16 cohorts, 14888)

fff Results presented are from the study by Den Ruitjer 2012 which had a predominant weight (14 cohorts with 45 828 individuals)

ggg Results presented are from the study by Nambi 2010 which had a predominant weight (13 145 individuals)
3 CLINICAL BENEFIT OF IMPROVED RISK PREDICTION

3.1 Objectives
We aimed at reviewing the evidence of clinical effectiveness of the 4 atherosclerosis markers displaying a substantial NRI in prediction studies (Table 15).

Table 15 – Research question on clinical effectiveness of atherosclerosis markers

<table>
<thead>
<tr>
<th>P (patient)</th>
<th>Individuals with no history of cardiovascular disease/event</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (Intervention)</td>
<td>CVD risk assessed by the measurement of coronary artery calcium (CAC) scores OR ankle-brachial index (ABI) OR aortic pulse wave velocity (PWV) OR carotid plaque (CP) measured in addition to traditional risk prediction models Systematic COronary Risk Evaluation (SCORE) OR Framingham model</td>
</tr>
<tr>
<td>C (comparison)</td>
<td>CVD risk assessed by traditional risk prediction models (Systematic COronary Risk Evaluation (SCORE) OR Framingham model)</td>
</tr>
<tr>
<td>O (outcome)</td>
<td>Cardiovascular disease/Coronary heart disease</td>
</tr>
<tr>
<td>S (settings)</td>
<td>Any</td>
</tr>
<tr>
<td>S (study type)</td>
<td>Randomized control trials</td>
</tr>
</tbody>
</table>

3.2 Methods
In the literature search performed in chapter 2, we retrieved no articles assessing the clinical benefit or harm of measuring the atherosclerosis markers reviewed. We re-run a new literature search to ensure that no important papers had been missed. The search strategies can be found in Appendix. Its main difference with the one run in chapter 2 was no reference to NRI (higher sensitivity) and a filter for RCT (higher specificity). We screened electronic databases (MEDLINE, EMBASE, the Cochrane Library, and the CRD databases) to retrieve evidence on the clinical benefit of adding coronary artery calcium (CAC) scores, ankle-brachial index (ABI), aortic pulse wave velocity (PWV), or carotid plaques (CP) to the measurement of the traditional CVD risk factors. We also manually screened the references of the included papers. There was no restriction on language or publication date.

Any original study combining the 2 following criteria was considered eligible:
- RCT study assessing benefit of adding a marker of atherosclerosis to the traditional risk markers to assess the individual risk score and to orientate the clinical management of asymptomatic individuals. Studies assessing combination of markers were eligible provided that the increment of individual markers was reported.
- Reporting on cardiovascular events (fatal or non-fatal) in the general population, i.e. reporting useful information in the field of primary CVD prevention. CVD was defined as coronary heart disease (fatal on non-fatal myocardial infarction) and stroke unless specified differently. CHD was defined as coronary death or myocardial infarction unless specified differently

Exclusion criteria were defined as follows:
- Observational studies
- Editorial or conference abstracts
3.3 Results
Our search yielded 136 hits after removal of duplicates. No study addressing the research question was retrieved.

3.4 Discussion
We retrieved no study on the clinical effectiveness of measuring atherosclerosis markers in addition to the traditional risk markers to assess the individual CVD risk and to orientate clinical management.

One RCT randomized 2137 volunteers to a private risk factor counseling session by a nurse practitioner based either on the FRS or the FRS+CAC. Subjects were informed that the presence of any calcium constituted evidence of atherosclerosis. Subjects were given 2 copies of their anonymized CAC scan report and were encouraged to share their results with their physician. The primary end point was 4-year change in coronary artery disease risk factors differences and in downstream medical resource utilization, i.e. no CVD or CHD outcomes. Although the authors reported a favorable change in systolic blood pressure, LDL-cholesterol and reduced waist circumference in the intervention group, these results should be considered with caution as the study presented high risk of bias (randomization process not described, no blinding of outcome assessment; lost-to-follow-up >10% and unbalanced among study groups).

Another large-scale (n=39 000) randomized trial to assess the clinical benefit and cost-effectiveness of a population-based screening for CVD either by SCORE or by CAC vs. no screening is currently under way. It is however not designed to assess the clinical benefit of measuring CAC in individuals classified at intermediated CVD risk by SCORE. The results will be available in 2019.

4 REVIEW OF EVIDENCE ON COST-EFFECTIVENESS

4.1 Methods

4.1.1 Search strategy
A systematic search for relevant publications was carried out with the consultation of electronic reference databases up to 18/12/2014. Medline (through OVID), EMBASE, Econlit (through OVID), NHSEED (CRD) and NHSHTA (CRD) were searched to retrieve primary full economic evaluations (studies comparing both costs and outcomes) and systematic reviews of economic evaluations (i.e. secondary economic evaluations) published in 2008 or after. This was a pragmatic approach aimed at focusing on those models based on clinical studies already considering NRI data. Nevertheless, in order to avoid missing other potentially informative economic evaluations, the references of all economic evaluations found via our search were checked and any full economic evaluations identified via those references were reviewed for completeness.

A hierarchical approach was followed:
- First, the analysis focused on recently published systematic reviews (SR) of economic studies published from 2008 to the 18th of December 2014.
- Second, the references of those studies identified via our search were checked and any relevant economic evaluations were included in our review.

An overview of the search strategy is given as an Appendix.

Furthermore, the websites of Health Technology Assessment (HTA) institutes listed on the INAHTA website (International Network of Agencies for Health Technology Assessment) and NICE (National Institute for Health and Care Excellence) were consulted to capture reports specifically focusing on those markers of subclinical atherosclerosis which showed to have a substantial NRI according to our clinical review (see section 2 of this report), (i.e. Ankle Brachial Index – ABI; Aortic Pulse Wave Velocity – aPWV; Carotid plaques and Coronary Artery Calcification – CAC). No restrictions were imposed for language.
4.1.2 Selection procedure

To identify potentially relevant studies for our analysis we first went through all titles and abstracts in order to exclude any obvious studies that did not match our research subject. All articles that appeared to be interesting, or for which there were some doubts, were read in full in order to select those relevant for inclusion in our review.

Study selection was completed by one researcher but any doubts that came up during the exercise were discussed and solved in collaboration with a second reviewer.

All studies finally included in our review were critically appraised by using an in-house structured data extraction sheet based on the check list originally developed by Drummond et al.75

4.1.3 Selection criteria

All full economic evaluations looking at ABI, aPWV, CAC or Carotid plaques, in addition to traditional prognostic tools, for facilitating an accurate identification of patients at risk of cardiovascular events, who could then benefit from preventive measures/medications, were considered for inclusion in our review. From those, only evaluations capturing their results as costs per quality adjusted life years (QALY) or costs per life years gained (LYG) were finally included.

Cost descriptive analyses or cost comparisons not taking into consideration effectiveness were discarded. Finally, publications in the form of letters, editorials or notes and abstracts were excluded, since these would not offer enough information to include them in our analysis and critically appraise their findings. An overview of the inclusion/exclusion criteria is given in Table 16.

<table>
<thead>
<tr>
<th>Selection criteria</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Individuals with no history of cardiovascular disease/events</td>
<td>Patients with a history of cardiovascular disease/events or suffering from comorbidities or diseases likely to significantly increase the risk of such events (eg diabetic or ESRD patients)</td>
</tr>
<tr>
<td>Intervention</td>
<td>ABI, aPWV – CAC and Carotid Plaques, measured in addition to other traditional tools (eg SCORE or Framingham)</td>
<td>Other markers</td>
</tr>
<tr>
<td>Comparator</td>
<td>Cardiovascular risk assessed by traditional tools (eg SCORE or Framingham)</td>
<td>No prognostic tool excluded</td>
</tr>
<tr>
<td>Outcome</td>
<td>LYG, or QALY</td>
<td>Number of cardiovascular events or other outcomes</td>
</tr>
<tr>
<td>Design</td>
<td>Full economic evaluations (primary or secondary)</td>
<td>Cost descriptive analysis, cost comparisons</td>
</tr>
<tr>
<td>Type of publication</td>
<td>Articles or reviews</td>
<td>Letters, editorials, notes, abstracts</td>
</tr>
</tbody>
</table>
Our search returned 980 citations, after eliminating duplicates. Of those, 957 did not meet our inclusion criteria based on a review of their title and/or abstract. In addition to the 23 citations left, which were read in full, 10 further potential references were identified by reference checking. From these 33 studies, 27 were excluded, primarily due to their design, which left us with 6 studies, 5 primary economic evaluations and 1 systematic review. The latter was purely used to further explore other potentially relevant references, but identified no additional full economic evaluations meeting our inclusion criteria.

Our literature selection process is illustrated in a flow chart in appendix 4.

4.2 Overview of economic evaluations

As shown in Table 17, all studies identified focus on CAC and were undertaken in the USA.

Only two of the studies were published after 2008 (Pletcher 2014, van Kempen 2011), and used as their input recent clinical studies providing net reclassification index (NRI) data. All studies were model-based (decision-tree and/or Markov models).

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>Type of analysis</th>
<th>Perspective</th>
<th>Time horizon (in years)</th>
<th>Discount rate; both costs and outcomes (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pletcher77</td>
<td>2014</td>
<td>USA</td>
<td>CUA</td>
<td>Third party payer</td>
<td>Lifetime</td>
<td>3%</td>
</tr>
<tr>
<td>Van Kempen50</td>
<td>2011</td>
<td>USA</td>
<td>CUA</td>
<td>Societal</td>
<td>Lifetime</td>
<td>3%</td>
</tr>
<tr>
<td>Taylor79</td>
<td>2005</td>
<td>USA</td>
<td>CUA</td>
<td>Third party payer</td>
<td>Lifetime</td>
<td>NA</td>
</tr>
<tr>
<td>O’Malley76</td>
<td>2004</td>
<td>USA</td>
<td>CUA</td>
<td>Third party payer</td>
<td>Lifetime</td>
<td>NA</td>
</tr>
<tr>
<td>Shaw78</td>
<td>2003</td>
<td>USA</td>
<td>CEA</td>
<td>NA</td>
<td>Lifetime</td>
<td>NA</td>
</tr>
</tbody>
</table>

CEA: Cost-effectiveness analysis; CUA: cost-utility analysis; NA: Not available/reported
4.2.1 Type of economic evaluation

Four of the studies performed cost-utility analyses\(^7\) expressing their clinical outcomes in QALYs. The remaining\(^8\) carried out a cost-effectiveness analysis with results reported in cost per LYG.

4.2.2 Time frame of analyses and discounting

All studies included in this analysis modelled costs and outcomes over a patient’s lifetime.

Out of the five studies, only the two most recent cost utility studies\(^7\),\(^8\) discounted costs and outcomes and gave information on the rates used. Three percent was used in both cases for both costs and outcomes. This is a frequent choice for economic evaluations undertaken in the USA\(^8\),\(^9\).

Three studies were performed from a third party payer perspective,\(^7\),\(^8\),\(^7\) while van Kempen et al.\(^8\) chose a broader perspective taking into consideration travel and patient time costs. Shaw et al.\(^8\) did not specify the perspective followed in their analysis.

All the studies that gave detailed information about the costs, with the only exception of van Kempen et al.\(^8\), used as inputs for their analyses direct medical costs only, defined as any costs falling within the health care system. Van Kempen et al. mentioned travel and patients time costs, but no further details or references were provided in this regard.

4.2.3 Population

Only one of the studies identified focused purely on asymptomatic adults\(^7\). Given the specific interest of the intermediate risk population according to conventional risk assessment tools, as a potential target population for the markers here studied, another study focused specifically in the asymptomatic elderly population at intermediate risk. In order to simplify their calculations Pletcher et al.\(^8\) modelled in their base case scenario women aged 55 with high cholesterol levels and no other CHD risk factors. Alternative scenarios varying sex, age or overall risks were also analysed.

Finally, both Taylor et al.\(^7\) and O’Malley et al.\(^7\) modelled healthy active-duty army personnel aged 40 - 55.

4.2.4 Intervention and comparator

All studies evaluated the cost-effectiveness of CAC screening. The comparator used varied depending on the study, with the most recent studies\(^7\),\(^8\) comparing amongst other strategies a “treat all” with statins approach for those “at risk”, according to conventional methods of risk stratification, to a more targeted approach in which CAC was used to facilitate a more accurate identification of those individuals “at risk” of CV events.

Less recent studies\(^7\),\(^8\),\(^7\) compared a treatment-decision strategy based on the risk classification of individuals via conventional tools versus an alternative strategy consisting of adding CAC to conventional tools. Only one study\(^7\), analysed different thresholds of CAC levels for treatment decisions and their potential effects on patients’ outcomes and costs.

4.2.5 Cost and outcome inputs

Costs were derived from the published literature or from reimbursement codes. With regard to outcomes, Shaw et al.\(^8\) captured theirs from a series of 676 asymptomatic patients and both Taylor and O’Malley used data from the PACC study\(^7\), although the former used a more complete set of data. More interesting are the studies published after 2008\(^7\),\(^8\) which used data from either the MESA study\(^8\) or the Rotterdam study\(^12\), already focusing on NRI as their main outcome.

Quality of life (QoL) is an important factor to bear in mind when studying prevention, giving the implications that being on long-term medication can have, as well as the potential QoL detriments that individuals may experience following non-fatal cardiovascular events. Utility values used in the four cost-utility studies here included\(^7\),\(^7\),\(^7\),\(^8\) were taken from the published literature with van Kempen et al.\(^8\) quoting a relatively recent study\(^8\), using the EQ-5D tool, as their main source of utility data.
4.2.6 Modelling

Four of the evaluations using decision analytic modelling gave enough detail to be able to offer an insight into their model structure and assumptions, while Shaw et al.78 failed to describe in enough detail their approaches. From the four evaluations offering enough transparency the two most recent77, 80 used Markov models illustrating multiple health states, moving from good health, through first and subsequent years after different cardiovascular events, and death. Both models used a life cycle of one year.

Pletcher et al.77, modelled the following interventions for a 55-year old woman with a mean 10-year Framingham risk score of angina, MI or CHD death of 7.5%:
1. Treat none with statins
2. Treat all with statins
3. Perform CAC screening and treat with statins if CAC>0
4. Perform CAC screening and treat with statins if CAC>100
5. Perform CAC screening and treat with statins if CAC>300.

In the case of van Kempen80, the strategies compared for intermediate-risk individuals were:
1. “Current practice”, described as the incidence of coronary heart disease (CHD) and non CHD events of individuals at intermediate-risk without any additional preventive measures, as observed in the Rotterdam study (some patients treated with statins and/or anti-hypertensives).
2. “Current guidelines”, reflecting current guidelines (i.e. lifestyle advice to all intermediate-risk patients, then statins if baseline LDL cholesterol >130mg/dl and antihypertensives if baseline systolic blood pressure >140mm Hg.)
3. “CAC screening”, followed by treatment with statins and anti-hypertensives for all of those re-classified via the combination of CAC scores and Framingham scores as having a high-risk. Treatment reflecting “current guidelines” (see bullet 2), for those remaining in the intermediate-risk group; and lifestyle advice for those reclassified as having low-risk followed by pharmacological treatment only if systolic blood pressure >140mm Hg and/or plasma LDL levels >160mg/dl.

4. “Statin therapy”, in which everyone at an intermediate risk and not on statins would receive a moderate dose of statins.

The model used by Taylor et al.79 is an update of the one presented by O’Malley et al.76, using a more complete set of data from the PACC study (n=2000 in Taylor’s et al. analysis versus n=1000 in O’Malley’s). The authors compared two interventions, one in which patients were classified by means of the Framingham score only or an alternative one in which CAC screening was added to Framingham scores. This was also the approach followed by Shaw et al.78

4.2.7 Results

Given the particular relevance of the two most recent cost-utility analyses77, 80, having made use of NRI data, we will start by describing in more detail their results and then comment on a more general basis the results from two of the other studies78, 79. The study of O’Malley would not be covered given that Taylor et al. offered an update of it with a more complete set of data from the PACC study, and thus the study from the latter should be more informative.

4.2.7.1 Costs

Costs from the study by Pletcher et al.77 and those of van Kempen et al.80 are difficult to compare because of their different approaches, with van Kempen considering a broader perspective, including travel and patient time costs and focusing specifically on intermediate-risk patients. The overall mean lifetime costs found by van Kempen et al.80 for their modelled population of men went from a low of US$7 551, for the most economical “current practice” strategy to a high of US$12 228 for the “CAC screening” strategy. Higher costs were found for women, for whom the most expensive strategy was not “CAC screening” but rather the “current guidelines” approach. The mean lifetime costs for women went from a low of US$8 553 for “current practice”, to a high of US$13 514 for “current guidelines”. “CAC screening” for women resulted in an overall costs of US$13 216.

Pletcher et al.77 presented their baseline calculations for a 55-year old women with high cholesterol and no other risk factors under two different sets of assumptions for the statins. First, a scenario with “favourable assumptions”, in which the cost of statins was US$ 0.13/pill and no disutility
was considered from taking this medication over the long-term. Under this scenario, the total cost of the different strategies for a hypothetical cohort of 10,000 women were very similar, with “treat none” or “treat all” patients with statins, or “treat if CAC > 100” costing less than other more conservative CAC strategies (i.e., “treat if CAC > 100” and “treat if CAC > 300”). Under less favourable conditions for the statins: US$1/pill and a disutility of 0.00384, the overall costs varied making the “treat all” with statins a less economically attractive option, while all others presented very similar overall costs.

Similar results were obtained for men. Shaw et al. reported mean lifetime costs of approximately US$1,863 for CAC screening versus US$1,347 (p<0.0001) for the alternative office-based risk assessment, based on Framingham scores.

Taylor et al. did not report mean cost estimates for the different approaches analysed.

With regard to the direct cost of the CAC test, the most recent studies quoted similar costs of US$105 (US$ of 2010) and US$114 (US$ of 2012) for van Kempen et al. and Fletcher et al. respectively. The evaluation by Taylor et al. from 2005 reported a noticeably higher cost estimate of US$400. No explicit mention of test costs was made in the publication by Shaw et al.

4.2.7.2 Outcomes

With regard to outcomes, the study of van Kempen et al. reported QALYs from 10.03 for “current practice” to 10.16 for “CAC screening” for men. “Statin therapy” lied in between with 10.12 QALYs, while “current guidelines” offered 10.14 QALYs.

The same model run for women reported QALYs from 9.26 for “current practice” to 9.41 for “current guidelines”. “Statin therapy” offered 9.36 QALYs and “CAT screening” 9.39.

The results from Fletcher et al. under the “favourable assumptions” scenario for the statins previously mentioned, showed that the approaches of “treat all” with statins or “treating if CAC > 0” were the most attractive offering mean QALY gains of 17.11 and 17.07 respectively for a theoretical 55-year-old women with high cholesterol levels. These figures decreased slightly when applying the disutility of long-term statin treatment under the “less favourable assumptions” scenario, but the overall picture remained the same with “treat all” offering most QALYs, followed by “treat if CAC > 100”. “Treat if CAC > 300” practically offered no gains compared to the strategy “treat none”. Similar results were obtained for men.

Shaw et al. reported incremental LYG for implementing CAC screening in addition to Framingham scoring, versus office-based Framingham risk assessment alone of 0.7 for low risk individuals (<0.6% annual risk of death of myocardial infarction) and of 11.3 for intermediate risk individuals (0.6%-2% annual risk of death of myocardial infarction).

No gains in QALYs were explicitly reported in the remaining study.

4.2.7.3 Incremental cost-effectiveness ratios (ICERs)

Van Kempen et al. reported an ICER in men for “statin therapy” compared to “current practice” of US$30,278/QALY and an ICER for “CAC screening” compared to “statin therapy” of US$48,800/QALY. “Current guidelines” was in this case dominated by extended dominance since, “CAC screening” yielded higher effectiveness at a lower incremental cost-effectiveness ratio. In women, the ICER for “statin therapy” versus “current practice” was US$23,910/QALY, while that of making treatment decisions by means of “current guidelines” offered an ICER of US$51,400/QALY when compared to “statins therapy”. This time, it was “CAC screening” that was dominated by extended dominance.

Probabilistic sensitivity analyses showed robust results in women (“CAC screening” highly unlikely to be cost-effective at willingness-to-pay thresholds of US$50,000), but uncertain in men with a probability of “CAC screening” captures situations in which a combination of two options shows greater cost-effectiveness than an alternative, which is said to be dominated by “extended dominance”.

---

In incremental cost effectiveness analysis alternatives that are less effective and more expensive are said to be “dominated” (strong dominance). There is however also the concept of “extended dominance” (weak dominance), which...
screening” being cost-effective for intermediate-risk individuals of less than 50% at a willingness-to-pay of US$50 000.

Pletcher et al.\(^{77}\) reported an ICER for their “treat all” scenario under “favourable assumptions for statins” versus the next cheaper non dominated strategy “treat if CAC>0” of US$100/QALY. Under such scenario, the price of statins is so low that their overall cost is practically offset by the benefits they offer, representing the most attractive therapeutic approach.

Under “less favourable assumptions” for the statins, the option of “treating if CAC>0” displays an ICER versus the “treat none” option of US$18 000/QALY, while “treating all” produced more QALYs when compared to “treat only if CAC>0”, but at a higher cost with an ICER of US$78 000/QALY. The “treat if CAC>100” and “treat if CAC>300” options were dominated under both scenarios.

They study by Taylor et al.\(^{79}\) gave an overall ICER of US$37 500 per QALY for adding CAC screening to Framingham versus implementing Framingham scoring alone in healthy individuals. However, these results were highly dependent on the efficacy of primary prevention.

Finally, the cost-effectiveness study of Shaw et al.\(^{78}\) showed ICERs sensitive to the population tested, with CAC screening in addition to Framingham risk assessment in low-risk patients (annual risk <0.6%) approaching ICERs of US$500 000 per LYG when compared to Framingham risk assessment alone, while this ICER fell to US$42 339 per LYG for patients with an annual coronary disease risk of 1% and to US$30 742 per LYG for those with an annual risk of 2% per year.

4.2.8 Sensitivity analysis

Uncertainty is intrinsic to any economic evaluations and should therefore always be accounted for. The two most relevant evaluations found via our review performed both one-way and probabilistic sensitivity analyses and showed their results to be highly uncertain and sensitive in particular to changes in the efficacy of primary prevention therapies\(^{80}\), variations on the utility of statins\(^{77}\), treatment adherence\(^{80}\) or statin costs (generic versus brand)\(^{77,80}\).

Older evaluations\(^{78,79}\) did not offer enough detail on the sensitivity analyses undertaken. Nevertheless, Shaw et al. concluded that their results were highly sensitive to the population screened, while Taylor et al. reported results sensitive to the efficacy of medication for primary prevention.

4.2.9 Conflict of interest

None of the relevant studies here included, reported a conflict of interest.

4.2.10 Discussion

The evidence published up to date on the cost-effectiveness of markers appear to be limited to CAC and present highly uncertain results for the general asymptomatic population. There are some important points worthwhile considering:

**Sources of clinical data**

Only two full economic evaluations up to date have used data on the NRI in their models. From those, one used the MESA study while the other made use of the Rotterdam study.

**Model assumptions**

In order to model the overall cost-effectiveness of the risk markers under evaluation (i.e. CAC), as potential prognostic tools that could help preventing future cardiovascular events in the general population, the authors of these evaluations had to make an important number of assumptions, in particular to do with the consequences of treating patients in response to CAC scores. Given the current lack of studies looking at the long-term consequences for the patient of making preventive treatment decisions on the basis of CAC, the two most recent studies reviewed\(^{77,80}\) assumed that statins would be given to patients at risk (in the case of van Kempen aspirin was also included as a treatment option under specific circumstances) and that their efficacy in patients with positive CAC scores in preventing cardiovascular events would be similar to their efficacy in more general, low or intermediate-risk populations (relative risk of myocardial infarction with statins versus placebo of 0.70 and 0.74 used by van Kempen and Pletcher respectively).
However, given that the only RCT to date assessing the effect of statin treatment on clinical outcomes specifically in patients with significantly elevated CAC scores\textsuperscript{84} reported no significant results, there is some uncertainty surrounding the accuracy of the estimations applied in these models in this regard.

Both studies included other important assumptions for which it was not always possible to back them-up with appropriate evidence. For example, the adherence to statins is an important factor, which could change the overall results. This was assumed to be 70\% of that of the original trials of statins in the case of van Kempen et al.\textsuperscript{80} but whether that would represent a fair estimate is unknown. When this specific assumption was tested during the sensitivity analysis the overall results changed, proving once more, the fragility of their findings.

### 4.2.11 Conclusion

In conclusion too many data gaps exist to be able to draw clear conclusions regarding the cost-effectiveness of CAC in the general population as a potential risk marker which could offer some additional value over that of current suboptimal risk classification systems.

These, coupled with the potential radiation risks linked to CAC screening, and the impossibility of performing such screening test in primary care, where most of the general population is followed by their general practitioner, makes CAC a far-from-optimal marker, for a broad use.

### 5 DISCUSSION

CVD remains the major cause of premature deaths in our country, and its primary prevention is a key public health objective. Targeting high-risk individuals is one possible strategy to achieve this objective. There is therefore considerable interest in improving the assessment of individual risk of CVD by adding new markers to the traditional risk factors. Markers of sub-clinical atherosclerosis could improve the predicted value of traditional prediction models. The most useful metrics to assess such improvement is the net reclassification index (NRI). The NRI is particularly informative in individuals classified at intermediate risk by traditional models, as it is in this group that the recommendations of clinical management are usually ill-defined.

We reviewed the evidence on the NRI of 6 markers of atherosclerosis. The performance of FMD could not be assessed given the paucity of data and contradictory findings. The reclassification performance of cIMT were poor, and this marker could unlikely serve as a useful CVD markers beyond the traditional risk factors. CAC score provided the best improvement in CVD stratification above the FRS\textsuperscript{44}. This was very clear in two good-quality studies which carried out intra-cohort comparison of the markers within a same cohort\textsuperscript{4, 12}.

Improvements in CVD risk reclassification with ABI and CP were lower than for CAC, but still reclassify correctly around 15\% of individuals in the intermediate risk group. aPWV may be in the same range as ABI and CP, but the evidence was less strong, as the meta-analysis by Ben-Shlomo et al. included 10\% of symptomatic individuals and the cut-offs to define risk categories were unusual \textsuperscript{40}.

There is currently no agreed cut-off of NRI above which a marker should be recommended. Such a recommendation should take into account other important criteria for an appropriate screening, such as cost, potential harm and easy use. For example, ABI is low cost, un-harmful and could be measured easily in general practice. Therefore, we reviewed the evidence on the clinical effectiveness cost-effectiveness of all 4 potential CVD markers previously mentioned for which there appears to be some benefit in terms of NRI.
We retrieved no study designed to assess the clinical benefit of refining the individual CVD risk by the measurement of one of the above-mentioned atherosclerosis markers. Cost-effectiveness studies were available only on CAC. Cost-effectiveness results were extremely unstable because of the many assumptions underlying the models which were not well backed-up with good quality evidence. The resulting uncertainty hampered the formulation of any operational recommendations and was a direct consequence of the lack of long-term clinical data on patient outcomes after reclassification and consequent treatment. No study has formally assessed whether the measurement of CAC, or another CVD marker, on top of the FRS or SCORE would result in better clinical management, including a better use of downstream testing and treatment, and in better health outcomes.

Such a dearth of good quality evaluation probably relates to the common belief that refining individual CVD risk will lead to a better clinical management through two main mechanisms. First, it is assumed that current preventative therapies, such as statins or aspirin, will be effective to reduce the additional CVD risk associated with the presence of a marker of atherosclerosis. Whether this assumption is correct is still a question mark, except for carotid plaques. For example, the majority of existing drugs do not seem to lower aPWV in a blood pressure–independent manner. The benefits of long-term blockade of the renin-angiotensin system and novel agents targeting elastic fiber cross-linking or calcification remain speculative. Provision of aspirin to people with a low ABI was not proven successful. Even for CAC, which yields the highest NRI and is the most studied atherosclerosis marker, good quality prospective studies on the clinical benefit are extremely rare. Statins do not seem to have an impact on CAC. The only RCT to date assessing the effect of statin treatment on clinical outcomes in patients with significantly elevated CAC scores reported no significant effect. Although this negative result might be due to the limited sample size, it underlines the need of large-scale outcomes trials to better define the role of medical treatment in individuals reclassified at high CVD risk. Such trials should also evaluate the potential harms associated with the CVD risk assessment, e.g. the potential risk attributable to radiation exposure associated with CAC scoring. Similarly, the outcomes of conservative approaches in individuals reclassified at low-risk should be properly assessed.

Second, measuring atherosclerosis markers could be a motivational tool for positive behavioural change, risk perception, and medication adherence. A recent systematic review suggested that CAC enhances medication utilization and adherence, with mixed results in the other domains. This review included 15 studies, only 3 of which were RCT, and the overall quality of evidence was low (notably the main outcome was self-reported in most studies). At any rate, whether using a prognosis test such as CAC to improve medication adherence is ethically justified is a debatable question.

Our review put in evidence another major shortcoming related to the design of the studies on NRI. It is well acknowledged that the predictive value of traditional models is modified by other parameters such as central obesity, parental history of premature CVD, sedentary, or social deprivation. For example, a parental history of premature CVD doubles the CVD risk obtained by SCORE. The BMI is associated with the presence of coronary plaques. The European Guidelines on CVD prevention recommends to take these elements into account when using SCORE for assessing the individual CVD risk of an individual. Unfortunately, SCORE does not integrate formally these elements, nor did the base model of the studies included in our review. What would be the NRI of atherosclerosis risk markers if this would have been the case is thus unknown. Studies assessing the predictive value of models integrating all the CVD risk factors which can be easily measured, either by the anamnesis or a simple clinical examination, and a revision of SCORE are urgently needed. Such studies should be based on new cohorts for two reasons. The first one is obvious: the information required to upgrade SCORE was not collected in past cohorts. The second reason is that the historical cohorts are not adapted for such an upgrade: unmeasured risk factors may have changed over time, such as salt and trans-fat intakes, and modern therapy (e.g. statins) may change the quantitative relationship between risk factors and CVD outcomes.

Without such evidence, i.e. the NRI of novel CVD markers when the base model integrate additional clinical information and the risk-benefit balance of adding novel CVD markers for measuring individual CVD risk, we cannot recommend the utilization of CVD markers in clinical practice. Meanwhile, first-line recommendation in asymptomatic subjects remains a healthy lifestyle including smoking cessation, regular physical activity, weight control, and a healthy diet, with medical treatment for controlling
hypertension, high blood cholesterol and diabetes when appropriate. Clinical skills remain central to adapt risk evaluation and management according to each individual situation.

Our review presented a number of limitations. We reviewed only a limited set of atherosclerosis markers and some other tests were not considered, e.g. cardiac stress test or retinal vessel caliber. However, we selected the most promising tests, and evidence on the other tests was often marginal as showed through an initial scoping review\textsuperscript{11}. Second, it was a rapid systematic review. However, study identification was done by 2 independent reviewers, one of whom was an information specialist applying a more extensive search strategy, and we are confident that the risk of selection bias is low. A meta-analysis was not possible because of the many methodological differences of included studies. The same reasons hampered direct comparison among CVD markers except in some studies comparing the CVD markers within the same cohort \textsuperscript{4, 12}. Finally, the generalizability of our findings to the Belgian context is limited as the vast majority of studies used the Framingham score as the base model. there is obviously a correspondence between the prediction models (for example, it has been computed that a 5\% SCORE risk of CVD death equates to a 10-25\% Framingham risk of total CVD\textsuperscript{96}) and we can reasonably assume that markers improving the predictive power of one model will do so for other models based on similar risk factors. Finally 2 over 5 cost-effectiveness studies focused specifically on a population of healthy, active duty army personnel aged 40-55.


59. Nambi V, Chambless L, He M, Folsom AR, Mosley T, Boerwinkle E, et al. Common carotid artery intima-media thickness is as good as carotid intima-media thickness of all carotid artery segments in improving prediction of coronary heart disease risk in the


74. Ijkema R, Van Aerde MA, Van Der Aalst CM, Van Ballegooijen M, Oudkerk M, De Koning HJ. A randomized controlled trial measuring the effectiveness of screening for cardiovascular disease using


